Solar Dynamics and Magnetism from the Interior to the Atmosphere

Nagi N. Mansour · Alexander G. Kosovichev ·
Rudolf Komm · Dana Longcope
Editors

Solar Dynamics and Magnetism from the Interior to the Atmosphere

Previously published in *Solar Physics* Volume 287,
Issues 1–2, 2013

Springer

Editors
Nagi N. Mansour
NASA Ames Research Center
Moffett Field, CA, USA

Rudolf Komm
National Solar Observatory
Tucson, AZ, USA

Alexander G. Kosovichev
Stanford University
Stanford, CA, USA

Dana Longcope
Montana State University
Bozeman, MT, USA

ISBN 978-1-4899-8004-5
ISBN 978-1-4899-8005-2 (eBook)
DOI 10.1007/978-1-4899-8005-2
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013953962

Cover illustration: NASA

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

DOI 10.1007/978-1-4899-8005-2_1
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-013-0377-6

Preface

Nagi N. Mansour · Alexander G. Kosovichev · Rudolf Komm · Dana Longcope · John W. Leibacher

Published online: 27 August 2013

This volume contains a collection of articles on the topic of *Solar Dynamics and Magnetism from the Interior to the Atmosphere* stimulated by the LWS/SDO-3/SOHO-26/GONG-2011 workshop of the same name, which was held 31 October – 4 November 2011, at Stanford, California, USA.

The goal of the workshop was to discuss recent advances and new problems in the exploration of the Sun's interior structure, solar dynamics and dynamo, mechanisms of sunspot and active regions formation, sources of solar irradiance variations, links between the subsurface dynamics, and flaring and CME activity. NASA's *Solar Dynamics Observatory*

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

N.N. Mansour (✉)
NASA Ames Research Center, NASA Advanced Supercomputing Division, Moffett Field, CA, USA
e-mail: nagi.n.mansour@nasa.gov

A.G. Kosovichev
Stanford University, Stanford, CA, USA
e-mail: sasha@sun.stanford.edu

R. Komm · J.W. Leibacher
National Solar Observatory, Tucson, AZ, USA

R. Komm
e-mail: rkomm@nso.edu

J.W. Leibacher
e-mail: john.leibacher@gmail.com

D. Longcope
Montana State University, Bozeman, MT, USA
e-mail: dana@solar.physics.montana.edu

J.W. Leibacher
Institut d'Astrophysique Spatial, Orsay, France

(SDO) mission has been providing a large amount of new data on solar dynamics and magnetic activities during the rising phase of the highly unusual Solar Cycle 24. These data are complemented by the *Solar and Heliospheric Observatory* (SOHO) mission and by ground-based observatories, which include the *Global Oscillation Network Group* (GONG) helioseismology network and the *New Solar Telescope* (NST). In addition, these observations are supported by realistic numerical simulations on high-end computers. This unprecedented amount of data provides a unique opportunity for multi-instrument investigations that address fundamental problems of the origin of solar magnetic activity at various spatial and temporal scales. The data are being used to develop new methods for forecasting solar cycles, emergence and evolution of active regions, and their flaring and CME activity.

This volume represents an attempt to develop a synergy among investigations of the solar interior by helioseismology, surface magnetism, and the atmospheric dynamics, by using both state-of-the-art observations and numerical simulations. The articles are organized in six chapters:

i) Local and Global Helioseismology.
ii) Sources of Solar Shape and Irradiance Variations.
iii) Large-Scale Dynamics, Magnetism, and Dynamo.
iv) Emerging Magnetic Flux and Subsurface Dynamics.
v) Formation, Structure, and Evolution of Sunspots and Active Regions.
vi) Magnetic Topology and Dynamics of the Solar Atmosphere.

Some articles fit in more than one topic, and in such cases the classification is not well-defined, and we recommend that readers consult the whole table of contents. Some of the articles are written as Invited Reviews, but all contain new results and ideas that reflect the current status of the field, new challenges, and future perspectives.

In recent years, our understanding of the solar interior has advanced substantially thanks to the development of helioseismology techniques and continuous observations of solar oscillations by the *Helioseismic and Magnetic Imager* (HMI) onboard SDO and by GONG. High-precision measurements of oscillation frequencies have provided the radial sound-speed profile and the distribution of the angular velocity through the whole interior, except perhaps the very inner core of the Sun. These measurements provided a test of stellar-evolution theory and the standard solar model. However, the discrepancy between the abundance of heavy elements determined spectroscopically on the solar surface and the abundance deduced from solar modeling and global helioseismology data provides a strong indication that our understanding of the basic physics of the solar interior is still incomplete. In particular, there are fundamental issues related to modeling of non-ideal properties of the solar plasma, and the equation of state, which directly affect the accuracy of the estimate for the heavy element abundance by helioseismology. In addition, there are indirect effects due to the uncertainty of the solar age and magnetic-field effects. Therefore, it is important to develop a synergy of more accurate frequency inversion methods and detailed description of the micro-physics in order to resolve the heavy element abundance problem. Also, it becomes increasingly important to investigate the global dynamics of the Sun. In addition to the traditional methods of global helioseismology, based on inversion of rotational frequency splitting, new methods of local helioseismology, which measure the meridional circulation and large-scale flows in the quiet-Sun and active regions, are being actively developed. In particular, the very long time series of solar oscillations observed by the *Michelson Doppler Imager* onboard the SOHO spacecraft and GONG during the whole solar cycle from 1995 – 2008, provided new measurements of the solar differential rotation down to 0.2 solar radii. However, while the results of these measurements are consistent with the hypothesis of a

uniformly rotating core, they cannot rule out that the energy-generating core rotates faster or slower than the outer radiative zone. This is also a problem of fundamental importance for understanding the formation and evolution of the Sun.

Local helioseismology based on measurements and inversion of frequency shifts and acoustic travel times is capable of providing the 3D structure and dynamics of the convection zone. The interpretation of these measurements and inversions is a subject of hot debate. The debate has been substantially improved with the development of realistic numerical simulations of randomly excited acoustic waves in 3D models of the Sun that include background flows and magnetic fields. The approach based on a synergy of a time–distance helioseismology method and numerical simulations provided new estimates of the subsurface flow structure of supergranulation, which led to a surprising result that strong upflows may be hidden 2 Mm beneath the surface supergranulation pattern dominated by outflows. In addition, local-helioseismology measurements based on unique multi-wavelength data from SDO provide new knowledge about wave propagation in magnetic regions of the solar atmosphere, and wave transformation and scattering. These new results challenge the existing simple models of MHD wave propagation and point to the need for further development of numerical simulations of waves in realistic conditions of turbulent and radiating plasma of magnetic regions. It becomes more and more clear that MHD waves excited by turbulent convection play a fundamental role in the dynamics and energetics of the solar atmosphere.

Variations of the solar diameter, shape, and irradiance are of particular interest because of the long and rich history of these investigations, their role in solar evolution, and global changes of the Earth's radiation and space environment. The high-precision measurements onboard SDO open new perspectives for improving our knowledge of solar variability on wide temporal and spatial scales, and advancing our understanding of the underlying physical mechanisms. In particular, these measurements encourage revisiting of old questions about solar oblateness, the relationship between oblateness changes and solar cycles, the constraints on relativistic celestial mechanics and alternative theories of gravitation. At the high-frequency end of the irradiance spectrum, accurate measurements from the *Extreme Ultraviolet Variability Experiment* (EVE) onboard SDO led to the detection of the global five-minute oscillations in the corona, and their modal structure, which corresponds to the global modes observed in the photosphere. If this surprising link between the interior and corona is confirmed by further observations, it will have important implications on theories of wave propagation on the Sun and mechanisms of coronal heating.

Helioseismology provides critical information for developing dynamo models of the solar magnetic cycle. After the discovery of the tachocline, a narrow rotational-shear layer at the bottom of the convection zone, dynamo theories assumed that the toroidal magnetic fields, which are the primary source of sunspot regions, are generated and stored in the tachocline. In order to explain the sunspot "butterfly" diagram, these theories assumed that the internal meridional circulation transports the toroidal field in the tachocline towards the Equator in the course of the solar cycle, and that this toroidal field emerges locally in the form of compact Ω-shaped magnetic loops and forms sunspot regions. However, there are serious, unsatisfactory aspects of such an advection-dominated dynamo because the strength of the return meridional flow is largely unknown, and because the required eddy diffusivity is about one order of magnitude greater than is predicted by standard mixing-length theory. One possible alternative to the "flux-transport" theory is a dynamo model that takes into account effects of the near-surface rotational-shear layer, also discovered by helioseismology. In the new theory, the magnetic field is generated in the bulk of the convection zone and forms the butterfly pattern in the subsurface-shear layer. This theory represents a paradigm shift in our understanding of the solar dynamo and cycles, and requires further detailed investigation. It is quite intriguing that after this workshop new local-helioseismology results

provided evidence that the meridional circulation may consist of two radial cells, making the flux-transport theory even more problematic. Ultimately, the issue will be resolved by improving the accuracy of the helioseismology measurements, and new approaches are being actively developed. In addition, important information is provided from synoptic analysis of magnetic patterns on the solar surface such as rotation of sunspot groups, and their inclination relative to the Equator (Joy's law). Future dynamo theories will have to explain these patterns.

Observations and modeling of the emergence of the dynamo-generated magnetic flux are of primary importance in heliophysics for two main reasons: First, the emerging flux is directly linked to the dynamo process and carries information about the dynamo; second, the emerging flux can trigger flares and CMEs, and thus is important for forecasting space weather and solar storms. Helioseismology made a substantial breakthrough in this topic by detecting acoustic travel-time variations associated with the emerging flux in the deep convection zone (at a depth of 45 – 70 Mm) 24 – 48 hours before the magnetic field becomes visible on the surface. The nature of the travel-time variations is not yet understood. These are probably related to scattering of acoustic waves on the emerging flux inhomogeneities. Solving this problem requires modeling of the flux emergence in realistic turbulent conditions of the convection zone. Simulations are being actively developed, and used for understanding links between the dynamo characteristics and properties of the emerging flux.

After emergence at the surface, the magnetic flux is a mixture of small elements of opposite polarity. However, very quickly, elements of each polarity merge and form compact and stable sunspot structures. After the formation, the sunspot magnetic fields continue to evolve, interact with magnetic fields of the opposite polarity, and form unstable magnetic configurations resulting in plasma eruptions. It seems that the magnetic-field dynamics observed on the surface is controlled by subsurface flows. Interaction of these flows, driven by the convective-energy flux, with magnetic fields is a key to understanding the solar activity, but the physics of this interaction is extremely complicated. Currently, only the first steps are being made by local time–distance helioseismology to map the sub-surface flow patterns and link these to the surface magnetic field dynamics. New results from the HMI instrument onboard SDO, obtained for a shallow 500 km deep region, which was previously inaccessible due to low spatial resolution of previous instruments, reveal converging plasma flows in the sunspot umbra area, surrounded by diverging flows. These measurements, supported by results of correlation tracking of the surface field, represent a significant challenge for some recent MHD simulations of sunspots, which predicted only diverging flow patterns. Another local-helioseismology technique, ring-diagram analysis, shows that the subsurface structure of sunspots is characterized by a shallow layer of reduced wave speed and a deeper layer of higher sound speed. Qualitatively, this is consistent with the previous time–distance helioseismology results. However, the depth estimates of these layers are significantly different. This difference may be related to differences in the spatial resolution, but ultimately will be resolved when large-scale realistic MHD simulations of sunspots will become available. So far, only relatively small, pore-like structures were modeled in the simulations self-consistently without specifying artificial boundary conditions to prevent the structure's decay. In these self-consistent simulations, the magnetic structures are maintained by converging downdrafts driven by surface cooling. An alternative model suggests that sunspot-like structure can form because of the suppression of turbulent pressure by the magnetic field. This suppression called "negative effective magnetic pressure instability" also results in converging downdrafts. In addition, the instability leads to a redistribution of turbulent intensity and gas pressure that could provide direct observational signatures.

The structure and dynamics of the atmosphere and corona is governed by magnetic fields generated by a dynamo in the deep convection zone and emerging on the solar

surface. Recent high-resolution observations from the NST of the Big Bear Solar Observatory led to a surprising discovery of ultra-fine structure of magnetic fields extending from the surface to the corona. Analysis of these ultra-fine magnetic fields suggest that they can serve as channels of the energy and mass flow from the interior to the corona and contribute to the coronal heating. The origin of the fine-scale structures and their dynamics are being investigated using new infrared and visible light spectro-polarimeters. The ultimate goal is to obtain high-resolution spectro-polarimetric data for diagnostics of solar dynamics and magnetism from the low photosphere to high chromosphere. Higher up in the corona, the magnetic fields cannot be measured by spectro-polarimetry. Therefore, the challenge is to infer the magnetic-field properties using EUV observations of the coronal magnetic loops from the SDO and the *Solar TErrestrial RElations Observatory* (STEREO) missions, and photospheric magnetograms. New algorithms of nonlinear force-free reconstruction have been developed and provide estimates of free magnetic energy, using only line-of-sight magnetic-field data. These are very important achievements for predicting flare and CME energy-release events.

In summary, recent high-resolution observations from ground and space, as well as realistic supercomputer simulations, have led to substantial progress in our understanding of solar dynamics and magnetism from the interior to the atmosphere. This collection of articles demonstrates that the synergy of high-resolution multi-wavelength observations and simulations is a key to uncovering long-standing puzzles of solar magnetism and dynamics.

Acknowledgements

The workshop was organized by Stanford University and the NASA Ames Research Center, with the primary support of NASA's Living With a Star Program.

Preface

Mansour, N.N., Kosovichev, A.G., Komm, R.W., Longcope, D.W., Leibacher, J.W.: 2013, Solar Dynamics and Magnetism Preface, *Solar Phys.* **287**, 1. doi:10.1007/s11207-013-0377-6.

Local and Global Helioseismology

Gough, D.: 2013, What Have We Learned from Helioseismology, What Have We Really Learned, and What Do We Aspire to Learn?, *Solar Phys.* **287**, 9. doi:10.1007/s11207-012-0099-1.

Eff-Darwich, A., Korzennik, S.G.: 2013, The Dynamics of the Solar Radiative Zone, *Solar Phys.* **287**, 43. doi:10.1007/s11207-012-0048-z.

Baldner, C.S., Basu, S., Bogart, R.S., Burtseva, O., González Hernández, I., Haber, D., Hill, F., Howe, R., Jain, K., Komm, R.W., Rabello-Soares, M.C., Tripathy, S.: 2013, Latest Results Found with Ring-Diagram Analysis, *Solar Phys.* **287**, 57. doi:10.1007/s11207-012-0171-x.

Duvall, T.L. Jr., Hanasoge, S.M.: 2013, Subsurface Supergranular Vertical Flows as Measured Using Large Distance Separations in Time-Distance Helioseismology, *Solar Phys.* **287**, 71. doi:10.1007/s11207-012-0010-0.

Komm, R., González Hernández, I., Hill, F., Bogart, R., Rabello-Soares, M.C., Haber, D.: 2013, Subsurface Meridional Flow from HMI Using the Ring-Diagram Pipeline, *Solar Phys.* **287**, 85. doi:10.1007/s11207-012-0073-y.

Rajaguru, S.P., Couvidat, S., Sun, X., Hayashi, K., Schunker, H.: 2013, Properties of High-Frequency Wave Power Halos Around Active Regions: An Analysis of Multi-height Data from HMI and AIA Onboard SDO, *Solar Phys.* **287**, 107. doi:10.1007/s11207-012-0180-9.

Woodard, M., Schou, J., Birch, A.C., Larson, T.P.: 2013, Global-Oscillation Eigenfunction Measurements of Solar Meridional Flow, *Solar Phys.* **287**, 129. doi:10.1007/s11207-012-0075-9.

Zhao, J., Chou, D.-Y.: 2013, Analysis of the Helioseismic Power-Spectrum Diagram of a Sunspot, *Solar Phys.* **287**, 149. doi:10.1007/s11207-012-0182-7.

Sources of Solar Shape and Irradiance Variations

Rozelot, J.P., Fazel, Z.: 2013, Revisiting the Solar Oblateness: Is Relevant Astrophysics Possible?, *Solar Phys.* **287**, 161. doi:10.1007/s11207-013-0245-4.

Didkovsky, L., Kosovichev, A., Judge, D., Wieman, S., Woods, T.: 2013, Variability of Solar Five-Minute Oscillations in the Corona as Observed by the *Extreme Ultraviolet Spectrophotometer* (ESP) on the *Solar Dynamics Observatory/Extreme Ultraviolet Variability Experiment* (SDO/EVE), *Solar Phys.* **287**, 171. doi:10.1007/s11207-012-0186-3.

Large-Scale Dynamics, Magnetism, and Dynamo

Bonanno, A.: 2013, Solar Dynamo and Toroidal Field Instabilities, *Solar Phys.* **287**, 185. doi:10.1007/s11207-012-0198-z.

Javaraiah, J.: 2013, A Comparison of Solar Cycle Variations in the Equatorial Rotation Rates of the Sun's Subsurface, Surface, Corona, and Sunspot Groups, *Solar Phys.* **287**, 197. doi:10.1007/s11207-013-0345-1.

McClintock, B., Norton, A.A.: 2013, Recovering Joy's Law as a Function of Solar Cycle, Hemisphere, and Longitude, *Solar Phys.* **287**, 215. doi:10.1007/s11207-013-0338-0.

Emerging Magnetic Flux and Subsurface Dynamics

Kholikov, S.: 2013, A Search for Helioseismic Signature of Emerging Active Regions, *Solar Phys.* **287**, 227. doi:10.1007/s11207-013-0321-9.

Weber, M.A., Fan, Y., Miesch, M.S.: 2013, Comparing Simulations of Rising Flux Tubes Through the Solar Convection Zone with Observations of Solar Active Regions: Constraining the Dynamo Field Strength, *Solar Phys.* **287**, 237. doi:10.1007/s11207-012-0093-7.

Formation, Structure, and Evolution of Sunspots and Active Regions

Baldner, C.S., Bogart, R.S., Basu, S.: 2013, The Sub-surface Structure of a Large Sample of Active Regions, *Solar Phys.* **287**, 263. doi:10.1007/s11207-012-0148-9.

Liu, Y., Zhao, J., Schuck, P.W.: 2013, Horizontal Flows in the Photosphere and Subphotosphere of Two Active Regions, *Solar Phys.* **287**, 277. doi:10.1007/s11207-012-0089-3.

Kemel, K., Brandenburg, A., Kleeorin, N., Mitra, D., Rogachevskii, I.: 2013, Active Region Formation through the Negative Effective Magnetic Pressure Instability, *Solar Phys.* **287**, 291. doi:10.1007/s11207-012-0031-8.

Magnetic Topology and Dynamics of the Solar Atmosphere

Goode, P.R., Cao, W.: 2013, Infrared Observations from the *New Solar Telescope* at Big Bear, *Solar Phys.* **287**, 313. doi:10.1007/s11207-013-0235-6.

Aschwanden, M.J.: 2013, A Nonlinear Force-Free Magnetic Field Approximation Suitable for Fast Forward-Fitting to Coronal Loops. I. Theory, *Solar Phys.* **287**, 321. doi:10.1007/s11207-012-0069-7.

Aschwanden, M.J., Malanushenko, A.: 2013, A Nonlinear Force-Free Magnetic Field Approximation Suitable for Fast Forward-Fitting to Coronal Loops. II. Numeric Code and Tests, *Solar Phys.* **287**, 343. doi:10.1007/s11207-012-0070-1.

Aschwanden, M.J.: 2013, A Nonlinear Force-Free Magnetic Field Approximation Suitable for Fast Forward-Fitting to Coronal Loops. III. The Free Energy, *Solar Phys.* **287**, 367. doi:10.1007/s11207-012-0203-6.

Panasenco, O., Martin, S.F., Velli, M., Vourlidas, A.: 2013, Origins of Rolling, Twisting, and Non-radial Propagation of Eruptive Solar Events, *Solar Phys.* **287**, 389. doi:10.1007/s11207-012-0194-3.

Petrie, G.J.D.: 2013, A Spatio-temporal Description of the Abrupt Changes in the Photospheric Magnetic and Lorentz-Force Vectors During the 15 February 2011 X2.2 Flare, *Solar Phys.* **287**, 413. doi:10.1007/s11207-012-0071-0.

Temmer, M., Vrsnak, B., Veronig, A.M.: 2013, The Wave-Driver System of the Off-Disk Coronal Wave of 17 January 2010, *Solar Phys.* **287**, 439. doi:10.1007/s11207-012-0088-4.

DOI 10.1007/978-1-4899-8005-2_2
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0099-1

What Have We Learned from Helioseismology, What Have We Really Learned, and What Do We Aspire to Learn?

Douglas Gough

Received: 9 February 2012 / Accepted: 3 August 2012 / Published online: 27 September 2012

Abstract Helioseismology has been widely acclaimed as having been a great success: it appears to have answered nearly all the questions that we originally asked, some with unexpectedly high precision. We have learned how the sound speed and matter density vary throughout almost all of the solar interior – something which not so very long ago was generally considered to be impossible – we have learned how the Sun rotates, and we have a beautiful picture, on a coffee cup, of the thermal stratification of a sunspot, and also an indication of the material flow around it. We have tried, with some success at times, to apply our findings to issues of broader relevance: the test of the General Theory of Relativity via planetary orbit precession (now almost forgotten because the issue has convincingly been closed, albeit no doubt temporarily) the solar neutrino problem, the manner of the transport of energy from the centre to the surface of the Sun, the mechanisms of angular-momentum redistribution, and the workings of the solar dynamo. The first two were of general interest to the broad scientific community beyond astronomy, and were, quite rightly, principally responsible for our acclaimed success; the others are still in a state of flux.

Keywords Helioseismology, heliophysics · Solar neutrinos · General relativity · Solar opacity · Equation of state

Invited Article.

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

D. Gough (✉)
Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK
e-mail: douglas@ast.cam.ac.uk

D. Gough
Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, Wilberforce Road, Cambridge, CB3 0WA, UK

D. Gough
Physics Department, Stanford University, Stanford, CA 94305, USA

1. Prelude

In the early heady days of helioseismology, the new techniques of inference, coupled with pertinent observations that had been stimulated by them, went hand-in-hand with consequent scientific discovery, and helioseismology was not unnaturally perceived to be almost a branch of science. Now, in its relative maturity, it is, or at least it should be, relegated to what it really is: a very valuable technique for drawing scientific inference. The basic physics of seismic oscillations is not new, is fundamentally quite simple, and is well understood; and the principles of inference from the observations should be straightforward to comprehend, even though the technicalities of putting them into practice may for some seem to be rather complicated. Therefore the robust raw conclusions are genuinely secure, more so than the broader issues to which they are intended to be applied. However, it is incumbent upon us to make the distinction between the inferences that really have been drawn reliably and the further wider inferences that might subsequently be, or have been, drawn, often with the aid of supplementary, possibly less secure, maybe non-seismic, information, and even, maybe, (sometimes unstated) surmise. Only if such distinction is made clear can the contributions of our subject to science be reaped to the full. Unfortunately, that aspiration has not always been achieved in the past, and misinformation has sometimes sullied the waters.

It takes only a brief scrutiny of the equations describing the structure and dynamical evolution of the Sun (it is not quite so brief to derive them) and the equations governing the low-amplitude seismic modes of oscillation to appreciate what broadly can, at least in principle, be reliably inferred. Anything further must depend on other criteria, such as general physical argument beyond seismology, traditional astronomical observation, or even prejudice. It is obligatory to be explicit about how such additional constraints are applied. The subject has advanced to a new level of sophistication; we are now trying to probe seismically (and otherwise) almost inaccessible aspects of the physics of the Sun, and the techniques for unravelling them are becoming more and more intricate, beyond the point at which most scientists wish to tread. There must necessarily be an increased trust in our findings, and it is our responsibility not to betray it. Many of the broader scientific community want to use our results in their research; for that they need to know not only the limitations of our inferences, and the caveats upon which they are based, but also which aspects of what we seismologists tell them can really be trusted.

Much of the emphasis of *Solar Dynamics Observatory* seismology concerns the workings of the convection zone. We want to know what controls the solar cycle, how magnetic field is amplified, modulated, and then suppressed, how sunspots are formed and destroyed – and what determines their lifespan. We want to know the geometry of at least the larger scales of convective motion, and how, beneath the seen superficial layers of the Sun, the processes that control the total radiative output are modulated. At least some of us want to understand how all these matters influence our procedures for inferring the gross properties of the Sun, and how they impinge on our broader ideas of the evolution of the Sun in particular, and of stars in general. Addressing such delicate issues with confidence may now seem an almost impossible task to us older scientists who have lived through the years of stumbling through the darkness, having finally emerged to bathe in the secure light illuminating the minute arena of knowledge that we have been instrumental in uncovering. It is now up to the younger community to proceed likewise: to grasp at the edge of our perception with initially insecure ideas, fully appreciating the uncertainty, of course; then moulding and strengthening them into a new body of secure scientific knowledge.

2. Introduction

Once the potential of solar oscillations to map the interior of the Sun was recognised (Christensen-Dalsgaard and Gough, 1976) there were two obvious serious issues that im-

mediately appeared accessible to resolution: the spherically symmetric component of the hydrostatic stratification and the internal angular velocity. The first of these was needed for investigating what has been called the solar neutrino problem; the second concerned the centrifugally induced oblateness of the Sun's gravitational equipotentials, and how that impinged upon an important test of theories of gravity, General Relativity in particular, via the precession of planetary orbits. These were the two most widely discussed issues in heliophysics at the time.

The stratification was originally addressed with the help of theoretical models. Deubner (1975) had published the first well resolved k–ω spectrum, and Ando and Osaki (1975) had shown that theoretical eigenfrequencies of p modes trapped in the outer layers of a solar model envelope were in quite good, but not perfect, agreement with Deubner's observations. What was required for bringing the theory more closely into line with observation was first estimated from the properties of the eigenfrequencies of a simple polytropic representation of the outer layers of the convection zone as a means of calibrating solar models (Gough, 1977); the conclusion was that the convection zone must be about 200 Mm deep, some 50 Mm deeper than the favoured value of the time. Basically, that conclusion was drawn from relating theoretical eigenfrequencies to the jump in the adiabatic "constant" p/ρ^{γ_1}, which is closely related to specific entropy, across the thin superadiabatic convective boundary layer. It was subsequently supported by more realistic, numerical, computations by Ulrich and Rhodes (1977). The principal implication of that result was that, according to complete solar models, a deeper convection zone implied greater helium and heavy-element abundances, a hotter, more centrally condensed, core, and a higher neutrino flux, thereby exacerbating the solar neutrino problem (*e.g.* Abraham and Iben, 1971; Bahcall and Ulrich, 1971). Indeed, one is tempted to speculate that the modellers in the past had adjusted the defining parameters of their models to minimise the theoretical neutrino flux, and that had prejudiced Ando and Osaki's calculations, although without repeating the calculations oneself (and maybe even if one did) one cannot be sure.

There was a great deal of healthy mistrust in the models at that time. The concern was that a seismological measurement – I should really say estimate – of merely the upper boundary layer of the convection zone, extending only a minute fraction of the solar radius beneath the photosphere, could hardly be a robust indicator of conditions inside the energy-generating core. Therefore, as soon as low-degree data became available (Claverie *et al.*, 1980; Grec, Fossat, and Pomerantz, 1980; Fossat, Grec, and Pomerantz, 1981) it became possible to use seismic indicators of more global properties; first the so-called large frequency separation (Christensen-Dalsgaard, Gough, and Morgan, 1979a, 1979b; Christensen-Dalsgaard and Gough, 1980), which measures the sound travel time from the centre of the Sun to the seismic surface[1] (*e.g.* Vandakurov, 1967; Tassoul, 1980; Gough, 1986), and later the small

[1] I consider the seismic surface $r = R$ of the Sun (assumed here to be spherically symmetrical) to be the radius at which c^2, regarded as a function of r, or c, regarded as a function of acoustic radius $\tau(r) = \int c^{-1}\,\mathrm{d}r$ – both of which are close to being linear functions in the outer adiabatically stratified layers of the convection zone (Balmforth and Gough, 1990; Lopes and Gough, 2001) – extrapolate to zero. In the Sun, according to Model S of Christensen-Dalsgaard *et al.* (1996), it lies about 1000 km above the photosphere, the precise value depending on exactly how the extrapolation is carried out. There is nothing special about the structure of the actual atmosphere in its vicinity, which is well inside the outer evanescent zone of most of the seismic modes and therefore has little significant influence on the dynamics. Instead, it acts simply as a (virtual) singularity in the acoustic wave equation, providing a convenient parametrisation of conditions (well below the photosphere) in the vicinity of the upper turning points of the modes. Put another way, it provides a convenient fiducial location with respect to which the acoustic phase in the propagating zone beneath is related. Unlike the photosphere, which has no acoustic significance, it shares a relation with the deeper solar interior

frequency separation (Gough, 1983c), which is a direct indicator of conditions in the core (*e.g.* Gough, 1983a, 1986). Unfortunately it was not possible to make all of the theoretical frequencies agree with the data, as was clear from an indiscriminant attempt to fit only the low-degree frequencies with whole-disc observations (Christensen-Dalsgaard and Gough, 1981): two solar models seemed to be favoured, one with a low initial helium abundance: $Y_0 = 0.18$ (coupled with a correspondingly low heavy-element abundance $Z = 0.003$ and a low neutrino flux, although not low enough to reproduce the neutrino detection rate), the other large: $Y_0 = 0.27$ and $Z_0 = 0.026$. The high-Y model fitted rather better, and one was tempted to prefer it, especially because its high-degree mode frequencies were closer to Deubner's observations. Moreover, the low-Y model had a helium abundance below what was thought to have been produced in the Big Bang (*cf.* footnote 8), which would call for some contrived explaining.

The two models required different identifications of the orders $[n]$ of the modes; unfortunately the orders were so high, and the fit so poor, that reliable extrapolation to $n = 1$ was not possible. It was not until Duvall and Harvey (1983) observed the frequencies of modes of intermediate degree that a secure connection between modes of low and high degree could be made (Gough, 1983b). The orders of the latter are determinable because the frequencies of high-degree f modes, with which is associated $n = 0$, are essentially independent of the structure of the Sun (Gough, 1982a). The high-Y alternative was thereby confirmed.

Thus we had learned that the solar neutrino problem was almost certainly not resolvable by adjusting solar models, and must be a matter for nuclear or particle physics.

Had we really learned that? Perhaps not yet. At the time there were still very serious doubts about the solar models, for they depended on many unproven assumptions, some of which are listed in Table 1. There had already been, and there were yet to be, many models computed in which some of these assumptions were relaxed in the hope of yielding lower neutrino fluxes (although none provided a satisfactory reconciliation of theory with observation). At the very least, a secure representation of the stratification throughout the Sun was surely required. That was soon to be provided from inversion analyses of frequency data from Duvall and Harvey (Christensen-Dalsgaard *et al.*, 1985; Christensen-Dalsgaard, Gough, and Thompson, 1989; Dziembowski, Pamyatnykh, and Sienkiewicz, 1990), from which it was possible to infer the (spherically averaged) sound speed $[c(r)]$ throughout almost all of the Sun. Details of the core were not yet within reach (and even today there is considerable uncertainty), but elsewhere the sound speed was essentially the same as that of Christensen-Dalsgaard and Gough's (1981) high-Y model. It was also possible to see the base of the convection zone. I recall one Friday morning (the second Friday of January 1984, a day on which I was due to deliver a lecture on my findings to the Royal Astronomical Society); it was about 4:30 in the morning – one could compute seriously only at night in those days – when I obtained my first plot of $c^2(r)$ that extended beneath the convection zone; I plotted c^2 and not c because it is related closely to temperature T – the equation of state in the solar interior is reasonably well approximated by the perfect-gas law, for which $c^2 \propto T/\mu$, where μ is the "mean molecular mass". A more modern (and rather

that is robust, and is insensitive to the non-seismic, thermal and radiative, properties of the outer convective boundary layer, whose structure changes with the solar cycle (Antia and Basu, 2004; Dziembowski and Goode, 2004, 2005; Lefebvre and Kosovichev, 2005; Lefebvre, Kosovichev, and Rozelot, 2007). In contrast to other, non-seismic, radii (*cf.* Bahcall and Ulrich, 1988), it provides a stable outer limit to the effective total acoustic-radius integral $\tau(R)$, which determines the large frequency separation; in Model S it is some 200 seconds or so greater than the actual acoustic radius of the photosphere.

Table 1 Standard model assumptions

Initially uniform chemical composition
Spherical symmetry
Hydrostatic equilibrium
Simple description of energy transport, using mixing-length theory in convection zones, and a simplified treatment of radiative transfer in the atmosphere
No serious internal mixing of chemical species except in convection zones, so no tachocline, yet gravitational settling and chemical diffusion
Rotation dynamically negligible
Maxwell stresses negligible (*i.e.* no magnetism)
No mass loss
No accretion
No large-scale instability
No (nonlinear) transport by waves
Thermal balance (almost)
Equation of state known
Nuclear reaction rates (reaction cross-sections) known
Opacity formula known

better) version of what I obtained is illustrated in Figure 1.[2] The location of the base of the convection zone is evident as a near discontinuity of the second derivative, as one can see more easily by holding the page almost in one's line of sight and looking along the curve. I confirmed the location by having a research student, the only person left in the computer room at that hour, repeat the exercise. I went home, slept a couple of hours, and then took a train to London. When I showed the plot at the RAS (Gough, 1984a) and pointed out the discontinuity, the audience was incredulous, even though I stood almost in the plane of the screen to reassure them. I then recounted the confirmation by the research student, and there was a sudden release of tension in the audience when I told them that the student was a cosmologist, for then they appreciated that such a student must undoubtedly have provided an unbiased opinion on such a matter.

The result was compared with theoretical solar models (Christensen-Dalsgaard *et al.*, 1985), and revealed a characteristic discrepancy (much larger than that evident in the more modern comparisons, such as those illustrated in Figures 3 and 4) immediately beneath the base of the convection zone. It was suggested that the discrepancy could have been caused

[2]Included for comparison is one of the models with low Y that was used in the original calibration with low-degree modes by Christensen-Dalsgaard and Gough (1981). The qualitative differences between the models can be appreciated by realising first that the radiative envelopes are roughly polytropic (with index 3.5), and that it is adequate to approximate the equation of state by the perfect-gas law. Then it is clear that the magnitude of $T(r)$ (and ρ) must be greater in the higher-Y model, because the total energy generation rate, which is an increasing function of X, ρ, and T, is the same for the two models. Polytropic scaling (*e.g.* Gough, 1990) indicates that Z is a steeply decreasing function of X, so Z is much greater in the higher-Y model. In addition, the polytropic radius scale is greater for the higher-Y model, as is evident from Figure 1 by imagining an extrapolation of the functional form of $T(r)$ outwards from the radiative zone. Consequently the convection zone, which steepens the gradient, has more truncating from the radiative structure to perform in order to maintain the observed photospheric radius, and is therefore deeper. An additional scaling in magnitude is required to convert T to $c^2 \propto T/\mu$, raising the dotted curve (corresponding to the model with the lower Y) relative to the continuous curve. It also depresses both curves near the centre of the star, in the energy-generating core where μ has been augmented by nuclear transmutation, providing a diagnostic of main-sequence age.

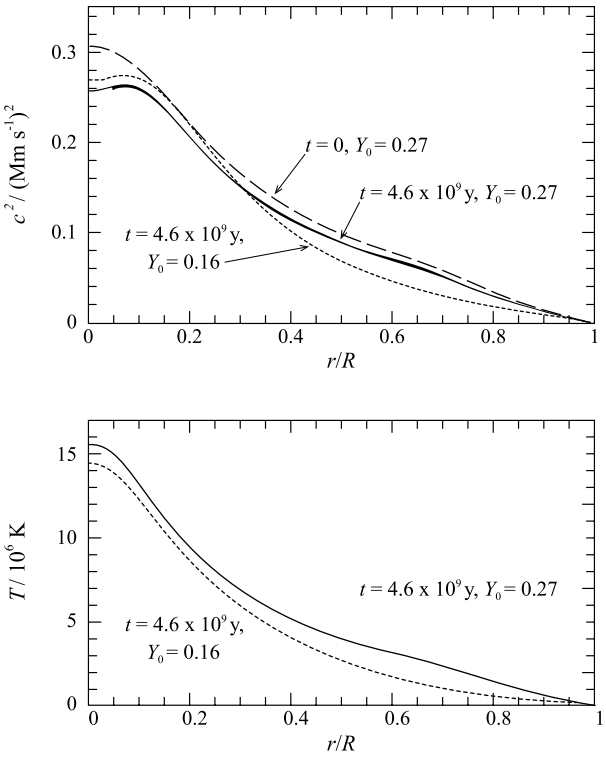


Figure 1 Upper panel: The continuous curve extending over the entire range of r/R is the square of the sound speed [c^2] in the standard solar model S of Christensen-Dalsgaard *et al.* (1996), which was computed with initial hydrogen and heavy-element abundances $X_0 = 0.7091$, $Z_0 = 0.0196$. The square of the sound speed in the Sun is also plotted (where $r/R > 0.1$) as a continuous curve; the difference between the two can be barely discerned by the variation in the thickness of the apparently single curve (but see Figures 3 and 4). The dip in c^2 at the centre of the Sun is a result of the augmentation of μ by nuclear transmutation, which increases with time on the main sequence (and which can therefore be used as a diagnostic of main-sequence age). The dashed curve is c^2 on the zero-age main sequence, and, except near the surface where abundant elements undergo ionisation, is (locally) proportional to temperature. The dotted curve is c^2 in a model with an initial heavy-element abundance $Z = 0.001$, but continuously contaminated at the surface at such a constant rate as to have a current heavy-element abundance $Z_S = 0.02$ in the convection zone today (Christensen-Dalsgaard, Gough, and Morgan, 1979a). Lower panel: Temperature [T] through the two present-day solar models.

by an error in the opacity (values of opacity in those days were available in tables provided by the Los Alamos National Laboratory in the USA). Such suggestions, in a rather broad sense, had been made in the past, for example by Simon (1982) in response to a failure to reproduce simultaneously observations relating to stellar evolution and observations of pulsational characteristics of classical variable stars. But here one was able to state quite precisely the thermodynamical conditions under which the opacity was in error, and the approximate magnitude (and sign) of the error. After some persuasion, Carlos Iglesias and Forrest Rogers at Livermore computed by moonlight the opacity for a few judiciously chosen values of the state variables ρ and T (and a plausibly appropriate chemical composition) using an independently written computer programme; they not only confirmed the helioseismic inference but also showed that the discrepancy was even greater at lower temperatures – of no concern

for the Sun because it is well inside the adiabatically stratified region of the convection zone where radiative energy transfer is negligible, but of great significance in reconciling with observation the theory of some classes of intrinsically variable stars such as β Cephei and slowly pulsating B stars (Cox *et al.*, 1992; Kiriakidis, El Eid, and Glatzel, 1992; Moskalik and Dziembowski, 1992; Dziembowski and Pamyatnykh, 1993; Dziembowski, Moskalik, and Pamyatnykh, 1993, 1994) and double-mode Cepheids (Moskalik, Buchler, and Marom, 1992). Largely as a result of efforts by Werner Däppen, scientists who were involved with the Los Alamos opacity computations were brought into the same room as the Livermore scientists, and errors in the Los Alamos calculations were identified. This led to the Lawrence Livermore National Laboratory providing the funds for more extensive opacity computations for astrophysical use (Iglesias, Rogers, and Wilson, 1990, Iglesias and Rogers, 1991, 1996, Rogers and Nayfonov, 2002). Thus we see the first example of helioseismology contributing directly to microscopic physics (if one considers the neutrino issue to be indirect). There have been further contributions, but I postpone discussion of those until later.

I emphasise that an important consequence of the sound-speed inversion was that it convincingly ruled out the low-Y models, reaffirming that the cause of the neutrino deficit was not in solar modelling. That conclusion was not fully appreciated by many of the solar modellers of the time, for they had not yet understood the power of helioseismological analysis.

I turn now to the angular velocity $[\Omega]$, which is measured by the odd (with respect to azimuthal order $[m]$) component of degeneracy splitting caused by advection and Coriolis acceleration. Strictly speaking, one needs the sound-speed inversion first, to establish the hydrostatic stratification with respect to which the splitting kernels are computed. However, the angular velocity was actually first obtained by using a theoretical solar model before the full sound-speed inversion had been carried out (Duvall *et al.*, 1984), partly, perhaps, because it was easier to perform with the data in hand, yet also with the knowledge that the error from which that procedure suffered was much less than that resulting from the errors in the measurements of the degeneracy splitting. The result was a big surprise.

There had been many prior discussions of how much faster than the photosphere the solar core must be rotating (due to spin-down), the most widely publicised being the discussions resulting from the measurements of the apparent oblateness of the solar surface by Dicke and Goldenberg (1967, 1974). The oblateness measurements had been made in the hope of confirming Dicke's theory of gravitation (Brans and Dicke, 1961)[3]: put naively, in the

[3]The gravitational attraction associated with the energy density $[-GM/r]$ of the gravitational field surrounding the Sun, absent in Newton's theory, causes the total gravitational attraction to increase: very roughly speaking, as a result of energy conservation the apparent gravitational mass of a planet at distance r from the Sun, in a flat representation of space, is augmented by approximately GM/c^2r per unit mass of planet above what it would have appeared to have been at infinity; M is the mass of the Sun, and here c is the speed of light. Similarly, the energy, hence the frequency, of a photon is multiplied by a factor $\Gamma = 1 + GM/c^2r$ – that causes the familiar gravitational redshift. Consequently, the orbit equation is modified simply by multiplying the Newtonian gravitational force on the planet by Γ^3. After linearisation and rewriting M in terms of the orbital specific angular momentum $h = \sqrt{(GMr)}$, valid for nearly circular orbits, the effective attractive force becomes $-(1 + 3h^2/c^2r^2)GM/r^2$; it increases with increasing proximity more rapidly than Newton's inverse square. It is easy to see also that the gravitational field in the equatorial plane of a rotating (oblate) axisymmetric self-gravitating body (like the Sun) also increases with decreasing distance faster than the field around a corresponding spherically symmetrical body: the act of flattening a spherical body takes equal amounts of material from the poles towards the near and the far sides of the Equator, the increase in gravitational attraction by closer nearside matter exceeding the lesser decrease by farside matter. Therefore the net gravitational attraction is increased, by an amount which increases as r decreases and the shape of the Sun becomes more apparent. The force is given approximately by $-(1 + \frac{3}{2}J_2R^2/r^2)GM/r^2$; J_2 is the quadrupole moment. In both cases, therefore, a planet is drawn towards the Sun more strongly with

Brans–Dicke theory Newton's gravitational constant [G] was regarded as a field satisfying a wave equation that couples to the matter, thereby relaxing the gravitational force field from the relative rigidity that is imposed when G is held constant, and hence reducing the rate of precession of planetary orbits from the value predicted by General Relativity; the fact that General Relativity predicts essentially the correct precession with a spherical Sun was regarded by Dicke as a fortuitous coincidence, and that in reality the shortfall predicted by his theory (which is not an absolute prediction, but depends on an unknown coupling constant) is made up by the oblateness of the Sun's gravitational field produced by a rapidly rotating interior. The oblateness measurements would calibrate the coupling constant. A debate ensued concerning two matters: the relation between the oblateness in surface brightness and the oblateness of the gravitational equipotentials (*e.g.* Dicke, 1970), and the fluid dynamics of the solar interior concerning angular-momentum transfer from centre to surface during spin-down (*e.g.* Howard, Moore, and Spiegel, 1967; Dicke, 1967; Bretherton and Spiegel, 1968). I shall enlarge on neither here, despite the intrinsic interest of each, because even the early seismological analysis (Duvall *et al.*, 1984) revealed that the Sun is unexpectedly rotating almost uniformly throughout its interior, except possibly in what might be its almost inaccessible core (Figure 2). Subsequent observations of low-degree modes revealed that the core rotates significantly no more rapidly than the envelope (Elsworth *et al.*, 1995; Chaplin *et al.*, 1999). From knowledge of Ω and the density and pressure stratification, the multipole moments [J_{2k}] of the gravitational equipotentials can easily be computed (*e.g.* Gough, 1981; Pijpers, 1998). I maintain that this is the most accurate way to determine J_{2k}, notwithstanding some claims in the literature to the contrary. The reason is partly that other methods involve relating surface brightness asphericity with gravity, which is not well understood, and partly because those measures rely on only the small quadratic centrifugal force rather than the more robust, and unambiguous, seismic effects of advection and, to a much lesser degree, Coriolis acceleration, both of which are predominantly linear in Ω and are therefore much larger than the centrifugal term; and, moreover, they do not depend on relating brightness to gravity. Moreover, the rotational distortion of the isobaric surfaces in the vicinity of the photosphere is dominated by the direct effect of the local centrifugal force, and exceeds the gravitational distortion by a factor of more than 20; determining the latter from the apparent shape of the solar disc therefore requires a subtraction of two measurements that differ by less than a mere 5 per cent. Even the most precise limb-shape measurements (*e.g.* Fivian *et al.*, 2008; Kuhn *et al.*, 2012) are far from that goal (Gough, 2012c).

The outcome is consistent with General Relativity; indeed, the miniscule contribution that J_2 does make to the precession of the orbit of Mercury (*e.g.* Gough, 1982b; Duvall *et al.*, 1984; Brown *et al.*, 1989; Antia, Chitre, and Gough, 2008) brought General Relativity into somewhat closer agreement with observation than assuming the Sun to be spherically symmetrical. The influence of higher moments is utterly negligible. The improvement is marginal, but with future, more delicate, orbital analyses the helioseismological determination of J_2 will evidently take on a crucial role. The almost uniform rotation of the radiative

increasing proximity than it would have been in an inverse-square field. Conserving its angular momentum, it is thereby caused to rotate through a greater angle near perihelion because its orbital angular velocity is augmented, distorting an otherwise approximately elliptical bound Newtonian orbit in such a manner as to appear to make it simply precess in the same direction as the angular velocity of the planet. (Near aphelion the oppositely directed contribution to the precession is lesser, therefore too small to annul the contribution from near perihelion.) Use of planetary (or spacecraft) orbital precession rates to calibrate the relativistically induced deviation from the inverse-square gravitational field surrounding the Sun therefore requires one to know the contribution from the distortion from spherical symmetry of the mass distribution in the Sun, such as is produced by centrifugal acceleration due to rotation. The precession rate is most easily calculated by perturbation theory (*e.g.* Ramsey, 1937).

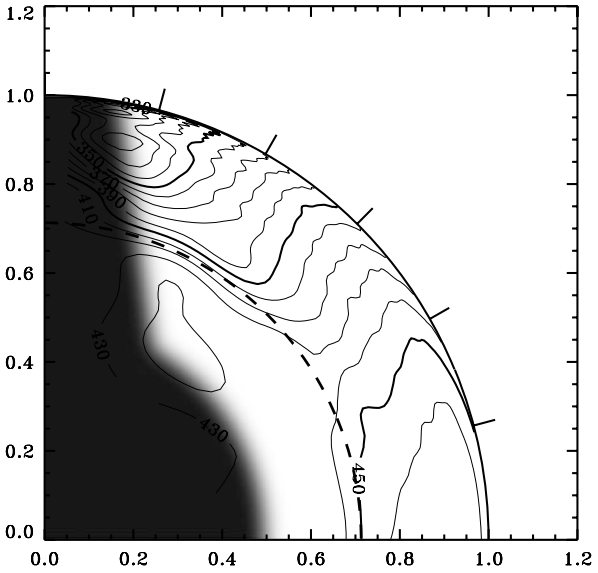


Figure 2 Optimally localised averages of the North – South symmetric component of the angular velocity [Ω] of the Sun, depicted as a contour plot in a quadrant. Some contours are labelled [nHz], and, for clarity, every fifth contour is drawn bold; the contour separation is 10 nHz. The outer quarter circle denotes the surface of the Sun, the dashed quarter circle indicates the base of convection zone, and the tick marks at the Sun's surface are drawn at latitudes 15, 30, 45, 60, and 75°. The Equator is the horizontal (relative to the page) axis and the Pole the vertical, each labelled with values of $x = r/R$. The shaded area indicates the region in the Sun where no reliable inference could be made from the data available (from Schou *et al.*, 1998). There has been little significant improvement of this inference, save for a demonstration that the core is not rotating rapidly (Elsworth *et al.*, 1995; Chaplin *et al.*, 1999); however, there have been studies of temporal variation associated with the solar cycle (*e.g.* Vorontsov *et al.*, 2002; Basu and Antia, 2003; Howe *et al.*, 2006a, 2006b; Antia, Chitre, and Gough, 2008).

interior of the Sun begs the question of how that can be, particularly because it interfaces with a differentially rotating convection zone. I shall return to that matter later.

3. Understanding Seismic Variables

It is extremely important to understand what is actually inferred from seismology, for only then can one draw reliable conclusions concerning the Sun. Seismic modes are essentially adiabatic. They result from forces – pressure and gravity predominantly – acting on matter with inertia. The dynamics therefore concerns only pressure [p] and matter density [ρ] which are related via gravity through hydrostatic equilibrium, and the relation between perturbations to them under adiabatic change, which is characterised by the first adiabatic exponent [$\gamma_1 = (\partial \ln p/\partial \ln \rho)_s$, the partial thermodynamic derivative being taken at constant specific entropy [s]]. The restoring force of acoustic modes (p modes) is principally pressure, that of gravity modes (g modes, including the fundamental f modes) is buoyancy. So the modes can provide information directly about only p, ρ, and γ_1 (and, of course, any function of them). These are the basic seismic variables. I should acknowledge that any magnetic field [**B**] that is present also contributes to the dynamics, and therefore is also a seismic variable. Unfortunately, its effect on the frequencies of the modes appears to be

indistinguishable from that of an appropriate variation in sound speed [c] – also a seismic variable because $c^2 = \gamma_1 p/\rho$ – which makes it difficult to unravel the two (Zweibel and Gough, 1995). In principle one might be able to do so from the eigenfunctions, which are configured differently by c and **B**, although supplementation by non-seismic arguments is likely to be more productive.

Leaving **B** aside for the time being, it should be appreciated that the relation between p and ρ through hydrostatics does not depend directly on γ_1, so p or ρ (or any function of only them) and γ_1 are structurally independent: it follows that although the relation between them, which physically is given by the equation of state, can be determined by seismology along the thermodynamic p–ρ path through the Sun, there is no remaining redundancy, so the veracity of the equation of state itself cannot therefore be probed by seismology alone; in order to investigate the equation of state, supplementary, non-seismic, information is required.

There are two approaches that one can take for drawing seismic inference. One is to adopt a parametrised model of the Sun, or some aspect of it, and from it calculate whatever seismic properties one wishes to compare with observation. The comparison calibrates the controlling parameters. That was the procedure that I described in my introduction for first estimating the depth of the convection zone, and consequently the helium abundance. The other approach is to ignore explicit models (almost) entirely, and combine the data in such a manner as to isolate certain properties of the seismic structure. Typically that involves adopting, at least at first, a reference model of the Sun, which it is hoped is sufficiently close to the Sun for linearisation in the small differences from it to be more-or-less valid; that simplifies the analysis enormously. Subsequent iteration can usually remove the dependence on the reference model. Nonlinear asymptotic methods have also been used successfully to yield approximate inferences without recourse to a reference model at all. The most common procedures that have been used to date are aimed at obtaining easily interpretable representations of the basic seismic variables. They are commonly called inversions.

When linearisation about a reference model is carried out, the frequency differences between the Sun and the model can each be expressed as a sum of spatial averages of independent seismic variables. A prudent procedure is then to seek suitable linear combinations of those averages which more easily inform one of whatever question one has chosen to pose. If it is the value of a particular seismic variable that one wants to investigate, then the most easily interpretable data combinations are those that represent averages of that variable with weighting functions (called kernels in this context) that are highly localised with hardly any sidelobes, and which at the same time are (almost) independent of any other variable, for then the result can be thought of as a blurred view of the variable of interest. Consequently one tries to tailor maximally localised kernels, although the attempted maximisation must be moderated by a requirement that the interference from extraneous seismic variables be kept low and that the influence of data errors not be excessive. Examples of such inversions are depicted in Figures 3 and 4. On the whole, increasing localisation requires data combinations with coefficients of greater and greater magnitude, which increases the influence of random data errors. The procedure to construct those kernels is now called optimally localised averaging (OLA). Just how low one demands the interference from other variables to be, and how much the influence of data errors one considers to be acceptable, is a matter of personal choice, which probably explains why the errors in the averages displayed in Figures 3 and 4 differ. This kind of tailoring was the first to be advocated for helioseismology (Gough, 1978), in contrast to the almost universal opinion of geoseismologists of the time who objected to the procedure on the ground that one might be tempted simply to draw a curve through the averages and mistake that curve for the actual variable, rather than the

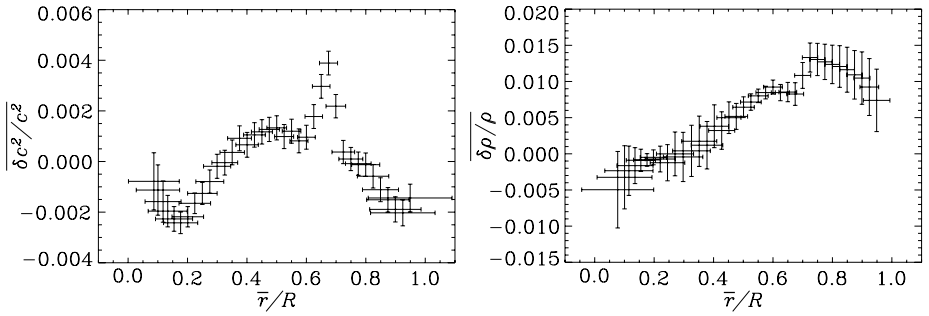


Figure 3 Optimally localised averages of the relative differences of the squared sound speed $[c^2]$ and the density $[\rho]$ in the Sun from those in Model S of Christensen-Dalsgaard *et al.* (1996), computed by M. Takata (Takata and Gough, 2001) from MDI 360-day data and plotted against the centres $[\bar{x} = \bar{r}/R]$ of the averaging kernels $[A(x;\bar{x})]$ (which here resemble Gaussian functions), defined by $\bar{x} = \int xA^2\,\mathrm{d}x \,/\, \int A^2\,\mathrm{d}x$. The length of each horizontal bar is *twice* the spread s of the corresponding averaging kernel, defined as $s = 12\int (x-\bar{x})^2 A^2\,\mathrm{d}x$ – an averaging kernel A that is well represented by a Gaussian function of variance Δ^2 has spread approximately $1.7\Delta \approx 0.72\,\mathrm{FWHM}$; were it to be a top-hat function, its spread would be the full width, which is why s has been so defined. The vertical bars extend to ± 1 standard deviation of the inversion errors, computed from the frequency errors quoted by the observers assuming them to be statistically independent; the errors in the averages are correlated (Gough, Sekii, and Stark, 1996; Howe and Thompson, 1996).

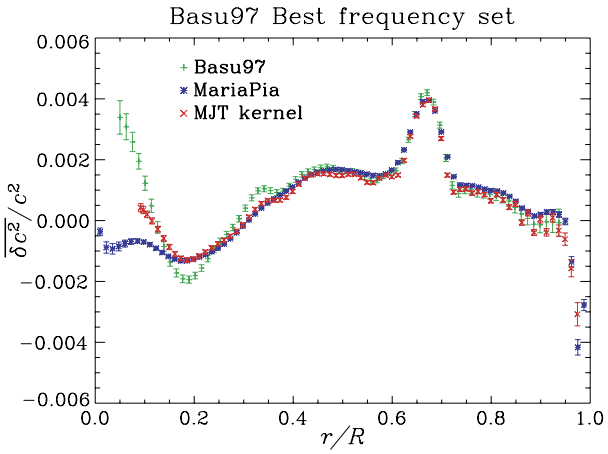


Figure 4 Several sets of optimally localised averages of the relative difference between the squared sound speed in the Sun and that in Model S of Christensen-Dalsgaard *et al.* (1996) computed from the same frequency data set, with kernels obtained by different inverters and by ignoring the contamination from other seismic variables. The vertical bars extend to $\pm$ one standard deviation of the errors computed from the frequency errors quoted by the observers. (The abrupt deviation close to the surface is a characteristic of having misrepresented the acoustic radius (*cf.* Takata and Gough, 2003).)

average that it actually is. We helioseismologists usually do not make that mistake. We explicitly draw a sequence of crosses, as in Figure 3, the horizontal components representing the averaging widths, the vertical bars representing the propagated standard errors in the data. Some geoseismologists are now changing their tune, and are following suit.

An alternative procedure, common in geoseismology, is to seek a putative "solution" that reproduces the data. Typically, one first expresses the seismic variables as linear combina-

tions of a prescribed set of basis functions, and then one chooses the coefficients in each combination to match the data the most closely, again moderated (regularised) to avoid excessive error magnification. The reason geoseismologists used to prefer that approach is that it might yield a curve that actually fits the data within the estimated errors, which OLA might not (and is not explicitly designed to). If it is a linearised perturbation to a reference model that is being fitted to the data, the outcome can also be expressed as averages with localised components. However, the averaging kernels often have severe sidelobes too, usually near the surface of the Sun, which renders interpretation more difficult. Moreover, the "solution", if it exists, is just one of infinitely many that satisfy the constraints imposed by the data. Usually, error-correlation information about the raw data is unavailable, and the data are usually fit by (regularised) least squares (RLS), regularisation being accomplished typically by imposing some criterion of smoothness on the ground that the data provide only finite spatial resolution.

If two seismic variables are involved, as is the case for structure inversions (in the absence of rotation and a magnetic field), then two separate complementary inversions must be carried out for OLA, each requiring the suppression of the contribution to the data combination from the variable not being sought. With the help of some regularising assumption, estimates of the seismic variables can then be made from their averages to enable one to gauge, for each seismic average, the contribution to the corresponding data combination from the other seismic variable; in most of the published "inversions", that contribution is ignored. In the case of RLS frequency fitting, the two seismic variables can be represented simultaneously.

I do not here go into how the inversions are carried out. That is not necessary for understanding the results, provided that adequate information about the averaging kernels is given. Unfortunately, that information is not always provided. Precisely how the centres of the optimally localised kernels were determined is rarely stated, although if the kernels are very narrow it doesn't much matter, unless there are large distant sidelobes. Sometimes authors use merely the locations where they tried to centre the kernels, rather than where they actually succeeded in centring them, which requires conjecture on the part of the reader relating to the information contained in the data set employed and the proficiency of the author in extracting it. One useful rule of thumb for helping to guess what might have been plotted is to recognise that with currently available frequency data well localised kernels without substantial sidelobes cannot be centred closer to the centre of the Sun than $r/R \approx 0.05$.

Inversion is not a well prescribed procedure. It is an art. And it can lead to different representations of the information, as is illustrated in Figure 4, which depicts different optimally localised sound-speed averages all derived from the same data. It should therefore be recognised that the differences between the averages plotted in the figure are not necessarily an indication of inversion error; instead, they can result from the differences in the averaging kernels selected. It is evident, therefore, that merely offering some measure of the widths of the kernels, which is absent from Figure 4, is insufficient for appreciating the results fully.

The objectives of most of our investigations do not concern seismic variables alone. I have already mentioned issues concerning the solar neutrino flux, and also the helium abundance $[Y]$ and the relation between its initial value $[Y_0]$ and Big-Bang nucleosynthesis. Investigation of these non-seismic quantities using seismology necessarily involves relating them to seismic variables via a theoretical model, which itself depends on the assumptions upon which that model has been built. That statement may seem obvious. But those not in our subject do not always appreciate just which of those assumptions are important. It therefore does no harm to state them. Evidently, which of the assumptions are the most important depends on the matter in hand.

I conclude this part of my discussion with another simple point, which is also not obvious to everybody: the dominant physics of helioseismology is extremely simple; it is simply the physics of the propagation and interference of well understood waves. Therefore, provided one is scrupulous in presenting the results with due care and attention, the direct conclusions cannot be questioned. When an observation appears to be in conflict with our seismological knowledge, as superficially seems to be the case with modern spectroscopic determinations of the photospheric heavy-element abundances, for example – an issue to which I shall turn my attention below – contrary to the opinion of some commentators (*e.g.* Guzik, Watson, and Cox, 2006) it is not the seismology itself that is to be challenged.[4] I hasten to add that there certainly are issues within seismology that have not been resolved, usually because a suitable way to analyse the data has not yet been found (possibly because the data do not even contain that information in a form that permits it to be readily extracted, possibly because it is impossible to do so, possibly because the extraction procedure itself still seems to be beyond our capabilities), but that is a different matter.

One further matter concerns the manner in which we use a diagnostic. That typically depends on the issue one wishes to address, and its relation to that issue may itself be subject to some doubt, perhaps due to untested assumptions in modelling. This means that the accuracy of the result may be much less than the precision, perhaps very much less. It is important to recognise the difference between the two. Precision can usually be estimated well from the precision of the data – although in order to trust the outcome one has to trust the estimated errors in the data. Accuracy requires recognising and assessing the influence of the assumptions – inaccurate assumptions can lead to a dispersion amongst the outcomes in any investigation that prudently utilises a variety of different analysis techniques (*cf.* Gough, 2012a). Therefore precision is never judged to be lesser when fewer analysis procedures are considered. Unfortunately, greater precision is often mistaken for greater accuracy.

4. Standard Solar Models

Standard solar models are constructed usually with the most sophisticated microphysics available, yet with the most primitive macrophysics – fluid dynamical processes are ignored wherever possible, probably because they are too difficult to model in an agreed standardised manner. Rotation and magnetic fields are normally ignored too, so the star is spherically symmetrical. The star is evolved hydrostatically either from somewhere on the Hayashi track, gravitationally contracting until it reaches the main sequence, whence radiative energy loss is balanced almost exactly by nuclear energy generated in the core; or it is evolved from an estimate of the zero-age main-sequence structure, which is obtained by equating radiant energy loss with nuclear generation in a completely homogeneous star. The second approach is not quite correct, because there was some nuclear transmutation prior to arrival on the main sequence, most notably conversion of ^{3}He to ^{4}He, which temporarily slows down the contraction (but is insufficient to halt it); however, that phase has little impact on the subsequent evolution of the star. Although the transition from gravitational to nuclear energy release is smooth, during subsequent evolution on the main sequence the central hydrogen abundance $[X_c]$ declines almost linearly with time (Gough, 1995), so backwards

[4] I am assuming that the seismic data processers have judged their errors correctly, and that the error correlations in the frequency data sets, normally ignored, are not unduly severe. I am also assuming that, where it is appropriate, asphericity of equilibrium structure is taken correctly into account.

extrapolation of X_c to the initial hydrogen abundance $[X_0]$ provides a useful fiducial origin of main-sequence age.

Evolution on the main sequence is hydrostatic, and essentially in thermal balance, energy being transported from centre to surface by optically dense radiative transfer wherever the stratification is convectively stable: I present the transfer equation in the next section. Nuclei undergoing thermonuclear reactions are usually considered to be screened by electrons according to the classical, non-relativistic, Debye–Hückel treatment by Salpeter (1954), although the validity of that has been questioned, most recently by Mussack and Däppen (2010, 2011). In convectively unstable regions, the transport of heat is modelled by a (usually local) mixing-length formalism; Reynolds stresses are usually, but not always, ignored. Chemical species are presumed to be homogenised in convection zones; and elsewhere, aside from nuclear transmutation, chemical differentiation is purely by gravitational settling (and possibly radiative levitation) against microscopic diffusion. Energy and material transport by the tachocline circulation is normally ignored, as is transport by acoustic and gravity waves generated by the turbulent convection. The total mass of the Sun is presumed to be conserved, and the "best" equation of state, opacity "formula" and nuclear reaction cross-sections are employed.

It goes without saying that in computing hydrostatic structure it is preferable (some would say mandatory) that the coordinate singularity at the centre of the star and the effective branch point at the radiative–convective interface (when a local mixing-length formalism is adopted) be correctly treated by the numerical integration, accurately enough for the resulting solar model to possess well defined seismic eigenfrequencies. Unfortunately, that is not always achieved, although I suspect that the numerical precision is greater than the observational accuracy.

5. The Chemical Composition of the Sun

Some time ago solar modelling (*e.g.* Christensen-Dalsgaard *et al.*, 1996) had achieved fair agreement with the measured seismic structure of the Sun (*e.g.* Gough *et al.*, 1996). Indeed, Bahcall (2001) considered that to be a "triumph for the theory of stellar evolution", which he illustrated by plotting the relative deviation of the solar sound speed from that of one of his standard solar models (*e.g.* Bahcall, Pinsonneault, and Basu, 2001). It is reproduced here as Figure 5. Yet might one not naively consider it to be an even greater triumph for helioseismology? After all, the seismological errors are much smaller than the modelling errors, as is more clearly evident in Figures 3 and 4. But I have already emphasised that the physics of seismology is very simple, whereas the physics of the structure of the Sun is not. The comparison can therefore be regarded in two opposite ways: i) one might marvel, as did Bahcall, that such complicated physics has been successfully reined to reproduce observation (more or less), or ii) one might consider it unremarkable that a mature theory, which has already survived a wide range of astronomical tests relating to other stars, and which depends on so many somewhat uncertain processes, can be adjusted to reproduce the measurable properties of the Sun. I think I side with Bahcall on this matter, because nobody has actually succeeded in reducing the discrepancy between the models and the Sun sufficiently to come within the much smaller seismological uncertainty. Moreover, the situation has been exacerbated by the results of recent spectroscopic re-analyses, by Asplund *et al.* (2009), Caffau *et al.* (2009, 2011), and Grevesse *et al.* (2011), of some of the chemical abundances in the solar atmosphere, those analyses now taking explicit account of spatial inhomogeneity in the Sun's atmosphere caused by convective motion as simulated by Asplund *et al.* (2000) based on a compressible hydrodynamical procedure described by Stein

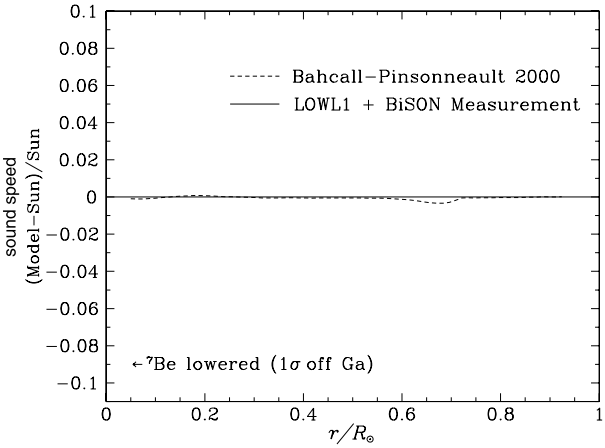


Figure 5 The dashed line is the relative difference between the sound speed in a standard solar model computed by Bahcall and Pinsonneault (1992) and that inferred by Basu *et al.* (1997) from a combination of oscillation frequency data obtained by Chaplin *et al.* (1996), Tomczyk, Schou, and Thompson (1995), and Tomczyk *et al.* (1995) (from Bahcall, Pinsonneault, and Basu, 2001). The arrow labelled ^{7}Be *etc.* is simply a very rough estimate of the sound-speed discrepancy in a theoretical solar model adjusted in a manner described by Bahcall, Basu, and Pinsonneault (1998) to yield a neutrino flux concomitant with that measured by the gallium detector at Gran Sasso (Hampel *et al.*, 1996).

and Nordlund (1998).[5] The early publications (Allende Prieto, Lambert, and Asplund, 2001, 2002; Asplund *et al.*, 2004, 2005a, 2005b; Asplund, Grevesse, and Sauval, 2005) suggested an enormous reduction in the total heavy-element abundance Z below previously accepted values: some 40 % or so. If that represents the abundance in the radiative interior, it implies, on average, a reduction in opacity of some 30 %. By present-day standards, that figure is enormous.

One of the best and most widely adopted solar models (combining all the generally accepted physics for so-called "standard" theory into a model computed accurately enough for reliable computations of its oscillation eigenfrequencies to be possible) is the Model S discussed by Christensen-Dalsgaard *et al.* (1996). It was computed with a mixing-length parameter [α] and initial hydrogen and heavy-element abundances [$X_0 = 0.7091$, $Z_0 = 0.0196$] chosen to yield the correct present-day radius and what was believed to be the correct luminosity, and also a resulting present surface abundance ratio [Z_s/X_s] of 0.0245, consistent with the spectroscopic analysis of Grevesse and Noels (1993), at an age $t_\odot = 4.60$ Gy (notwithstanding the different value quoted by Christensen-Dalsgaard *et al.* (1996), and subsequently by Doğan, Bonanno, and Christensen-Dalsgaard (2010)). The individual surface abundances are $X_s = 0.7373$, $Z_s = 0.0181$; the present surface helium abundance is $Y_s = 0.2447$, which agrees with observation within the limits of accuracy of the equation of state and helioseismological analyses of the depression of γ_1 caused by the second ionisation of helium in the adiabatically stratified convection zone (Gough, 1984b; Kosovichev *et al.*, 1992; Vorontsov, Baturin, and Pamyatnykh, 1992; Basu and Antia, 1995;

[5]The original analysis employed simulations in a model atmosphere with the previously accepted chemical composition; from 2009 onwards, newer simulations by Asplund and his collaborators were carried out with the Fe abundance inferred from the 2005 abundance determinations; the abundances of C, N, and O were not changed because their spectral lines are weak and have no significant impact on the radiative energy flux (R. Trampedach, personal communication, 2012).

Däppen, 2007), as I shall discuss briefly later. Figures 3 and 4 illustrate how well the seismic structure of the model corresponds to that of the Sun.

The heavy-element abundance now favoured by Grevesse *et al.* (2011) is somewhat greater than the original announcement, namely $Z_s = 0.0134 \pm 0.0005$; an independent and somewhat different analysis by Caffau *et al.* (2011) that also takes convective inhomogeneity into account has yielded $Z_s = 0.0153 \pm 0.0011$. Both of these values were estimated from abundances of the major opacity-producing elements except neon, whose abundance in the Sun cannot be measured spectroscopically. A recent compilation by Lodders, Palme, and Gail (2009) using a wider range of solar data, with an eye also on meteoritic abundances, yet using more simple hydrostatics, has yielded the recommendation $Z_s/X_s = 0.0191$. Accepting $Y_s = 0.249$ from a calibration of model envelopes by Basu and Antia (2004) using seismically accessible (asymptotic) integrals that are sensitive to the depression of γ_1 due to He II ionisation, or $Y_s = 0.224$ from a calibration of complete solar models (Houdek and Gough, 2011), yields $Z_s = 0.0141$ and $Z_s = 0.0145$, respectively. (It is unclear whether calibrating complete solar models in this fashion is more reliable or less reliable that calibrating only model envelopes.) These values are all substantially lower than that of Model S, on average by about 20 %. What are the implications of these results?

5.1. The Abundances in Context

Many of the discussions that have ensued have carried out solar-evolution computations and then compared, in some manner, implied seismic data, or seismic structure (principally sound speed), with the inferences from the Sun. They have been catalogued by Basu and Antia (2008). However, seismic structure is merely a diagnostic of the issue, not the issue itself. It is instead much more lucid to the average physicist to be confronted with the direct implication of the newly reported abundances. So let us accept (most of) the assumptions of solar modelling, as is usually done in the discussions of the problem, and ask what they really imply. I start with the most obvious: that the Sun approached the main sequence fully mixed, having just come down the Hayashi track. Therefore the heavy-element abundance $[Z(r, t_\odot)]$ in the radiative interior today is (almost) the same as it is in the photosphere, aside from a relative excess of 3 % or so due to differential gravitational settling. Heavy elements influence the structure of the Sun primarily via their effect on the opacity $[\kappa]$, which controls the relation between luminosity $[L]$ and temperature $[T]$ through the equation of radiative transfer: $L(r,t) = -(16\pi a c r^2 T^3/3\kappa\rho)\,\mathrm{d}T/\mathrm{d}r$. However, T is not a seismic variable, and therefore cannot be measured directly. It is related to the seismic variables c^2 and ρ by the equation of state: $T = \mathcal{T}(c^2, \rho; Y, Z)$ – I set aside, for the moment, consideration of the relative abundances of the heavy elements. The contribution of Z to the equation of state is only about 0.5 %, so for the purposes of this discussion its uncertainty can safely be ignored. There remains only the helium abundance $[Y]$ which is to be estimated from the theory of stellar evolution. It is important to appreciate that this is the only stage in the argument where the details of the evolution theory come into serious play. However, it must be realised that knowledge of the functional form of $Y(r)$ today is crucial for assessing the stratification of the Sun's radiative interior. Therefore, some attention must be paid to how it is determined, appreciating the assumptions to which the outcome is sensitive.

As the Sun evolves on the main sequence, hydrogen is transmuted into helium in the core, adding a spatially and temporally varying component $\delta_\varepsilon Y$ to the helium abundance: $Y(r,t) = [1 + s_Y(r,t)]Y_0 + \delta_\varepsilon Y$, where $s_Y Y_0$ is the change in Y produced by gravitational settling moderated by diffusion. It is known (*e.g.* Gough, 1983c; 1990) that the time dependence of the total luminosity $L_s(t) := L(R,t)$ of the solar model, calibrated to satisfy

$L_s(t_\odot) = L_\odot$, is not very sensitive to the details of the theory (such as the choice of Z_0, or whether or not there has been some small degree of material mixing in the core – we know from seismology that the core cannot have been homogenised, a conclusion which is consistent with HR diagrams of solar-like stars), so the total amount of helium that has been produced by today $[\int \delta_\varepsilon Y \, \mathrm{d}r]$ is proportional to $\int L_s \, \mathrm{d}t$, which is essentially known (provided $t_\odot$ is known). Therefore it is adequate for the current discussion to accept the function $\delta_\varepsilon Y$ from any standard model. Also, the function $s_Y(r,t)$ is only very weakly dependent on Y_0, and may also safely be taken as given by the model. Likewise, one can take $Z = [1 + s_Z(r,t)]Z_0$ with s_Z given by the model (it is broadly similar to the function s_Y). Whence $T(r,t) = \mathcal{T}(c^2, \rho; (1+s_Y)Y_0 + \delta_\varepsilon Y, (1+s_Z)Z_0)$, where c^2 and ρ have the values determined seismologically from the Sun. Only the initial abundance Y_0 now remains unknown. How do we determine it?

I recommend calibrating Y_0 by accepting the nuclear reaction rates – after all, all the pertinent nuclear cross-sections have been carefully assessed experimentally during the investigations of the neutrino problem, save the controlling p–p reaction cross-section which is determined only theoretically (however, it is the simplest of all nuclear reactions). There remains also an issue concerning the screening of energetic particles (*e.g.* Mussack and Däppen, 2010, 2011). These matters are probably, for the purposes of this discussion, minor. Therefore the total rate of thermal energy production, hence the luminosity, is determined in terms of $t_\odot$ and Y_0. Granted that $L_\odot$ and $t_\odot$ are known quite well, Y_0 can thus be calibrated to determine $T(r)$ today,[6] which can then be substituted into the equation of radiative transfer to evaluate $\kappa(r)$. That procedure can either be carried out explicitly from a seismological inversion to determine the seismic stratification, or implicitly within the inversion itself, such as in the manner advocated by Elliott (1995). The outcome is illustrated in Figure 6, in which the continuous curve is the relative difference $(\kappa_\odot - \kappa_S)/\kappa_S$ between the Sun's opacity $[\kappa_\odot]$ determined in the manner that I have just described and that of Christensen-Dalsgaard's Model S. Not surprisingly, $\kappa_\odot$ is very similar to the opacity in Model S. The issue posed by

[6]Implications from earlier estimates of Y to determine $T(r)$ have been discussed by Elliott (1995) and Tripathy and Christensen-Dalsgaard (1998), the first under the assumption that $Y(r)$ differs from that in a reference solar model by just a constant, the second that it can be obtained simply by scaling the reference-model value by a constant factor, both of which are inaccurate in the core, although Elliott notes that the opacity perturbations produce a predominantly local response, so that these analyses should provide a fair estimate in the radiative envelope. The Tripathy–Christensen-Dalsgaard scaling was used by Tripathy, Basu, and Christensen-Dalsgaard (1998) to obtain the opacity difference from Model S by RLS frequency fitting (J. Christensen-Dalsgaard, personal communication, 2012) assuming that difference can be expressed as a function of T alone, yielding a superficially similar functional form to the continuous curve in Figure 6, but with a magnitude about 50 per cent greater. Bahcall *et al.* (2005) have estimated the opacity difference by adjusting κ by hand in solar models. Their preferred model had a constant 11 % augmentation over the OPAL values using the most recent abundance determinations (Asplund *et al.*, 2000, 2004; Asplund, 2005; Allende Prieto, Lambert, and Asplund, 2001, 2002; Asplund personal communication with Bahcall *et al.*, 2004) in the radiative envelope down to $T = 5 \times 10^6$ K, beneath which the augmentation was smoothly reduced to zero (probably by a half Lorentzian function with half-width at half maximum of 2×10^5 K), and has a seismic structure as close to that of the Sun as does Model S. Regarding that model as a proxy Sun, one would expect the relative opacity difference between it and a model with the unmodified abundances to be comparable with the inference by Christensen-Dalsgaard *et al.* (2009). The two estimates are depicted in Figure 6. Korzennik and Ulrich (1989) had earlier estimated opacity errors by RLS (L_2 norm) data fitting, and Saio (1992) by L_1 data fitting, each by scaling κ by a function of T and ignoring the dependence of the relation between T and the seismic variables on chemical composition; they expressed their results as deviations from different reference models, so they cannot easily be compared with those presented in Figure 6. An estimate by OLA (Takata and Gough, 2001) of the absolute structure of the Sun, including opacity, using the procedure for determining Y described in the text (together with tachocline homogenisation as calibrated by Elliott, Gough, and Sekii (1998)) is presented by Gough and Scherrer (2002) and Gough (2004, 2006).

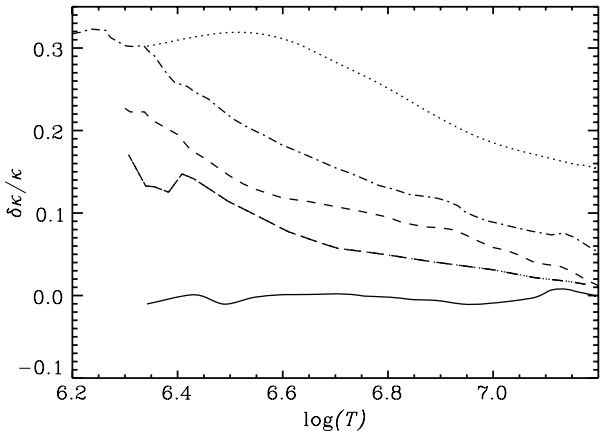


Figure 6 The dot–dashed curve is the estimate by Christensen-Dalsgaard *et al.* (2009) of the relative opacity augmentation required when adopting the Asplund (2005) chemical composition in OPAL opacity computations (Iglesias and Rogers, 1996) to produce a solar model with the same sound speed as in Model S of Christensen-Dalsgaard *et al.* (1996). The dashed curve represents the augmentation, computed by Christensen-Dalsgaard and Houdek (2010), required of OPAL opacities with the revised Asplund *et al.* (2009) abundances. The triple-dot–dashed curve is a comparable difference between two solar models computed by Bahcall *et al.* (2005) with OPAL opacities, the first with the latest abundances by Asplund *et al.* at the time, the second with abundances that have been artificially chosen to yield what one might regard as a seismically acceptable proxy Sun. For comparison, the dotted curve is a linearised estimate of the same quantity were the relative abundances of the heavy elements to have been preserved, using opacity derivatives obtained from the tables of Bahcall and Pinsonneault (1992). The continuous curve is the relative difference $(\kappa_\odot - \kappa_S)/\kappa_S$ inferred by Gough (2004) between the Sun's opacity $[\kappa_\odot]$ and that of Model S.

Asplund *et al.* and Caffau *et al.* is therefore simply an abundance–opacity problem: how can $\kappa_\odot(r)$ be reconciled with their abundance measurements?

5.2. Suggestions for Reconciliation

An obvious naive suggestion is that opacity calculations are in error by just the appropriate factor required to compensate for the proposed revision in Z, leaving the functional dependence on the other variables unscathed. Given the testing that has been undertaken by Iglesias and Rogers, that does not seem so very likely. It is also unlikely that energy transport is opposed by some other mechanism, such as gravity-wave transport, because that would require a completely different physical process to mimic the functional form of the effect of radiatively induced atomic transitions. Of course, it could be that the new abundance determinations are in error, despite the extra care that has been taken; we have seen an increase since the early work by Allende Prieto, Lambert, and Asplund (2001, 2002) and Asplund *et al.* (2004), Asplund, Grevesse, and Sauval (2005), Asplund *et al.* (2005a, 2005b), so a further increase might not be too surprising. However, the independent investigation by Caffau *et al.* (2009, 2011) has yielded results not very dissimilar from those of Asplund *et al.* (2009) and Grevesse *et al.* (2011), which adds to the credibility of the new abundances. Of course, another possibility is that early in its post-Hayashi days the Sun was contaminated by metal-poor material, as Guzik, Watson, and Cox (2005) have suggested. If that were the case, it may just be possible to detect seismologically a relic compositional discontinuity at the location of the base of the zero-age convection zone, provided that it has not been destroyed. The discontinuity is at least stable to double-diffusive convection, so it would not

be destroyed spontaneously. Yet a fourth possibility is that the relative abundances of the heavy elements are different from what is generally believed, particularly that of neon, as has been discussed, for example, by Drake and Testa (2005), leaving open the possibility that Z is greater than the values suggested by Asplund *et al.* and Caffau *et al.* based on the assumption that the relation of the neon abundance to the abundances of the other opacity-producing elements is preserved, and that therefore κ is perhaps proportionately greater too. That too is regarded by many spectral analysts as being intrinsically unlikely, and in any case neon alone does not mimic the heavy-element mixture adequately to restore the solar models to their pre-Asplund state.

The apparent opacity discrepancy has evolved with the abundance revisions. Christensen-Dalsgaard *et al.* (2009) estimated the amount $[\delta\kappa]$ by which the opacity according to OPAL (Iglesias and Rogers, 1996) with Asplund (2005) chemical composition would need to be augmented in order to produce an otherwise standard solar model with the same sound speed as in Model S, and hence about the same sound speed as in the Sun. It is depicted as the dot–dashed curve in Figure 6, where it is plotted against $\log(T)$ along the thermodynamic ρ–T path of one of the models. One might have expected the outcome to have been given approximately by $(\partial \ln \kappa / \partial \ln Z)_{\rho,T,X} \delta \ln Z$, in which $\delta \ln Z \approx \ln Z_{\rm S} - \ln Z_{\rm Asp} \approx \ln Z_{\rm s,S} - \ln Z_{\rm s,Asp} = 0.370$, the subscript s denoting surface value and the subscripts S and Asp denoting Model S (Grevesse and Noels, 1993) and Asplund (2005) values, respectively, and where the opacity derivative is evaluated at constant relative heavy-element abundances. However, it is evident from Figure 6 that is not the case. According to Jørgen Christensen-Dalsgaard (personal communication, 2011) the quite substantial difference arises because the relative heavy-element abundances in the two models differ. Unfortunately, that renders back-of-the-envelope estimates suspect. The later abundance determinations by Asplund *et al.* (2009) have brought the OPAL opacities closer to those of the Sun, as the dashed curve in Figure 6 depicts. It is interesting that Antia and Basu (2011), using the chemical composition proposed by Caffau *et al.* (2011), whose relative abundances do differ from the earlier values given by Grevesse and Sauval (1998), Grevesse and Noels (1993), and presumably also the more modern values of Grevesse *et al.* (2011), report that they have obtained solar models that are seismically almost as good as models constructed with the abundances of Grevesse and Sauval (1998), and they found even closer agreement with the Sun if the assumed abundance of neon were artificially enhanced by a factor $\sqrt{2}$. That essentially reduces the problem of reproducing the Sun's seismic stratification theoretically to the state in which it was prior to Asplund's original announcement. It must be realised, however, that that does not close the matter, because merely reproducing previous partial results does not necessarily prove the veracity of the physics behind the new models. Moreover, the seismic structures of the models, new and old, deviate from that of the Sun by many standard errors, as is evinced by Figure 3.

5.3. Possible Flaws in the Argument

The foregoing discussion is predicated on the presumption that the sole significantly discrepant ingredient in the solar modelling is opacity. That need not be the case; modification of any non-seismic variable might, at least in principle, bring a theoretical model into line with observation. Evidently, a direct change in the relation between the seismic variables pressure and density, for example, via a modification to the abundances of the abundant chemical elements H and He, can alter the sound speed: Gough and Kosovichev (1988, 1990) found that the stratification of the energy-generating core can be reproduced by a slight smoothing of the abundance profile that had been produced by nuclear transmutation, suggesting that a slight degree of mixing has taken place, possibly by disturbances

that have been shear-generated in a manner analogous to clear-air turbulence in the Earth's atmosphere. Alternatively, there could be an additional source of energy transport without mixing, such as by accreted weakly interacting massive particles (WIMP), whose presence in the Sun was postulated originally in order to try to account for the observed low neutrino flux without neutrino transitions (Spergel and Press, 1985); WIMP modify the temperature distribution, and consequently the distribution of helium produced by nuclear transmutation, leading to a sound-speed modification whose general functional form is broadly similar to what would be required (Gilliland *et al.*, 1986), although a serious attempt to reproduce the seismic structure of the radiative interior appears not to have been made (see also Faulkner, Gough, and Vahia, 1986; Däppen, Gilliland, and Christensen-Dalsgaard, 1986; Christensen-Dalsgaard, 1992).

It is worth commenting that a reduction in Z_s at fixed relative abundances of only about the 20 % below that of Model S implies, via the theory of solar evolution, a reduction of about 0.02 in Y_s. This appears to me to be not dissimilar to the margins of the true uncertainty in the helioseismological determinations of the helium abundance of the convection zone – although not within the precision quoted by those who have attempted to determine Y_s by only a single procedure (*e.g.* Basu and Antia, 1995, 2004; Basu, 1998; Richard *et al.*, 1998; Di Mauro *et al.*, 2002), even granted that systematic errors had been recognised.

5.4. Seismological Investigation

Chemical abundances can in principle be measured directly by seismology by analysing the non-ideal properties of the solar plasma. Specifically, γ_1 is depressed in regions of partial ionisation, by an amount that is almost proportional to the abundance of the ionising element. Däppen and Gough (1984) proposed measuring the helium abundance by calibrating the ionisation-induced variation of a thermodynamic function $[\Theta]$ expressible in terms of γ_1 and its derivatives and which is easily accessible to seismological probing in the adiabatically stratified layers of the convection zone (Gough, 1984b). A procedure for so doing was developed by Däppen, Gough, and Thompson (1988), who found Y_s to be lower than typical values obtained from calibrating evolved solar models (Däppen *et al.*, 1991), which triggered Christensen-Dalsgaard, Proffitt, and Thompson (1993) to investigate the seismological implications of gravitational settling of heavy elements, with notable success. Other forms of calibration, often based more directly on γ_1, have been pursued (*e.g.* Däppen *et al.*, 1991; Basu and Antia, 1995, 2004; Basu, 1998; Richard *et al.*, 1998; Di Mauro *et al.*, 2002; Houdek and Gough, 2007). None has been completely satisfactory, not least because there are serious uncertainties in the equation of state. There have been attempts to estimate Z by calibrating solar models against low-degree frequency-separation ratios assuming $t_\odot$ to be known (*e.g.* Chaplin *et al.*, 2007); that is equivalent to calibrating Y, of course.

A direct seismological estimate of Z in the convection zone, and thus a direct estimate of Z_s, via the ionisation-induced depressions in γ_1 or in the variation of the seismologically more accessible function Θ of γ_1 and its derivatives, is very delicate. It would probably be necessary to average the individual variations in the different ionisation zones, so the result would depend to some degree on the assumed values of the relative abundances (Mussack and Gough, 2009). Nevertheless, if successful, the accuracy of the result could probably be assessed from the seismic frequency uncertainties, because the relation between the magnitude of the γ_1 depression and the abundance of any ionising element producing it, at least at the level required to judge the abundance–opacity problem, is fairly robust. The seismologically apparently more straightforward procedure of measuring a spatially averaged absolute value of Θ and relating that to the slightly greater value expected in the absence of heavy

elements, as Antia and Basu (2006) have tried, is more uncertain because the change in Θ due to heavy-element ionisation that is sought appears to be less than the uncertainty in the equation of state arising, for example, from the commonly neglected, at least in the chemical picture (Däppen, 2004, 2007), finite volume occupied by bound species (*e.g.* Baturin *et al.*, 2000).

5.5. The Equation of State

My statements about the uncertainty in Y_s and the equation of state call for some justification. It goes without saying that chemical composition is not a seismic variable, not even the abundances X and Y of the abundant elements H and He. In order to determine Y, say, it is evidently necessary to use an equation of state, which relates the seismic variables to composition via another, necessarily non-seismic, variable, such as T or s, and subsequently, at some point, to confine attention to the adiabatically stratified region of the convection zone where that variable can be eliminated from the spatial variation of the seismic variables. Once that has been accomplished one can learn about chemical composition through the effect of ionisation on γ_1. This has been approached either directly by measuring some appropriate property of γ_1 itself, or by working with some thermodynamic function Θ of it. The reliability of the outcome necessarily rests on the reliability of the equation of state, which is very difficult to assess. That should be obvious because there is no redundancy in the dependence of the seismic modes on the seismic variables. Therefore, as has been alluded to in the past (*e.g.* Gough, 2004), perhaps too obliquely, any intrinsic error in the equation of state cannot be assessed by seismology alone. To make progress, non-seismic information must be incorporated. In practice, that information comes from some prior appreciation of the reliability of some aspects of the equation of state, coupled with the additional non-seismic constraint that the convection zone is chemically homogeneous (sometimes augmented with the constraint that deep down the stratification is adiabatic). Suppose, for example, one is estimating the deviation of the Sun from a reference solar model, represented by $\delta\ln\gamma_1$ and $\delta\ln u$, where $u(r)$ is a complementary seismic variable. From seismology one can relate averages $\overline{\delta\ln\gamma_1} + \overline{\overline{\delta\ln u}}$ of those deviations to data combinations d, where the single and double overbars indicate simply that the averaging kernels are different. Then, one can write the first average as the sum of $\overline{(\partial\ln\gamma_1/\partial\ln u)_Y\,\delta\ln u}$, which can be incorporated into the second average, of $\overline{(\partial\ln\gamma_1/\partial Y)_u\delta Y}$, where δY is the difference in Y between the Sun and the reference model, and of a component $\overline{\delta_{\rm int}\ln\gamma_1}$ resulting from intrinsic error in the equation of state. One would like to be able to distinguish between those components. However, that is strictly impossible because they have the same averaging kernel. Therefore one cannot unambiguously determine δY from the data $[d]$. What has been attempted in the past is to design kernels such that those defining the overbar and the double overbar are both in some sense small, and then to estimate δY by neglecting $\overline{\delta_{\rm int}\ln\gamma_1}$ and $\overline{\overline{\delta\ln u}}$, a procedure which is evidently not strictly valid; after that, $\overline{\delta_{\rm int}\ln\gamma_1}$ can be computed with presumably different and well localised averaging kernels. The results of such a procedure have been presented by Basu, Däppen, and Nayfonov (1999), Di Mauro *et al.* (2002), and Däppen (2004), using both OPAL and MHD, the two most popular, and probably the best, equations of state available today. Not only do their inferred intrinsic errors from the two equations differ, as they surely must, but so too do the estimates of the helium abundance $[Y_s]$ in the Sun's convection zone.[7] It must be appreciated that any seismological inference about $\delta_{\rm int}\ln\gamma_1$ such as this is

[7] Di Mauro *et al.* quote $Y_s = 0.2539 \pm 0.0005$ when OPAL is used, $Y_s = 0.2457 \pm 0.0005$ when MHD is used, the errors being merely formal, representing a precision error that takes no account of the error in the

susceptible to potential errors in the inferred values of Y_s, which are difficult to appreciate because the precise manner in which δY was obtained is unclear. However, well outside the hydrogen and helium ionisation zones $\partial \ln \gamma_1 / \partial Y$ is small, so errors in δY can hardly contaminate $\overline{\delta_{\rm int} \ln \gamma_1}$ there. This is a robust property of any realistic equation of state, because under normal stellar conditions γ_1 is essentially independent of Y where hydrogen and helium are both fully ionised, errors in the small dependency that does remain (arising from processes such as the influence of the electron density on the ionisation of heavy elements) being characterised by the quantity $\overline{\delta_{\rm int} \ln \gamma_1}$ itself. However, the situation is different within the ionisation zones, where in general $\overline{\delta_{\rm int} \ln \gamma_1}$ and $\overline{(\partial \ln \gamma_1 / \partial Y) \delta Y}$ contribute comparably, and cannot be entirely recognised apart. I must add, however, that $\overline{\delta \ln \gamma_1}$ and $\overline{(\partial \ln \gamma_1 / \partial Y)}$ are normally functionally different, so $\overline{\delta_{\rm int} \ln \gamma_1}$ cannot vanish everywhere.

An alternative, more transparent, approach to an attempt at separation could be to adopt the attitude that one's chosen equation of state is as good as it can be, at least in the adiabatically stratified region where $\overline{\delta \ln \gamma_1}$ has been inferred. Therefore one could choose for each equation of state the value of δY that minimises the integral with respect to acoustic radius of $(\delta_{\rm int} \ln \gamma_1)^2$ over that region – I choose the acoustic radius because that is the natural seismic frequency-controlling independent variable. The outcome, using the inferences of $\overline{\delta_{\rm int} \ln \gamma_1}$ and Y_s by Di Mauro *et al.* (2002), is $Y_s = 0.2430$ for OPAL and $Y_s = 0.2280$ for MHD; I do not quote a precision because it is hardly material to estimating accuracy in this context. The range of all these values, roughly 0.02, must surely be regarded as a lower bound to the range of values within which the solar Y_s is likely to lie, because the discrepancies are systematic; that value is some 40 – 100 times greater than typical errors quoted by, for example, Basu (1998) or Di Mauro *et al.* (2002), based on the precision of the specific procedures that were carried out. Based solely on the calibrations described in this paragraph, one might optimistically conclude that $Y_s = 0.243 \pm 0.01$, while recognising that the accuracy might well have been overestimated. Consequent estimates of $\overline{\delta_{\rm int} \ln \gamma_1}$ are illustrated in Figure 7.

5.6. Other Matters to Consider

There are uncertainties in other quantities used in estimations of the chemical composition that still need to be addressed. Amongst them are the current luminosity $[L_\odot]$ and the age $[t_\odot]$, a matter to which I have already alluded. The luminosity is inferred from the total solar irradiance, and is currently obtained by assuming that the radiant output is spherically symmetric (*e.g.* Willson and Hudson, 1988). That assumption is not strictly correct, as models of solar-cycle irradiance variation indicate (*e.g.* Taylor *et al.*, 1998; Foukal *et al.*, 2006; Fröhlich, 2011); accounting for the asphericity augments $L_\odot$ by some 1.5 % or so. However, there has been a recent downward revision in the measured value of the irradiance from an end-to-end instrumental recalibration (Kopp and Lean, 2011), which is partially compensating. The age $[t_\odot]$ is more uncertain. It is likely to be about the same or only slightly greater than the age of the oldest meteorites, which seem to lie between 4.563 and 4.576 Gy (Amelin *et al.*, 2002; Jacobsen *et al.*, 2008, 2009; Bouvier and Wadhwa, 2010). Attempts

relation between Y_s and γ_1 in the reference model. With those values the magnitude of the inferred $\overline{\delta_{\rm int} \ln \gamma_1}$ was found to be the greater for the MHD equation of state beneath the helium ionisation zones, and the lesser above $r/R \approx 0.97$ in the He II ionisation zone. That is perhaps not surprising because MHD is possibly better at taking into account the complicated chemistry that dominates higher up in the solar envelope, whereas the virial expansion used for OPAL is perhaps more reliable where such complications make only a minor non-ideal contribution. The difference between OPAL and MHD must surely offer some estimate of the total uncertainty.

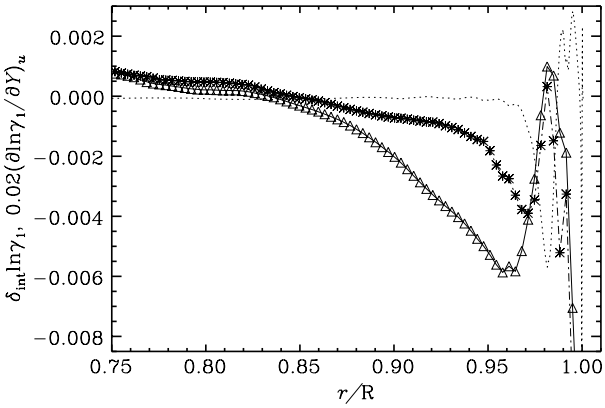


Figure 7 Inferences concerning the intrinsic error in the OPAL (asterisks) and MHD (triangles) equations of state, based on the assumption that the helium abundance, in each case, is such as to minimise its L_2 norm with respect to acoustic radius over the region of the convection zone that is stratified adiabatically to better than 1 part in 10^3, computed from the analysis of Di Mauro *et al.* (2002). Because the helium abundance is uncertain, the symbols should not be meant to represent reliable estimates of the actual intrinsic errors $\overline{\delta_{\mathrm{int}}\ln\gamma_1}$. Formal error bars are not drawn because they are enormously smaller than the true uncertainties, and might therefore confuse. The dotted curve is $0.02(\partial\ln\gamma_1/\partial Y)_u$, computed with the OPAL equation of state.

to calibrate solar models seismically have yielded values equal to or slightly lower than the age [t_{S}] adopted for Christensen-Dalsgaard's model S, namely 4.60 Gy. Those calibrations in which a value of Z_{s} is assumed to be that of Model S yield a value rather lower than t_{S} (Dziembowski *et al.*, 1999; Bonanno, Schlattl, and Paternò, 2002; Doğan, Bonanno, and Christensen-Dalsgaard, 2010), the most recent, based on BiSON data (Chaplin *et al.*, 2007) being 4.57 Gy; a simultaneous calibration of Z_{s} and $t_\odot$ by Houdek and Gough (2011), also with BiSON data (Basu *et al.*, 2007), has yielded precisely t_{S}. Interestingly, the heavy-element abundance in the latter investigation was calibrated to be $Z_{\mathrm{s}} = 0.0142$, essentially the same as the recent value preferred by Lodders, Palme, and Gail (2009), and lying between the values preferred by Asplund *et al.* (2009) and Caffau *et al.* (2011). However, the model was calculated with the relative heavy-element abundances of Grevesse and Noels (1993) (neon not being artificially enhanced); therefore the opacity is too low and the seismic structure cannot be correct.[8] Whether that has yielded a superior or an inferior estimate of $t_\odot$ is unclear (Gough, 2012a). The calibration relies partially on estimating Y_{s} seismologically from the oscillatory component of the low-degree frequency distribution caused by the acoustic glitch associated with the depression of γ_1 in the helium ionisation zones (Houdek and Gough, 2007). To accomplish that estimate it was assumed that hydrostatic support is solely a balance between pressure gradient and gravity, as is usual: magnetic stresses and the centrifugal force were ignored. (It was also assumed that the Sun is spherically symmetric, as in other calibrations, which we know is not strictly correct (*e.g.* Goode *et al.*, 2002; Komm, Howe, and Hill, 2006; Emilio *et al.*, 2007; Fivian *et al.*, 2008; Kuhn *et al.*,

[8]Moreover, the initial helium abundance [Y_0] of the calibrated model is about 0.250, which is dangerously close to the amount [Y_{p}] believed to have been created by Big-Bang nucleosynthesis, whose estimated value has been climbing over the last two decades (Steigman, 2007): the latest estimate is $Y_{\mathrm{p}} = 0.2478 \pm 0.0006$ (G. Steigman and M. Pettini, personal communication, 2011). Subsequent contamination of the interstellar medium by supernovae exacerbates the situation.

2012); by how much the structural asymmetry contaminates the assumed relation between the degree-dependence of the seismic frequencies and the locations of their lower turning points, of crucial importance to the structural calibrations, is yet to be ascertained.)

The validity of ignoring magnetic stress has been implicitly questioned by Basu and Mandel (2004) and Verner, Chaplin, and Elsworth (2006), who reported a 10 % solar-cycle variation in the glitch signal. That variation is huge, and, if correct, cannot possibly result from a temporal abundance variation. It has been pointed out that were the glitch variation to be magnetic, an intensity variation of some ten or so Tesla in the second He ionisation zone would be implied (Gough, 2006). Moreover, it would introduce an error in the two-parameter age calibration of similar relative magnitude (Houdek and Gough, 2011). However, Christensen-Dalsgaard *et al.* (2011) have recently failed to detect a change of such magnitude, although they admit that further analysis is necessary in order to be sure. Meanwhile, Basu *et al.* (2010) maintain that the change does occur. The whole issue is evidently very important, and we would like to know the answer. It may be some time before we do.

6. What of Importance Is There Left to Be Learnt?

Helioseismology was amazingly successful in the early days. It demonstrated that the resolution of the solar neutrino problem was not to be found in adjustments to parameters of a standard solar model, and later, as precision increased, that the resolution must be sought in nuclear or particle physics. As we all now know, the matter is resolved by neutrino transitions. Helioseismology has also provided estimates of the solar helium abundance in the convection zone, perhaps not as accurately as we would like, yet which, granted gravitational settling, are at least not seriously at variance with Big-Bang nucleosynthesis. Coupled with that is a precise measure of the location of the base of the convection zone (Christensen-Dalsgaard, Gough, and Thompson, 1991; Basu and Antia, 1997), which has been used extensively as a simple diagnostic for calibrating theoretical solar models (*e.g.* Turck-Chièze *et al.*, 1993; Bahcall, Pinsonneault, and Basu, 2001; Guzik, Watson, and Cox, 2005), and is important for defining the boundary conditions in numerical simulations of the convection zone. In addition, we know that the quadrupole moment $[J_2]$ of the exterior gravitational equipotentials is about 2.2×10^{-7} (*e.g.* Schou *et al.*, 1998; Antia, Chitre, and Gough, 2008), contributing to the precession of planetary orbits, particularly that of Mercury, by an amount that is compatible with General Relativity. The latitudinal variation of the angular velocity $[\Omega]$ in the photosphere persists, approximately, to the base of the convection zone, beneath which is a thin interface, the tachocline, and a rigidly rotating interior (admittedly with some doubt about the rotation of the energy-generating core).

We seem to have answered the major outstanding questions accessible to seismology that interest most of the scientific community. So is global helioseismology finished? Is helioseismology finished? To be sure, there is the unfinished business I talked about in the previous section, but it may require only minor details to sort that out. It is now often suggested by commentators that the days of true excitement in the subject, and of important discovery, are over.

I dissent vehemently from that view. Firstly, the issues that I discussed in the previous section are related to the microphysics of opacity and the equation of state, which are important to plasma physics and may end up being very important also to the study of stars other than the Sun. More detailed seismological investigation may be able to unveil, or at least provide clues to, some fundamental unresolved issues in plasma physics, such as, in

the chemical picture, how the energy difference between the continuum of "unbound" electrons and the strongly bound states is determined. This is related to the whole matter of electron screening of charged species, a reliable consistent quantum-mechanical study of which is still wanting, and has implications also for thermonuclear reaction rates (*e.g.* Däppen, 1998). But more obvious is the whole matter of the internal macroscopic dynamics of the Sun: large- and medium-scale angular-momentum transport; material redistribution by meridional flow in the radiative zone; the pattern of the large-scale meridional flow in the convection zone; augmentation, distortion, and decay of the magnetic field and its back-reaction on the flow; the formation, evolution, and final decay of sunspots and other forms of activity; and how all these conspire to drive and control the solar cycle, and, perhaps most prominently, modulate the solar outputs of particles and electromagnetic radiation that influence the climate on Earth. Many of these are the object of ongoing local helioseismological techniques – time–distance and seismic holography, and ring analysis. I have not discussed those techniques here. Ring analysis is relatively straightforward, and has given us views of horizontal flow not far beneath the photosphere. Telechronoseismology has given us a view of a sunspot, complete with the flow around it: deep divergent horizontal flow carrying away the excess heat rising around the obstructing spot and an associated convergent subphotospheric counterflow above (Zhao, Kosovichev, and Duvall, 2001; Zhao and Kosovichev, 2003; Kosovichev, 2009; Zhao, Kosovichev, and Sekii, 2010). Divergent (Evershed) motion has now been detected even closer to the photosphere (Zhao *et al.*, 2011) – a critical feature that was missing from the original seismological analysis, and which caused some onlookers to harbour grave doubts. The qualitative picture, which has made its way onto the well-known SOI/MDI coffee mug, is now not only physically plausible, but also in accord with photospheric observations. A dynamical picture of the Evershed flow is now emerging (*e.g.* Weiss *et al.*, 2004; Kitiashvili *et al.*, 2009, 2010). More recent helioseismic observations (Kosovichev and Duvall, 2011) have revealed the cessation of the deeper subphotospheric convergent flows as sunspots decay, flows which one might presume had previously held the spots intact. However, different seismological analyses are not yet all entirely consistent (*e.g.* Gizon *et al.*, 2009; Hindman, Haber, and Toomre, 2009), so the picture is not wholly secure. Also, more extensive observations will be needed for understanding the entire life-cycle of the spots.

Another matter of importance for understanding the global internal dynamics is the structure of the tachocline and its associated meridional flow, and how that influences the convection above. It is extremely difficult to detect deep meridional flow seismologically, because the frequency perturbations are small; however, some progress appears to be possible with the use of eigenfunction distortions (Schou, Woodard, and Birch, 2009; Gough and Hindman, 2010), and Antia, Chitre, and Gough (2012) are trying to advance simplified seismologically constrained dynamical arguments. If the meridional tachocline flow is downwelling near the Equator and the Poles, and upwelling near the latitude of zero tachocline shear, as Spiegel and Zahn (1992) and Gough and McIntyre (1998) have argued, how does that flow connect to the general meridional flow in the convection zone? Or is that flow so slow that it is simply swept aside unnoticed by an independent Reynolds-stress-driven circulation? Does the upwelling dredge up a primordial magnetic field from beneath, and if so, does the shearing of that field in the tachocline react back on the rotation to produce a mid-latitude shear-free region? Does the dredged field stoke a (non)dynamo in the convection zone (*cf.* Byington *et al.*, 2012)? Does the equatorward tachocline flow in the polar regions and the poleward flow in the equatorial regions advect the periphery of the predominantly dipole remnant field in the radiative interior to align with the upwelling flow-convergence zone (Gough, 2012b), or do the torques from the rotational shear dominate the

dynamics, as Wood and McIntyre (in preparation) quite plausibly presume? Is such a putative inclined dipole responsible for the active longitudes? – and the emergence of sunspots? Most previous dynamical studies have presumed the tachocline flow to be essentially steady; the evidence presented recently by Antia, Chitre, and Gough (2012) suggests that it varies with the solar cycle. So these issues will need to be readdressed in a new light. The rise of sunspots through the convection is now becoming accessible to seismic observation (Ilonidis, Zhao, and Kosovichev, 2011). There are also questions related to structural changes associated with the cycle, and how they are related to the variations in the Sun's outputs. And what of the rotation of the solar core, and the associated circulation, if Ω differs substantially from the angular velocity of the surrounding envelope? That is likely to influence our ideas about other stars.

It must be appreciated that the Sun is an important benchmark for the whole of stellar physics, and helioseismology for asteroseismology. Finding answers to many of these questions is likely to be assisted by new helioseismological findings, although seismology alone will not be enough. Answers will not come easily – the easy questions were answered quickly (although not necessarily easily) in the early days. But the rewards from answering the new questions – designing subtle, hardly detectable, seismic diagnostics and developing techniques to analyse them – are potentially great. Their pursuit is to a large degree the task of the new generation of young seismologists.

Acknowledgements I am grateful to Jørgen Christensen-Dalsgaard, Werner Däppen, Günter Houdek, and Sasha Kosovichev for interesting discussion. I thank Jeannette Gilbert and Paula Younger for typing the first draft of this paper, Günter Houdek for providing Figure 4, and Amanda Smith for her help in producing Figures 1 and 6. I thank the Leverhulme Trust for an Emeritus Fellowship, and P.H. Scherrer for support from HMI NASA contract NAS5-02139. This article has benefited from comments on the first draft by the referee who raised the matter of the reliability of the equation of state.

References

Abraham, Z., Iben, I. Jr: 1971, More solar models and neutrino fluxes. *Astrophys. J.* **170**, 157. doi:10.1086/151197.

Allende Prieto, C., Lambert, D.L., Asplund, M.: 2001, The forbidden abundance of oxygen in the Sun. *Astrophys. J. Lett.* **556**, L63 – L66. doi:10.1086/322874.

Allende Prieto, C., Lambert, D.L., Asplund, M.: 2002, A reappraisal of the solar photospheric C/O ratio. *Astrophys. J. Lett.* **573**, L137 – L140. doi:10.1086/342095.

Amelin, Y., Krot, A.N., Hutcheon, I.D., Ulyanov, A.A.: 2002, Lead isotopic ages of chondrules and calcium-aluminium-rich inclusions. *Science* **297**, 1678 – 1683. doi:10.1126/science.1073950.

Ando, H., Osaki, Y.: 1975, Nonadiabatic nonradial oscillations – an application to the five-minute oscillation of the Sun. *Publ. Astron. Soc. Japan* **27**, 581 – 603.

Antia, H.M., Basu, S.: 2004, Temporal variations in the solar radius? In: Danesy, D. (ed.) *SOHO 14 Helio- and Asteroseismology: Towards a Golden Future* **SP-559**, ESA, Noordwijk, 301.

Antia, H.M., Basu, S.: 2006, Determining solar abundances using helioseismology. *Astrophys. J.* **644**, 1292 – 1298. doi:10.1086/503707.

Antia, H.M., Basu, S.: 2011, Are recent solar heavy element abundances consistent with helioseismology? *J. Phys. Conf. Ser.* **271**, 012034. doi:10.1088/1742-6596/271/1/012034.

Antia, H.M., Chitre, S.M., Gough, D.O.: 2008, Temporal variations in the Sun's rotational kinetic energy. *Astron. Astrophys.* **477**, 657 – 663. doi:10.1051/0004-6361:20078209.

Antia, H.M., Chitre, S.M., Gough, D.O.: 2012, On the magnetic field required for driving the observed angular-velocity variations in the solar convection zone. *Mon. Not. R. Astron. Soc.*, in press.

Asplund, M.: 2005, New light on stellar abundance analyses: Departures from LTE and homogeneity. *Annu. Rev. Astron. Astrophys.* **43**, 481 – 530. doi:10.1146/annurev.astro.42.053102.134001.

Asplund, M., Grevesse, N., Sauval, A.J.: 2005, The solar chemical composition. In: Barnes, T.G. III, Bash, F.N. (eds.) *Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis* **CS-336**, Astron. Soc. Pac., San Francisco, 25.

Asplund, M., Nordlund, Å., Trampedach, R., Stein, R.F.: 2000, Line formation in solar granulation. II. The photospheric Fe abundance. *Astron. Astrophys.* **359**, 743 – 754.

Asplund, M., Grevesse, N., Sauval, A.J., Allende Prieto, C., Kiselman, D.: 2004, Line formation in solar granulation. IV. [O I], O I and OH lines and the photospheric O abundance. *Astron. Astrophys.* **417**, 751 – 768. doi:10.1051/0004-6361:20034328.

Asplund, M., Grevesse, N., Sauval, A.J., Allende Prieto, C., Kiselman, D.: 2005a, Line formation in solar granulation. IV. [O I], O I and OH lines and the photospheric O abundance. *Astron. Astrophys.* **435**, 339 – 340. doi:10.1051/0004-6361:20034328e.

Asplund, M., Grevesse, N., Sauval, A.J., Allende Prieto, C., Blomme, R.: 2005b, Line formation in solar granulation. VI. [C I], C I, CH and C_2 lines and the photospheric C abundance. *Astron. Astrophys.* **431**, 693 – 705. doi:10.1051/0004-6361:20041951.

Asplund, M., Grevesse, N., Sauval, A.J., Scott, P.: 2009, The chemical composition of the Sun. *Annu. Rev. Astron. Astrophys.* **47**, 481 – 522. doi:10.1146/annurev.astro.46.060407.145222.

Bahcall, J.N.: 2001, High-energy physics: neutrinos reveal split personalities. *Nature* **412**, 29 – 31.

Bahcall, J.N., Basu, S., Pinsonneault, M.H.: 1998, How uncertain are solar neutrino predictions? *Phys. Lett. B* **433**, 1 – 8. doi:10.1016/S0370-2693(98)00657-1.

Bahcall, J.N., Pinsonneault, M.H.: 1992, Standard solar models, with and without helium diffusion, and the solar neutrino problem. *Rev. Mod. Phys.* **64**, 885 – 926. doi:10.1103/RevModPhys.64.885.

Bahcall, J.N., Pinsonneault, M.H., Basu, S.: 2001, Solar models: current epoch and time dependences, neutrinos, and helioseismological properties. *Astrophys. J.* **555**, 990 – 1012. doi:10.1086/321493.

Bahcall, J.N., Ulrich, R.K.: 1971, Solar neutrinos. III. Composition and magnetic-field effects and related inferences. *Astrophys. J.* **170**, 593. doi:10.1086/151245.

Bahcall, J.N., Ulrich, R.K.: 1988, Solar models, neutrino experiments, and helioseismology. *Rev. Mod. Phys.* **60**, 297 – 372. doi:10.1103/RevModPhys.60.297.

Bahcall, J.N., Basu, S., Pinsonneault, M.H., Serenelli, A.M.: 2005, Helioseismological implications of recent solar abundance determinations. *Astrophys. J.* **618**, 1049 – 1056. doi:10.1086/426070.

Balmforth, N.J., Gough, D.O.: 1990, Effluent stellar pulsation. *Astrophys. J.* **362**, 256 – 266. doi:10.1086/169262.

Basu, S.: 1998, Effects of errors in the solar radius on helioseismic inferences. *Mon. Not. Roy. Astron. Soc.* **298**, 719 – 728. doi:10.1046/j.1365-8711.1998.01690.x.

Basu, S., Antia, H.M.: 1995, Helium abundance in the solar envelope. *Mon. Not. Roy. Astron. Soc.* **276**, 1402 – 1408.

Basu, S., Antia, H.M.: 1997, Seismic measurement of the depth of the solar convection zone. *Mon. Not. Roy. Astron. Soc.* **287**, 189 – 198.

Basu, S., Antia, H.M.: 2003, Changes in solar dynamics from 1995 to 2002. *Astrophys. J.* **585**, 553 – 565. doi:10.1086/346020.

Basu, S., Antia, H.M.: 2004, Constraining solar abundances using helioseismology. *Astrophys. J. Lett.* **606**, L85 – L88. doi:10.1086/421110.

Basu, S., Antia, H.M.: 2008, Helioseismology and solar abundances. *Phys. Rep.* **457**, 217 – 283. doi:10.1016/j.physrep.2007.12.002.

Basu, S., Däppen, W., Nayfonov, A.: 1999, Helioseismic analysis of the hydrogen partition function in the solar interior. *Astrophys. J.* **518**, 985 – 993. doi:10.1086/307312.

Basu, S., Mandel, A.: 2004, Does solar structure vary with solar magnetic activity? *Astrophys. J. Lett.* **617**, L155 – L158. doi:10.1086/427435.

Basu, S., Christensen-Dalsgaard, J., Chaplin, W.J., Elsworth, Y., Isaak, G.R., New, R., Schou, J., Thompson, M.J., Tomczyk, S.: 1997, Solar internal sound speed as inferred from combined BiSON and LOWL oscillation frequencies. *Mon. Not. Roy. Astron. Soc.* **292**, 243.

Basu, S., Chaplin, W.J., Elsworth, Y., New, R., Serenelli, A.M., Verner, G.A.: 2007, Solar abundances and helioseismology: fine-structure spacings and separation ratios of low-degree p-modes. *Astrophys. J.* **655**, 660 – 671. doi:10.1086/509820.

Basu, S., Broomhall, A.-M., Chaplin, W.J., Elsworth, Y., Fletcher, S., New, R.: 2010, Differences between the current solar minimum and earlier minima. In: Cranmer, S.R., Hoeksema, J.T., Kohl, J.L. (eds.) *SOHO-23: Understanding a Peculiar Solar Minimum* **CS-428**, Astron. Soc. Pac., San Francisco, 37.

Baturin, V.A., Däppen, W., Gough, D.O., Vorontsov, S.V.: 2000, Seismology of the solar envelope: sound-speed gradient in the convection zone and its diagnosis of the equation of state. *Mon. Not. Roy. Astron. Soc.* **316**, 71 – 83. doi:10.1046/j.1365-8711.2000.03459.x.

Bonanno, A., Schlattl, H., Paternò, L.: 2002, The age of the Sun and the relativistic corrections in the EOS. *Astron. Astrophys.* **390**, 1115 – 1118. doi:10.1051/0004-6361:20020749.

Bouvier, A., Wadhwa, M.: 2010, The age of the Solar System redefined by the oldest Pb–Pb age of a meteoritic inclusion. *Nat. Geosci.* **3**, 637 – 641. doi:10.1038/ngeo941.

Brans, C., Dicke, R.H.: 1961, Mach's principle and a relativistic theory of gravitation. *Phys. Rev.* **124**, 925 – 935. doi:10.1103/PhysRev.124.925.

Bretherton, F.P., Spiegel, A.E.: 1968, The effect of the convection zone on solar spin-down. *Astrophys. J. Lett.* **153**, L77. doi:10.1086/180224.

Brown, T.M., Christensen-Dalsgaard, J., Dziembowski, W.A., Goode, P., Gough, D.O., Morrow, C.A.: 1989, Inferring the Sun's internal angular velocity from observed p-mode frequency splittings. *Astrophys. J.* **343**, 526 – 546. doi:10.1086/167727.

Byington, B.M., Brummell, N.H., Stone, J., Gough, D.O.: 2012, Stoked nondynamos: sustaining field in magnetically non-closed systems. *Phys. Fluids*. in preparation.

Caffau, E., Maiorca, E., Bonifacio, P., Faraggiana, R., Steffen, M., Ludwig, H.-G., Kamp, I., Busso, M.: 2009, The solar photospheric nitrogen abundance. Analysis of atomic transitions with 3D and 1D model atmospheres. *Astron. Astrophys.* **498**, 877 – 884. doi:10.1051/0004-6361/200810859.

Caffau, E., Ludwig, H.-G., Steffen, M., Freytag, B., Bonifacio, P.: 2011, Solar chemical abundances determined with a CO5BOLD 3D model atmosphere. *Solar Phys.* **268**, 255 – 269. doi:10.1007/s11207-010-9541-4.

Chaplin, W.J., Elsworth, Y., Howe, R., Isaak, G.R., McLeod, C.P., Miller, B.A., van der Raay, H.B., Wheeler, S.J., New, R.: 1996, BiSON performance. *Solar Phys.* **168**, 1 – 18. doi:10.1007/BF00145821.

Chaplin, W.J., Christensen-Dalsgaard, J., Elsworth, Y., Howe, R., Isaak, G.R., Larsen, R.M., New, R., Schou, J., Thompson, M.J., Tomczyk, S.,: 1999, Rotation of the solar core from BiSON and LOWL frequency observations. *Mon. Not. Roy. Astron. Soc.* **308**, 405 – 414.

Chaplin, W.J., Serenelli, A.M., Basu, S., Elsworth, Y., New, R., Verner, G.A.: 2007, Solar heavy-element abundance: constraints from frequency separation ratios of low-degree p-modes. *Astrophys. J.* **670**, 872 – 884. doi:10.1086/522578.

Christensen-Dalsgaard, J.: 1992, Solar models with enhanced energy transport in the core. *Astrophys. J.* **385**, 354 – 362. doi:10.1086/170944.

Christensen-Dalsgaard, J., Gough, D.O.: 1976, Towards a heliological inverse problem. *Nature* **259**, 89 – 92. doi:10.1038/259089a0.

Christensen-Dalsgaard, J., Gough, D.O.: 1980, Implications of the whole-disk Doppler observations of the Sun. In: Hill, H.A., Dziembowski, W.A. (eds.) *Nonradial and Nonlinear Stellar Pulsation*, *Lecture Notes in Phys.* **125**, Springer, Berlin, 184 – 190. doi:10.1007/3-540-09994-8_18.

Christensen-Dalsgaard, J., Gough, D.O.: 1981, Comparison of observed solar whole-disk oscillation frequencies with the predictions of a sequence of solar models. *Astron. Astrophys.* **104**, 173 – 176.

Christensen-Dalsgaard, J., Houdek, G.: 2010, Prospects for asteroseismology. *Astrophys. Space Sci.* **328**, 51 – 66. doi:10.1007/s10509-009-0227-z.

Christensen-Dalsgaard, J., Gough, D.O., Morgan, J.G.: 1979a, Dirty solar models. *Astron. Astrophys.* **73**, 121 – 128.

Christensen-Dalsgaard, J., Gough, D.O., Morgan, J.G.: 1979b, Erratum: "Dirty solar models" [Astron. Astrophys. **79**, 260]. *Astron. Astrophys.* **79**, 121 – 128.

Christensen-Dalsgaard, J., Gough, D.O., Thompson, M.J.: 1989, Differential asymptotic sound-speed inversions. *Mon. Not. Roy. Astron. Soc.* **238**, 481 – 502.

Christensen-Dalsgaard, J., Gough, D.O., Thompson, M.J.: 1991, The depth of the solar convection zone. *Astrophys. J.* **378**, 413 – 437. doi:10.1086/170441.

Christensen-Dalsgaard, J., Proffitt, C.R., Thompson, M.J.: 1993, Effects of diffusion on solar models and their oscillation frequencies. *Astrophys. J. Lett.* **403**, L75 – L78. doi:10.1086/186725.

Christensen-Dalsgaard, J., Duvall, T.L. Jr, Gough, D.O., Harvey, J.W., Rhodes, E.J. Jr: 1985, Speed of sound in the solar interior. *Nature* **315**, 378 – 382. doi:10.1038/315378a0.

Christensen-Dalsgaard, J., Däppen, W., Ajukov, S.V., Anderson, E.R., Antia, H.M., Basu, S., Baturin, V.A., Berthomieu, G., Chaboyer, B., Chitre, S.M., Cox, A.N., Demarque, P., Donatowicz, J., Dziembowski, W.A., Gabriel, M., Gough, D.O., Guenther, D.B., Guzik, J.A., Harvey, J.W., Hill, F., Houdek, G., Iglesias, C.A., Kosovichev, A.G., Leibacher, J.W., Morel, P., Proffitt, C.R., Provost, J., Reiter, J., Rhodes, E.J. Jr, Rogers, F.J., Roxburgh, I.W., Thompson, M.J., Ulrich, R.K.: 1996, The current state of solar modeling. *Science* **272**, 1286 – 1292. doi:10.1126/science.272.5266.1286.

Christensen-Dalsgaard, J., Di Mauro, M.P., Houdek, G., Pijpers, F.: 2009, On the opacity change required to compensate for the revised solar composition. *Astron. Astrophys.* **494**, 205 – 208. doi:10.1051/0004-6361:200810170.

Christensen-Dalsgaard, J., Monteiro, M.J.P.F.G., Rempel, M., Thompson, M.J.: 2011, A more realistic representation of overshoot at the base of the solar convective envelope as seen by helioseismology. *Mon. Not. Roy. Astron. Soc.* **414**, 1158 – 1174. doi:10.1111/j.1365-2966.2011.18460.x.

Claverie, A., Isaak, G.R., McLeod, C.P., van der Raay, H.B., Roca-Cortes, T.: 1980, The latest results of the velocity spectroscopy of the Sun. In: Hill, H.A., Dziembowski, W.A. (eds.) *Nonradial and Nonlinear Stellar Pulsation*, *Lecture Notes in Phys.* **125**, Springer, Berlin, 181. doi:10.1007/3-540-09994-8_17.

Cox, A.N., Morgan, S.M., Rogers, F.J., Iglesias, C.A.: 1992, An opacity mechanism for the pulsations of OB stars. *Astrophys. J.* **393**, 272 – 277. doi:10.1086/171504.

Däppen, W.: 1998, Microphysics: equation of state. *Space Sci. Rev.* **85**, 49 – 60. doi:10.1023/A:1005176317912.

Däppen, W.: 2004, Equations of state for solar and stellar modeling. In: Celebonovic, V., Gough, D., Däppen, W. (eds.) *Equation-of-State and Phase-Transition in Models of Ordinary Astrophysical Matter* **CS-731**, AIP, New York, 3 – 17. doi:10.1063/1.1828391.

Däppen, W.: 2007, Seismic abundance determination in the Sun and in stars. In: Stancliffe, R.J., Houdek, G., Martin, R.G., Tout, C.A. (eds.) *Unsolved Problems in Stellar Physics: A Conference in Honour of Douglas Gough* **CS-948**, AIP, New York, 179 – 190. doi:10.1063/1.2818968.

Däppen, W., Gough, D.O.: 1984, On the determination of the helium abundance of the solar convection zone. In: *Liège Internat. Astrophys. Coll.* **25**, Université de Liège, 264 – 268.

Däppen, W., Gilliland, R.L., Christensen-Dalsgaard, J.: 1986, Weakly interacting massive particles, solar neutrinos, and solar oscillations. *Nature* **321**, 229 – 231. doi:10.1038/321229a0.

Däppen, W., Gough, D.O., Thompson, M.J.: 1988, Further progress on the helium abundance determination. In: Rolfe, E.J. (ed.) *Seismology of the Sun and Sun-Like Stars* **SP-286**, ESA, Noordwijk, 505 – 510.

Däppen, W., Gough, D.O., Kosovichev, A.G., Thompson, M.J.: 1991, A new inversion for the hydrostatic stratification of the Sun. In: Gough, D., Toomre, J. (eds.) *Challenges to Theories of the Structure of Moderate-Mass Stars*, *Lecture Notes in Phys.* **388**, Springer, Berlin, 111. doi:10.1007/3-540-54420-8_57.

Deubner, F.-L.: 1975, Observations of low wavenumber nonradial eigenmodes of the Sun. *Astron. Astrophys.* **44**, 371 – 375.

Di Mauro, M.P., Christensen-Dalsgaard, J., Rabello-Soares, M.C., Basu, S.: 2002, Inferences on the solar envelope with high-degree modes. *Astron. Astrophys.* **384**, 666 – 677. doi:10.1051/0004-6361:20020020.

Dicke, R.H.: 1967, The solar spin-down problem. *Astrophys. J. Lett.* **149**, L121. doi:10.1086/180072.

Dicke, R.H.: 1970, The solar oblateness and the gravitational quadrupole moment. *Astrophys. J.* **159**, 1. doi:10.1086/150286.

Dicke, R.H., Goldenberg, H.M.: 1967, Solar oblateness and general relativity. *Phys. Rev. Lett.* **18**, 313 – 316. doi:10.1103/PhysRevLett.18.313.

Dicke, R.H., Goldenberg, H.M.: 1974, The oblateness of the Sun. *Astrophys. J. Suppl. Ser.* **27**, 131. doi:10.1086/190292.

Doğan, G., Bonanno, A., Christensen-Dalsgaard, J.: 2010, Near-surface effects and solar-age determination. arXiv:1004.2215.

Drake, J.J., Testa, P.: 2005, The 'solar model problem' solved by the abundance of neon in nearby stars. *Nature* **436**, 525 – 528. doi:10.1038/nature03803.

Duvall, T.L. Jr., Harvey, J.W.: 1983, Observations of solar oscillations of low and intermediate degree. *Nature* **302**, 24 – 27. doi:10.1038/302024a0.

Duvall, T.L. Jr., Dziembowski, W.A., Goode, P.R., Gough, D.O., Harvey, J.W., Leibacher, J.W.: 1984, Internal rotation of the Sun. *Nature* **310**, 22 – 25. doi:10.1038/310022a0.

Dziembowski, W.A., Goode, P.R.: 2004, Helioseismic probing of solar variability: the formalism and simple assessments. *Astrophys. J.* **600**, 464 – 479. doi:10.1086/379708.

Dziembowski, W.A., Goode, P.R.: 2005, Sources of oscillation frequency increase with rising solar activity. *Astrophys. J.* **625**, 548 – 555. doi:10.1086/429712.

Dziembowski, W.A., Moskalik, P., Pamyatnykh, A.A.: 1993, The opacity mechanism in b-type stars – part two – excitation of high-order G-modes in main sequence stars. *Mon. Not. Roy. Astron. Soc.* **265**, 588.

Dziembowski, W.A., Moskalik, P., Pamyatnykh, A.A.: 1994, G-mode instability in the main sequence B-type stars. In: Balona, L.A., Henrichs, H.F., Le Contel, J.M. (eds.) *Pulsation; Rotation; and Mass Loss in Early-Type Stars*, *IAU Symp.* **162**, Kluwer, Dordrecht, 69.

Dziembowski, W.A., Pamyatnykh, A.A.: 1993, The opacity mechanism in B-type stars. I – Unstable modes in Beta Cephei star models. *Mon. Not. Roy. Astron. Soc.* **262**, 204 – 212.

Dziembowski, W.A., Pamyatnykh, A.A., Sienkiewicz, R.: 1990, Solar model from helioseismology and the neutrino flux problem. *Mon. Not. Roy. Astron. Soc.* **244**, 542 – 550.

Dziembowski, W.A., Fiorentini, G., Ricci, B., Sienkiewicz, R.: 1999, Helioseismology and the solar age. *Astron. Astrophys.* **343**, 990 – 996.

Elliott, J.R.: 1995, Opacity determination in the solar radiative interior. *Mon. Not. Roy. Astron. Soc.* **277**, 1567.

Elliott, J.R., Gough, D.O., Sekii, T.: 1998, Helioseismic determination of the solar tachocline thickness. In: Korzennik, S. (ed.) *Structure and Dynamics of the Interior of the Sun and Sun-like Stars* **SP-418**, ESA, Noordwijk, 763.

Elsworth, Y., Howe, R., Isaak, G.R., McLeod, C.P., Miller, B.A., New, R., Wheeler, S.J., Gough, D.O.: 1995, Slow rotation of the Sun's interior. *Nature* **376**, 669 – 672.

Emilio, M., Bush, R.I., Kuhn, J., Scherrer, P.: 2007, A changing solar shape. *Astrophys. J. Lett.* **660**, L161 – L163. doi:10.1086/518212.

Faulkner, J., Gough, D.O., Vahia, M.N.: 1986, Weakly interacting massive particles and solar oscillations. *Nature* **321**, 226 – 229. doi:10.1038/321226a0.

Fivian, M.D., Hudson, H.S., Lin, R.P., Zahid, H.J.: 2008, A large excess in apparent solar oblateness due to surface magnetism. *Science* **322**, 560 – 562. doi:10.1126/science.1160863.

Fossat, E., Grec, G., Pomerantz, M.: 1981, Solar pulsations observed from the geographic South Pole – Initial results. *Solar Phys.* **74**, 59 – 63. doi:10.1007/BF00151274.

Foukal, P., Fröhlich, C., Spruit, H., Wigley, T.M.L.: 2006, Variations in solar luminosity and their effect on the Earth's climate. *Nature* **443**, 161 – 166. doi:10.1038/nature05072.

Fröhlich, C.: 2011, Total solar irradiance: what have we learned from the last three cycles and the recent minimum? *Space Sci. Rev.* doi:10.1007/s11214-011-9780-1.

Gilliland, R.L., Faulkner, J., Press, W.H., Spergel, D.N.: 1986, Solar models with energy transport by weakly interacting particles. *Astrophys. J.* **306**, 703 – 709. doi:10.1086/164380.

Gizon, L., Schunker, H., Baldner, C.S., Basu, S., Birch, A.C., Bogart, R.S., Braun, D.C., Cameron, R., Duvall, T.L., Hanasoge, S.M., Jackiewicz, J., Roth, M., Stahn, T., Thompson, M.J., Zharkov, S.: 2009, Helioseismology of sunspots: a case study of NOAA region 9787. *Space Sci. Rev.* **144**, 249 – 273. doi:10.1007/s11214-008-9466-5.

Goode, P.R., Didkovsky, L.V., Libbrecht, K.G., Woodard, M.F.: 2002, Evolution of the Sun's near-surface asphericities over the activity cycle. *Adv. Space Res.* **29**, 1889 – 1898. doi:10.1016/S0273-1177(02)00240-5.

Gough, D., Hindman, B.W.: 2010, Helioseismic detection of deep meridional flow. *Astrophys. J.* **714**, 960 – 970. doi:10.1088/0004-637X/714/1/960.

Gough, D.O.: 1977, Random remarks on solar hydrodynamics. In: Bonnet, R.M., Delache, P. (eds.) *The Energy Balance and Hydrodynamics of the Solar Chromosphere and Corona*, *IAU Colloq.* **36**, G. de Bussac, Clermont-Ferrand, 3 – 36.

Gough, D.O.: 1978, The relevance of solar oscillations to theories of the rotation of the Sun. In: Belvedere, G., Paternò, L. (eds.) *Proc. EPS Workshop on Solar Rotation*, Catania Univ., 87 – 103.

Gough, D.O.: 1981, A new measure of the solar rotation. *Mon. Not. Roy. Astron. Soc.* **196**, 731 – 745.

Gough, D.O.: 1982a, A review of the theory of solar oscillations and its implications concerning the internal structure of the Sun. In: Cox, J.P., Hansen, C.J. (eds.) *Pulsations in Classical and Cataclysmic Variable Stars*, JILA, Boulder, 117.

Gough, D.O.: 1982b, Inferences from solar oscillations. *Ir. Astron. J.* **15**, 118 – 119.

Gough, D.O.: 1983a, Our first inferences from helioseismology. *Phys. Bull.* **34**, 502 – 507.

Gough, D.O.: 1983b, Solar structure: a bridge in a gap in solar oscillations. *Nature* **302**, 18. doi:10.1038/302018a0.

Gough, D.O.: 1983c, The protosolar helium abundance. In: Shaver, P.A., Kunth, D., Kjar, K. (eds.) *Workshop on the Primordial Helium (A83-50030 24-90)*, ESO, Garching, 117 – 136.

Gough, D.O.: 1984a, Helioseismology. *Observatory* **104**, 118 – 119.

Gough, D.O.: 1984b, Towards a solar model. *Mem. Soc. Astron. Ital.* **55**, 13.

Gough, D.O.: 1986, EBK quantization of stellar waves. In: Osaki, Y. (ed.) *Hydrodynamic and Magnetodynamic Problems in the Sun and Stars*, Univ. Tokyo, 117 – 143.

Gough, D.O.: 1990, The internal structure of late-type main-sequence stars. In: Gustafsson, B., Nissen, P.E. (eds.) *Astrophysics: Recent Progress and Future Possibilities (A91-15054 03-90)*, Kongelige Danske Videnskabernes Selskab, Copenhagen, 13 – 50.

Gough, D.O.: 1995, Prospects for asteroseismic inference. In: Ulrich, R.K., Rhodes, E.J. Jr., Däppen, W. (eds.) *GONG 1994: Helio- and Astro-Seismology from the Earth and Space* **CS-76**, Astron. Soc. Pac., San Francisco, 551.

Gough, D.O.: 2004, The power of helioseismology to address issues of fundamental physics. In: Celebonović, V., Gough, D., Däppen, W. (eds.) *Equation-of-State and Phase-Transition in Models of Ordinary Astrophysical Matter* **CS-731**, AIP, New York, 119 – 138. doi:10.1063/1.1828398.

Gough, D.O.: 2006, Helioseismological determination of the state of the solar interior. In: Lacoste, H., Ouwehand, L. (eds.) *SOHO-17: 10 Years of SOHO and Beyond* **SP-617**, ESA, Noordwijk, 1 – 17.

Gough, D.O.: 2012a, Heliophysics gleaned from seismology. In: Shibahashi, H., Takata, M. (eds.) *Progress in Solar/Stellar Physics with Helio- and Asteroseismology*; *Proc. 65th Fujihara Seminar*, Astron. Soc. Pac., San Francisco, in press.

Gough, D.O.: 2012b, Pattern formation in rapidly oscillating peculiar A stars. *Geophys. Astrophys. Fluid Dyn.* **106**, 429 – 449.

Gough, D.O.: 2012c, How oblate is the Sun? *Science*. in press

Gough, D.O., Kosovichev, A.G.: 1988, An attempt to understand the Stanford p-mode data. In: Rolfe, E.J. (ed.) *Seismology of the Sun and Sun-Like Stars* **SP-286**, ESA, Noordwijk, 195 – 201.

Gough, D.O., Kosovichev, A.G.: 1990, Using helioseismic data to probe the hydrogen abundance in the solar core. In: Berthomieu, G., Cribier, M. (eds.) *Inside the Sun, IAU Colloq.* **121**, Kluwer, Dordrecht, 327. Astrophys. Space Sci. Lib. **159**.

Gough, D.O., McIntyre, M.E.: 1998, Inevitability of a magnetic field in the Sun's radiative interior. *Nature* **394**, 755 – 757. doi:10.1038/29472.

Gough, D.O., Scherrer, P.H.: 2002, The solar interior. In: Bleeker, J.A., Geiss, J., Huber, M.C.E. (eds.) *The Century of Space Science* **I**. Kluwer, Dordrecht, 1035.

Gough, D.O., Sekii, T., Stark, P.B.: 1996, Inferring spatial variation of solar properties from helioseismic data. *Astrophys. J.* **459**, 779. doi:10.1086/176942.

Gough, D.O., Kosovichev, A.G., Toomre, J., Anderson, E., Antia, H.M., Basu, S., Chaboyer, B., Chitre, S.M., Christensen-Dalsgaard, J., Dziembowski, W.A., Eff-Darwich, A., Elliott, J.R., Giles, P.M., Goode, P.R., Guzik, J.A., Harvey, J.W., Hill, F., Leibacher, J.W., Monteiro, M.J.P.F.G., Richard, O., Sekii, T., Shibahashi, H., Takata, M., Thompson, M.J., Vauclair, S., Vorontsov, S.V.: 1996, The seismic structure of the Sun. *Science* **272**, 1296 – 1300. doi:10.1126/science.272.5266.1296.

Grec, G., Fossat, E., Pomerantz, M.: 1980, Solar oscillations – Full disk observations from the geographic South Pole. *Nature* **288**, 541 – 544. doi:10.1038/288541a0.

Grevesse, N., Noels, A.: 1993, Cosmic abundances of the elements. In: Prantzos, N., Vangioni-Flam, E., Casse, M. (eds.) *Origin and Evolution of the Elements*, Cambridge Univ. Press, Cambridge, 15 – 25.

Grevesse, N., Sauval, A.J.: 1998, Standard solar composition. *Space Sci. Rev.* **85**, 161 – 174. doi:10.1023/A:1005161325181.

Grevesse, N., Asplund, M., Sauval, A.J., Scott, P.: 2011, The new solar composition and the solar metallicity. In: Miralles, M.P., Sánchez Almeida, J. (eds.) *The Sun, the Solar Wind, and the Heliosphere*, Springer, Berlin, 51.

Guzik, J.A., Watson, L.S., Cox, A.N.: 2005, Can enhanced diffusion improve helioseismic agreement for solar models with revised abundances? *Astrophys. J.* **627**, 1049 – 1056. doi:10.1086/430438.

Guzik, J.A., Watson, L.S., Cox, A.N.: 2006, Implications of revised solar abundances for helioseismology *Mem. Soc. Astron. Ital.* **77**, 389.

Hampel, W., Heusser, G., Kiko, J., Kirsten, T., Laubenstein, M., Pernicka, E., Rau, W., Rönn, U., Schlosser, C., Wojcik, M., Zakharov, Y., v. Ammon, R., Ebert, K.H., Fritsch, T., Heidt, D., Henrich, E., Stieglitz, L., Weirich, F., Balata, M., Sann, M., Hartmann, F.X., Bellotti, E., Cattadori, C., Cremonesi, O., Ferrari, N., Fiorini, E., Zanotti, L., Altmann, M., v. Feilitzsch, F., Mößbauer, R., Berthomieu, G., Schatzman, E., Carmi, I., Dostrovsky, I., Bacci, C., Belli, P., Bernabei, R., D'Angelo, S., Paoluzi, L., Bevilacqua, A., Cribier, M., Gosset, L., Rich, J., Spiro, M., Tao, C., Vignaud, D., Boger, J., Hahn, R.L., Rowley, J.K., Stoenner, R.W., Weneser, J.: 1996, GALLEX solar neutrino observations: Results for GALLEX III. *Phys. Lett. B* **388**, 384 – 396. doi:10.1016/S0370-2693(96)01121-5.

Hindman, B.W., Haber, D.A., Toomre, J.: 2009, Subsurface circulations within active regions. *Astrophys. J.* **698**, 1749 – 1760. doi:10.1088/0004-637X/698/2/1749.

Houdek, G., Gough, D.O.: 2007, An asteroseismic signature of helium ionisation. *Mon. Not. Roy. Astron. Soc.* **375**, 861 – 880. doi:10.1111/j.1365-2966.2006.11325.x.

Houdek, G., Gough, D.O.: 2011, On the seismic age and heavy-element abundance of the Sun. *Mon. Not. Roy. Astron. Soc.* **418**, 1217 – 1230. doi:10.1111/j.1365-2966.2011.19572.x.

Howard, L.N., Moore, D.W., Spiegel, E.A.: 1967, *Nature* **214**, 1297.

Howe, R., Thompson, M.J.: 1996, On the use of the error correlation function in helioseismic inversions. *Mon. Not. Roy. Astron. Soc.* **281**, 1385.

Howe, R., Komm, R., Hill, F., Ulrich, R., Haber, D.A., Hindman, B.W., Schou, J., Thompson, M.J.: 2006a, Large-scale zonal flows near the solar surface. *Solar Phys.* **235**, 1 – 15. doi:10.1007/s11207-006-0117-2.

Howe, R., Rempel, M., Christensen-Dalsgaard, J., Hill, F., Komm, R., Larsen, R.M., Schou, J., Thompson, M.J.: 2006b, Solar convection zone dynamics: how sensitive are inversions to subtle dynamo features? *Astrophys. J.* **649**, 1155 – 1168. doi:10.1086/506931.

Iglesias, C.A., Rogers, F.J.: 1991, Opacities for the solar radiative interior. *Astrophys. J.* **371**, 408 – 417. doi:10.1086/169902.

Iglesias, C.A., Rogers, F.J.: 1996, Updated opal opacities. *Astrophys. J.* **464**, 943. doi:10.1086/177381.

Iglesias, C.A., Rogers, F.J., Wilson, B.G.: 1990, Opacities for classical Cepheid models. *Astrophys. J.* **360**, 221 – 226. doi:10.1086/169110.

Ilonidis, S., Zhao, J., Kosovichev, A.: 2011, Detection of emerging sunspot regions in the solar interior. *Science* **333**, 993. doi:10.1126/science.1206253.

Jacobsen, B., Yin, Q.-Z., Moynier, F., Amelin, Y., Krot, A.N., Nagashima, K., Hutcheon, I.D., Palme, H.: 2008, ^{26}Al ^{26}Mg and ^{207}Pb ^{206}Pb systematics of Allende CAIs: canonical solar initial ^{26}Al/^{27}Al ratio reinstated. *Earth Planet. Sci. Lett.* **272**, 353 – 364. doi:10.1016/j.epsl.2008.05.003.

Jacobsen, B., Yin, Q.-Z., Moynier, F., Amelin, Y., Krot, A.N., Nagashima, K., Hutcheon, I.D., Palme, H.: 2009, Erratum to "^{26}Al ^{26}Mg and ^{207}Pb ^{206}Pb systematics of Allende CAIs: Canonical solar initial

^{26}Al/^{27}Al ratio reinstated" [*Earth Planet Sci. Lett.* **272** (2008) 353 – 364]. *Earth Planet. Sci. Lett.* **277**, 549. doi:10.1016/j.epsl.2008.12.001.

Kiriakidis, M., El Eid, M.F., Glatzel, W.: 1992, Heavy element opacities and the pulsations of Beta Cepheid stars. *Mon. Not. Roy. Astron. Soc.* **255**, 1P – 5P.

Kitiashvili, I.N., Kosovichev, A.G., Wray, A.A., Mansour, N.N.: 2009, Traveling waves of magnetoconvection and the origin of the Evershed effect in sunspots. *Astrophys. J. Lett.* **700**, L178 – L181. doi:10.1088/0004-637X/700/2/L178.

Kitiashvili, I.N., Bellot Rubio, L.R., Kosovichev, A.G., Mansour, N.N., Sainz Dalda, A., Wray, A.A.: 2010, Explanation of the sea-serpent magnetic structure of sunspot penumbrae. *Astrophys. J. Lett.* **716**, L181 – L184. doi:10.1088/2041-8205/716/2/L181.

Komm, R., Howe, R., Hill, F.: 2006, Helioseismic sensing of the solar cycle. *Adv. Space Res.* **38**, 845 – 855. doi:10.1016/j.asr.2005.07.034.

Kopp, G., Lean, J.L.: 2011, A new, lower value of total solar irradiance: evidence and climate significance. *Geophys. Res. Lett.* **38**, 1706. doi:10.1029/2010GL045777.

Korzennik, S.G., Ulrich, R.K.: 1989, Seismic analysis of the solar interior. I – Can opacity changes improve the theoretical frequencies? *Astrophys. J.* **339**, 1144 – 1155. doi:10.1086/167369.

Kosovichev, A.G.: 2009, Photospheric and subphotospheric dynamics of emerging magnetic flux. *Space Sci. Rev.* **144**, 175 – 195. doi:10.1007/s11214-009-9487-8.

Kosovichev, A.G., Duvall, T.L. Jr: 2011, Investigation of a sunspot complex by helioseismology. In: Choudhary, D.P., Strassmeier, K.G. (eds.) *The Physics of Sun and Star Spots*, *IAU Symp.* **273**, Cambridge Univ. Press, Cambridge, 320 – 324. doi:10.1017/S1743921311015456.

Kosovichev, A.G., Christensen-Dalsgaard, J., Däppen, W., Dziembowski, W.A., Gough, D.O., Thompson, M.J.: 1992, Sources of uncertainty in direct seismological measurements of the solar helium abundance. *Mon. Not. Roy. Astron. Soc.* **259**, 536 – 558.

Kuhn, J.R., Bush, R., Emilio, M., Scholl, I.F.: 2012, The precise solar shape and its variability. *Science*. doi:10.1126/science.1223231.

Lefebvre, S., Kosovichev, A.G.: 2005, Changes in the subsurface stratification of the Sun with the 11-year activity cycle. *Astrophys. J. Lett.* **633**, L149 – L152. doi:10.1086/498305.

Lefebvre, S., Kosovichev, A.G., Rozelot, J.P.: 2007, Helioseismic test of nonhomologous solar radius changes with the 11 year activity cycle. *Astrophys. J. Lett.* **658**, L135 – L138. doi:10.1086/515394.

Lodders, K., Palme, H., Gail, H.-P.: 2009, Abundances of the elements in the solar system. In: Trümper, J.E. (ed.) *Landolt-Börnstein – Group VI Astron. Astrophys. Numerical Data and Functional Relationships in Science and Technology Volume*, Springer, Berlin, 44. doi:10.1007/978-3-540-88055-4_34.

Lopes, I.P., Gough, D.O.: 2001, Seismology of stellar envelopes: probing the outer layers of a star through the scattering of acoustic waves. *Mon. Not. Roy. Astron. Soc.* **322**, 473 – 485. doi:10.1046/j.1365-8711.2001.03940.x.

Moskalik, P., Buchler, J.R., Marom, A.: 1992, Toward a resolution of the bump and beat Cepheid mass discrepancies. *Astrophys. J.* **385**, 685 – 693. doi:10.1086/170975.

Moskalik, P., Dziembowski, W.A.: 1992, New opacities and the origin of the Beta Cephei pulsation. *Astron. Astrophys.* **256**, L5 – L8.

Mussack, K., Däppen, W.: 2010, Dynamic screening in solar and stellar nuclear reactions. *Astrophys. Space Sci.* **328**, 153 – 156. doi:10.1007/s10509-009-0245-x.

Mussack, K., Däppen, W.: 2011, Dynamic screening correction for solar p–p reaction rates. *Astrophys. J.* **729**, 96. doi:10.1088/0004-637X/729/2/96.

Mussack, K., Gough, D.O.: 2009, Measuring solar abundances with seismology. In: Dikpati, M., Arentoft, T., González Hernández, I., Lindsey, C., Hill, F. (eds.) *Solar-Stellar Dynamos as Revealed by Helio- and Asteroseismology: GONG 2008/SOHO 21* **CS-416**, Astron. Soc. Pac., San Francisco, 203.

Pijpers, F.P.: 1998, Helioseismic determination of the solar gravitational quadrupole moment. *Mon. Not. Roy. Astron. Soc.* **297**, L76 – L80. doi:10.1046/j.1365-8711.1998.01801.x.

Ramsey, A.S.: 1937, *Dynamics Part II*, Cambridge Univ. Press, Cambridge.

Richard, O., Dziembowski, W.A., Sienkiewicz, R., Goode, P.R.: 1998, Precise determination of the solar helium abundance by helioseismology. In: Korzennik, S. (ed.) *Structure and Dynamics of the Interior of the Sun and Sun-Like Stars* **SP-418**, ESA, Noordwijk, 517.

Rogers, F.J., Nayfonov, A.: 2002, Updated and expanded OPAL equation-of-state tables: implications for helioseismology. *Astrophys. J.* **576**, 1064 – 1074. doi:10.1086/341894.

Saio, H.: 1992, Opacity in the solar radiative interior inferred from 5-min oscillations. *Mon. Not. Roy. Astron. Soc.* **258**, 491 – 496.

Salpeter, E.E.: 1954, Electron screening and thermonuclear reactions. *Aust. J. Phys.* **7**, 373.

Schou, J., Woodard, M.F., Birch, A.C.: 2009, Large-scale flows from eigenfunction fitting. *Bull. Am. Astron. Soc.* **40**, 07.05.

Schou, J., Antia, H.M., Basu, S., Bogart, R.S., Bush, R.I., Chitre, S.M., Christensen-Dalsgaard, J., di Mauro, M.P., Dziembowski, W.A., Eff-Darwich, A., Gough, D.O., Haber, D.A., Hoeksema, J.T., Howe, R., Korzennik, S.G., Kosovichev, A.G., Larsen, R.M., Pijpers, F.P., Scherrer, P.H., Sekii, T., Tarbell, T.D., Title, A.M., Thompson, M.J., Toomre, J.: 1998, Helioseismic studies of differential rotation in the solar envelope by the solar oscillations investigation using the Michelson Doppler Imager. *Astrophys. J.* **505**, 390 – 417. doi:10.1086/306146.

Simon, N.R.: 1982, A plea for reexamining heavy element opacities in stars. *Astrophys. J. Lett.* **260**, L87 – L90. doi:10.1086/183876.

Spergel, D.N., Press, W.H.: 1985, Effect of hypothetical, weakly interacting, massive particles on energy transport in the solar interior. *Astrophys. J.* **294**, 663 – 673. doi:10.1086/163336.

Spiegel, E.A., Zahn, J.-P.: 1992, The solar tachocline. *Astron. Astrophys.* **265**, 106 – 114.

Steigman, G.: 2007, Primordial nucleosynthesis in the precision cosmology era. *Annu. Rev. Nucl. Part. Sci.* **57**, 463 – 491. doi:10.1146/annurev.nucl.56.080805.140437.

Stein, R.F., Nordlund, A.: 1998, Simulations of solar granulation. I. General properties. *Astrophys. J.* **499**, 914. doi:10.1086/305678.

Takata, M., Gough, D.O.: 2001, The influence of uncertainties in the Sun's radius on inversions for the solar structure. In: Wilson, A., Pallé, P.L. (eds.) *SOHO 10/GONG 2000 Workshop: Helio- and Asteroseismology at the Dawn of the Millennium* **SP-464**, ESA, Noordwijk, 543 – 546.

Takata, M., Gough, D.O.: 2003, The seismic radius of the Sun, and structure inversions. In: Sawaya-Lacoste, H. (ed.) *GONG+ 2002. Local and Global Helioseismology: The Present and Future* **SP-517**, ESA, Noordwijk, 397 – 400.

Tassoul, M.: 1980, Asymptotic approximations for stellar nonradial pulsations. *Astrophys. J. Suppl. Ser.* **43**, 469 – 490. doi:10.1086/190678.

Taylor, S.F., Varsik, J.R., Woodard, M.F., Libbrecht, K.G.: 1998, Spatial dependence of solar-cycle changes in the Sun's luminosity. *Solar Phys.* **178**, 1 – 12.

Tomczyk, S., Schou, J., Thompson, M.J.: 1995, Measurement of the rotation rate in the deep solar interior. *Astrophys. J. Lett.* **448**, L57. doi:10.1086/309598.

Tomczyk, S., Streander, K., Card, G., Elmore, D., Hull, H., Cacciani, A.: 1995, An instrument to observe low-degree solar oscillations. *Solar Phys.* **159**, 1 – 21. doi:10.1007/BF00733027.

Tripathy, S.C., Christensen-Dalsgaard, J.: 1998, Opacity effects on the solar interior. I. Solar structure. *Astron. Astrophys.* **337**, 579 – 590.

Tripathy, S.C., Basu, S., Christensen-Dalsgaard, J.: 1998, Helioseismic determination of opacity corrections. In: Provost, J., Schmider, F.-X. (eds.) *Sounding Solar and Stellar Interiors*, *IAU Symp.* **181**, Poster volume, Université de Nice, Côte d'Azur, 129 – 130.

Turck-Chièze, S., Däppen, W., Fossat, E., Provost, J., Schatzman, E., Vignaud, D.: 1993, The solar interior. *Phys. Rep.* **230**, 57 – 235. doi:10.1016/0370-1573(93)90020-E.

Ulrich, R.K., Rhodes, E.J. Jr.: 1977, The sensitivity of nonradial P mode eigenfrequencies to solar envelope structure. *Astrophys. J.* **218**, 521 – 529. doi:10.1086/155705.

Vandakurov, Y.V.: 1967, The frequency distribution of stellar oscillations *Astron. Ž.* **44**, 786.

Verner, G.A., Chaplin, W.J., Elsworth, Y.: 2006, BiSON data show change in solar structure with magnetic activity. *Astrophys. J. Lett.* **640**, L95 – L98. doi:10.1086/503101.

Vorontsov, S.V., Baturin, V.A., Pamyatnykh, A.A.: 1992, Seismology of the solar envelope – towards the calibration of the equation of state. *Mon. Not. Roy. Astron. Soc.* **257**, 32 – 46.

Vorontsov, S.V., Christensen-Dalsgaard, J., Schou, J., Strakhov, V.N., Thompson, M.J.: 2002, Helioseismic measurement of solar torsional oscillations. *Science* **296**, 101 – 103. doi:10.1126/science.1069190.

Weiss, N.O., Thomas, J.H., Brummell, N.H., Tobias, S.M.: 2004, The origin of penumbral structure in sunspots: downward pumping of magnetic flux. *Astrophys. J.* **600**, 1073 – 1090. doi:10.1086/380091.

Willson, R.C., Hudson, H.S.: 1988, Solar luminosity variations in solar cycle 21. *Nature* **332**, 810 – 812. doi:10.1038/332810a0.

Zhao, J., Kosovichev, A.G.: 2003, Helioseismic observation of the structure and dynamics of a rotating sunspot beneath the solar surface. *Astrophys. J.* **591**, 446 – 453. doi:10.1086/375343.

Zhao, J., Kosovichev, A.G., Duvall, T.L. Jr.: 2001, Investigation of mass flows beneath a sunspot by time-distance helioseismology. *Astrophys. J.* **557**, 384 – 388. doi:10.1086/321491.

Zhao, J., Kosovichev, A.G., Sekii, T.: 2010, High-resolution helioseismic imaging of subsurface structures and flows of a solar active region observed by Hinode. *Astrophys. J.* **708**, 304 – 313. doi:10.1088/0004-637X/708/1/304.

Zhao, J., Couvidat, S., Bogart, R.S., Parchevsky, K.V., Birch, A.C., Duvall, T.L., Beck, J.G., Kosovichev, A.G., Scherrer, P.H.: 2011, Time-distance helioseismology data-analysis pipeline for Helioseismic and Magnetic Imager onboard Solar Dynamics Observatory (SDO/HMI) and its initial results. *Solar Phys.* **275**, 375 – 390. doi:10.1007/s11207-011-9757-y.

Zweibel, E.G., Gough, D.O.: 1995, Is there a seismic signature of the Sun's magnetic field? In: Hoeksema, J.T., Domingo, V., Fleck, B., Battrick, B. (eds.) *Helioseismology* **SP-376**, ESA, Noordwijk, 73.

DOI 10.1007/978-1-4899-8005-2_3
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0048-z

The Dynamics of the Solar Radiative Zone

A. Eff-Darwich · S.G. Korzennik

Received: 13 February 2012 / Accepted: 1 June 2012 / Published online: 3 July 2012

Abstract The dynamics of the solar radiative interior are still poorly constrained by comparison to the convective zone. This disparity is even more marked when we attempt to derive meaningful temporal variations. Many data sets contain a small number of modes that are sensitive to the inner layers of the Sun, but we found that the estimates of their uncertainties are often inaccurate. As a result, these data sets allow us to obtain, at best, a low-resolution estimate of the solar-core rotation rate down to approximately $0.2R_{\odot}$. We present inferences based on mode determination resulting from an alternate peak-fitting methodology aimed at increasing the amount of observed modes that are sensitive to the radiative zone, while special care was taken in the determination of their uncertainties. This methodology has been applied to MDI and GONG data, for the whole Solar Cycle 23, and to the newly available HMI data. The numerical inversions of all these data sets result in the best inferences to date of the rotation in the radiative region. These results and the method used to obtain them are discussed. The resulting profiles are shown and analyzed, and the significance of the detected changes is discussed.

Keywords Helioseismology · Inverse modeling · Observations · Interior · Radiative zone

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

A. Eff-Darwich
Dept. Edafologia y Geologia, Univ. La Laguna, 38206, Tenerife, Spain
e-mail: adarwich@ull.es

A. Eff-Darwich
Instituto de Astrofisica de Canarias, 38205, Tenerife, Spain
e-mail: adarwich@iac.es

S.G. Korzennik (✉)
Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA
e-mail: skorzennik@cfa.harvard.edu

1. Introduction

Ground-based helioseismic observations (*e.g.* GONG: Harvey *et al.*, 1996; BiSON: Broomhall *et al.*, 2009) and space-based ones (*e.g.* MDI: Scherrer *et al.*, 1995; GOLF: Gabriel *et al.*, 1995; or HMI: Scherrer *et al.*, 2012), have allowed us to derive a good description of the dynamics of the solar interior (*e.g.* Eff-Darwich, Korzennik, and Jiménez-Reyes, 2002; Thompson *et al.*, 2003; García *et al.*, 2007; Eff-Darwich *et al.*, 2008; Howe, 2009). Helioseismic inferences have confirmed that the differential rotation observed at the surface persists throughout the convection zone. The outer radiative zone ($0.3 < r/R_\odot < 0.7$) appears to rotate approximately as a solid body at an almost constant rate ($\approx$430 nHz), whereas it is not possible to rule out a different rotation rate for the innermost core ($0.19 < r/R_\odot < 0.3$). At the base of the convection zone, a shear layer – known as the tachocline – separates the region of differential rotation throughout the convection zone from the one with rigid rotation in the radiative zone. Finally, there is a subsurface shear layer between the fastest-rotating layer, located at about $0.95R_\odot$, and the surface. Of course, this rotation profile is not constant; the time-varying component of the rotation displays clear variations near the surface (known as the torsional oscillations), while we see hints of variations at the base of the convection zone, both being likely related to the driving mechanisms of the solar-activity cycle.

Our understanding of the dynamics of the solar interior has undoubtedly improved; however, we still need to constrain the rotation profile near the core and fully analyze the nature of the torsional oscillations. We still do not know how thin the tachocline really is and what is keeping it this way. Understanding the tachocline should help discern if there is a fossil magnetic field in the radiative zone that prevents the spread of the tachocline (Zahn, Brun, and Mathis, 2007), or an oscillating magnetic field (Forgács-Dajka and Petrovay, 2001). No purely fluid-dynamics mechanism can explain the tachocline, resulting in a compelling argument for the presence of a strong magnetic field (Gough and McIntyre, 1998).

The proper knowledge of the relationship between the solar dynamics and its structure is not important only in order to understand the present conditions of the Sun, but also to understand the temporal evolution of our star and other solar-like stars. It is usually assumed that the main characteristics of the dynamics of the Sun were established during its contraction phase (Turck-Chièze *et al.*, 2010), hence the Sun was not a rapid rotator when it entered the Zero Age Main Sequence (ZAMS). The transport of momentum during the contraction phase might have been carried out by a magnetic field in the core and the diffusion of this field flattened the rotation profile in the rest of the radiative zone (Duez, Mathis, and Turck-Chièze, 2010). In any case, theories about the mechanisms that drive the solar rotation and its spatial and temporal variations remain to be tightly constrained by improved helioseismic inversion results. Better rotation profiles mean not only improved inversion methodologies but improved estimates of rotational-frequency splittings.

We present here results derived using an improved inversion methodology that i) adjusts the inversion grid (over both depth and latitude) based on the data set and its precision, and ii) solves the inversion problem iteratively. But first we review recent developments in global-mode characterization (Korzennik, 2008) that allowed us to infer with better confidence the internal-rotation rate and its time-varying patterns. We describe in detail the inversion methodology and show the resulting profiles.

Table 1 List of instruments and time span from which data sets were used in the work presented here.

Instrument	Time span
BiSON (ground-based)	01 Jan 1992 – 31 Dec 2002
GONG (ground-based)	07 May 1995 – 11 Feb 2011
GOLF (SOHO)	21 May 1996 – 07 Jun 2007
MDI (SOHO)	01 May 1996 – 12 Dec 2008
HMI (SDO)	30 Apr 2010 – 16 Sep 2011

2. The Data Sets

2.1. Introduction

We have used rotational-frequency splittings determined from fitting data acquired with five different instruments. Two are ground-based: the *Birmingham Solar Oscillations Network* (BiSON) and the *Global Oscillation Network Group* (GONG) and three are onboard spacecraft: the *Global Oscillations at Low Frequencies* (GOLF) and the *Michelson Doppler Imager* (MDI) onboard the *Solar and Heliospheric Observatory* (SOHO), and the *Helioseismic and Magnetic Imager* (HMI) onboard the *Solar Dynamics Observatory* (SDO). For all but the last instrument, the available data sets span well over a decade of observations. Table 1 summarizes what data, from which instrument, and for what time span are included in this study.

The data from these five instruments were fitted with various techniques, and in some cases the same data were fitted with more than one methodology. The GOLF and BiSON data were fitted, using "Sun-as-a-star" fitting techniques, as described by García *et al.* (2008) and Broomhall *et al.* (2009), respectively. The fitting, in both cases, is limited by these instruments' lack of spatial resolution to low-degree modes ($\ell \leq 3$).

The methodology for the mode-fitting pipeline used by the GONG project is described by Anderson, Duvall, and Jefferies (1990). It processes 108-day long overlapping time series, each 36 days apart, and individually fits each mode. It does it without including any spatial leakage matrix information and uses a symmetric profile for the mode power-spectral density. When resolved, spatial leaks are independently fitted, but when they are not resolved (in most cases), blended leaks are fitted as a single peak. Since there is no inclusion of any leakage information, the blending affects the result, skewing the mode frequency and the mode line width.

The mode-fitting pipelines used by the MDI team (both the standard and the "improved" pipelines) fit non-overlapping 72-day long epochs. That fitting methodology fits multiplets, using a polynomial expansion in m to model the frequency splitting, and includes the leakage-matrix information (as described by Schou, 1992). The improved pipeline (Larson and Schou, 2008) includes an improved spatial decomposition, where the effective instrument plate scale and our best model of the image distortion is included, as well as an improved leakage computation that incorporates the distortion of the eigenmodes by differential rotation (Woodard, 1989). The improved pipeline is set up to fit either a symmetric or an asymmetric mode power-spectral density profile.

2.2. Our Alternate Peak-Fitting Method

Korzennik (2005, 2008) has developed and implemented an alternative fitting methodology, which has processed GONG, MDI, and HMI data. The key elements of this method are

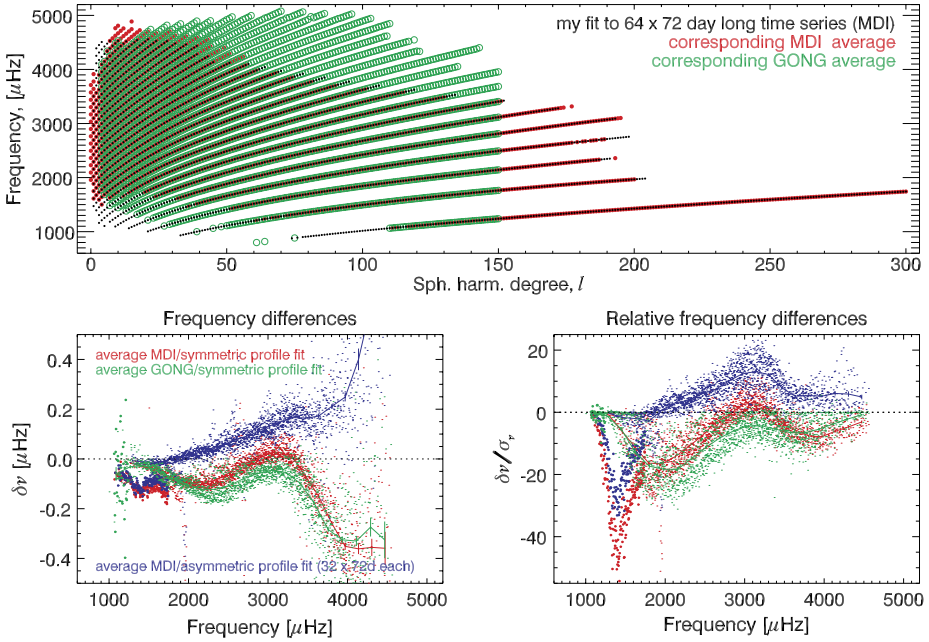


Figure 1 Comparison of fitted frequencies (singlets): top panel shows the coverage in the ℓ–ν plane, bottom panels shows frequency and relative frequency differences (*i.e.*, differences divided by the uncertainty), relative to the alternate fitting methodology. Black dots correspond to modes fitted using the alternate fitting methodology, applied to 64 × 72-day long time series, the red and green dots correspond to MDI and GONG pipeline fitting, respectively, while the blue dots correspond to MDI "improved" fitting, using an asymmetric profile. The large dots correspond to the f-modes, the curves are the p-mode corresponding binned values.

as follow: it fits individual multiplets, simultaneously for all the azimuthal orders while including the leakage information. The leakage matrix includes the effect of the distortion of the eigenmodes by differential rotation (Woodard, 1989). The spectral estimator is a sine multi-tapered one, whose number of tapers is adjusted to be *optimal*, a value derived from the mode line width. The mode power-spectral density profile is asymmetric, the procedure is iterative so as to include mode contamination (mode with a different radial order $[n]$ present in the fitting window), and it includes a rejection factor, where modes with too low an amplitude are not fitted.

The other major difference in the implementation of this method is that we choose to fit time series of varying lengths. The gain in signal-to-noise ratio when using longer time series allows us to derive more accurate mode parameters, while trading precision for temporal resolution. We used $64\times$, $32\times$, $16\times$, $8\times$, $4\times$, and 2×72-day long, overlapping, time series, as well as a 1×72-day long non-overlapping epochs (note that the longer segments all start on 01 May 1996, *i.e.* the start of science-quality observations for MDI). This extensive analysis of some 13 years of data was carried out on the Smithsonian Institution High Performance Cluster.

This method was used to fit GONG time series, using a leakage matrix specifically computed for that instrument, although the change in leakage resulting from the 2001 camera upgrade was not yet included (Schou, private communication, 2003). That same method was used to fit MDI data, for the exact same epochs, but using an MDI-specific leakage matrix. In fact, we fitted the data using a leakage matrix supplied by the MDI team, as well as our

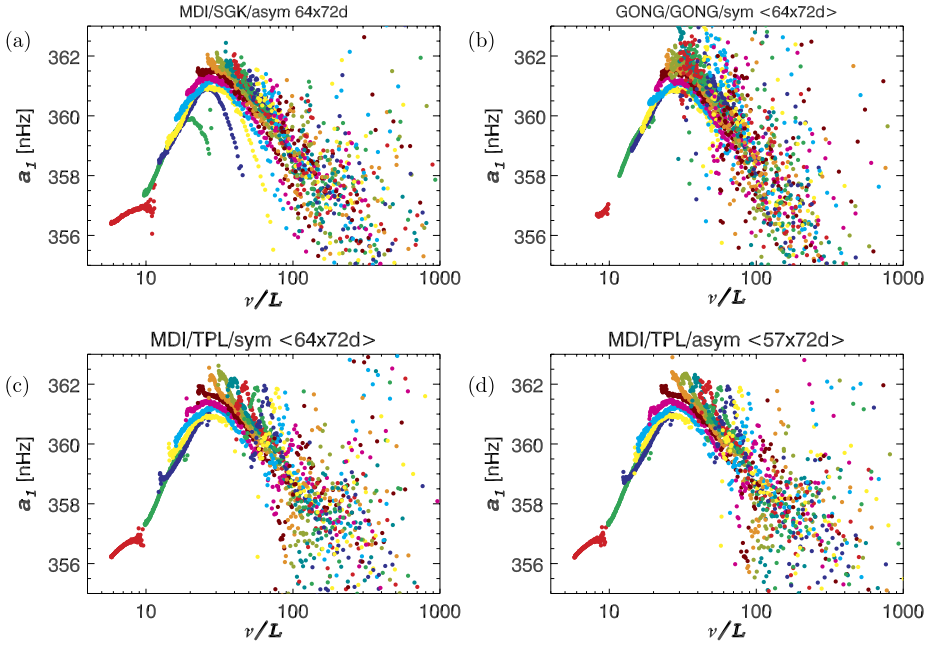


Figure 2 Comparison of the frequency splitting leading Clebsch–Gordan coefficient derived from four mode-fitting procedures: (a) results from using our alternative fitting methodology, (b) GONG pipeline, (c) MDI improved symmetric fit, and (d) MDI improved asymmetric fit. The resulting a_1 coefficients are plotted *versus* $\frac{\nu}{L}$, where $L^2 = \ell(\ell+1)$, while the symbol's color corresponds to the mode order, n.

own independent leakage-matrix computation. We used the "improved" MDI time series, where the spatial decomposition includes the effective instrument plate scale and our best model of the image distortion. We also fitted HMI *provisional* time series (as the HMI processing pipeline is yet to be finalized). The HMI instrumental image distortion and precise plate scale are included at the filtergram processing level, and the data were fitted using a *provisional* leakage matrix (*i.e.* the one derived for the full-disk MDI observations).

Figures 1 and 2 and Table 2 compare results from fitting GONG and MDI data with the respective project's analysis pipeline and the above described alternate fitting method. The table lists the mean and standard deviation of the differences in the a_1 Clebsch–Gordan coefficients (linear term) estimated by various fitting procedures. The resulting fits show systematic differences, which are not simply explained by the inclusion or not of an asymmetric profile, with even larger and systematic differences for the f-mode. The comparisons of the rotational-splitting coefficients show less of a scatter for the linear term, when using the alternate peak-fitting method, and differences at the few σ level.

But also important, if not more important, is the difference in mode attrition, when using the various fitting methods. Figure 3 illustrates that mode attrition, *i.e.* how often a mode is successfully fitted for each epoch analyzed. That figure shows clearly that the project pipeline methods produce large attrition, while the alternate peak-fitting method results display a more consistent fitting pattern. In order to be confident that we deduce significant changes of the solar rotation, when inverting rotational-frequency splittings for various epochs, we ought not to inject changes resulting from using different mode sets in the inversions. The estimated solutions of an inversion problem are some weighted spatial average of

Table 2 Comparison of resulting a_1 Clebsch–Gordan coefficients (linear term) derived from four mode-fitting procedures. The table lists the mean and standard deviation of the differences in a_1.

	δa_1 [nHz]	$\delta a_1/\sigma_{a_1}$
GONG (sym.) *vs.* alternate (asym.) 64 × 72-day long	-0.277 ± 0.984	-0.917 ± 1.279
MDI (sym.) *vs.* alternate (asym.) 64 × 72-day long	0.051 ± 0.635	0.534 ± 2.888
MDI (asym.) *vs.* alternate (asym.) 32 × 72-day long	0.096 ± 0.769	1.398 ± 2.384

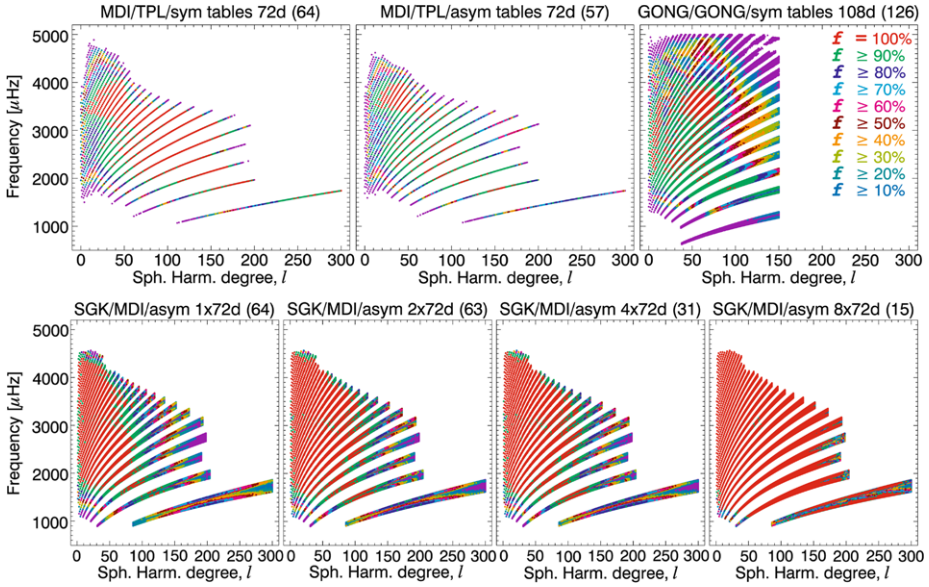


Figure 3 Mode attrition in the ℓ–ν plane. The color represents how often a mode is fitted, with red indicating all the time ($f = 100$ %), green 90 %, *etc*. The top panels correspond to MDI improved symmetric and asymmetric fit and GONG pipeline fit. The bottom panels correspond to our alternative fitting methodology, for 1×, 2×, 4×, 8 × 72-day long time series (left to right).

the "real" underlying solution. Those weights (also known as resolution kernels) depend on the extent of the input set, and thus change when the input sets change.

3. Inversion Methodology

The starting point of all helioseismic, linear rotational-inversion methodologies is the functional form of the perturbation in frequency $[\Delta\nu_{n\ell m}]$ induced by the rotation of the Sun, $\Omega(r,\theta)$:

$$\Delta\nu_{n\ell m} = \int_0^{R_\odot}\int_0^{\pi} K_{n\ell m}(r,\theta)\Omega(r,\theta)\,\mathrm{d}r\,\mathrm{d}\theta \pm \epsilon_{n\ell m}. \tag{1}$$

The perturbation in frequency $[\Delta\nu_{n\ell m}]$ with the observational error $[\epsilon_{n\ell m}]$, which corresponds to the rotational component of the frequency splittings, is given by the integral of the product of a sensitivity function, or kernel $[K_{n\ell m}(r,\theta)]$, with the rotation rate $[\Omega(r,\theta)]$ over

the radius [r] and the co-latitude [θ]. The kernels [$K_{n\ell m}(r,\theta)$] are known functions of the solar model.

Equation (1) defines a classical inverse problem for the Sun's rotation. The inversion of this set of M integral equations – one for each measured $\Delta\nu_{n\ell m}$ – allows us to infer the rotation-rate profile as a function of radius and latitude from a set of observed rotational-frequency splittings (hereafter referred to as splittings).

Our inversion method requires the discretization of the integral relation to be inverted. In our case, Equation (1) is transformed into a matrix relation

$$\mathbf{D} = \mathsf{A}\mathbf{x} + \boldsymbol{\epsilon}, \tag{2}$$

where $\mathbf{D}$ is the data vector, with elements $\Delta\nu_{n\ell m}$ and dimension M, $\mathbf{x}$ is the solution vector to be determined at N model grid points, A is the matrix with the kernels of dimension $M \times N$, and $\boldsymbol{\epsilon}$ is the vector containing the corresponding observational uncertainties. The number and location of the N model grid nodes are calculated according to the effective spatial resolution of the inverted data set. Such a procedure produces a non-equally spaced (*i.e.* unstructured) mesh distribution. A complete description and examples of the gridding methodology can be found in Eff-Darwich and Pérez-Hernández (1997) and Eff-Darwich, Korzennik, and García (2010).

The resulting unstructured grid is used to compute the matrix A in Equation (2). That equation is then solved with a modified version of the iterative method developed by Starostenko and Zavorotko (1996). This approach calculates x according to the following algorithm:

$$\mathbf{x}^{k+1} = \mathbf{x}^{k} - \mathsf{B}^{-1}\mathsf{A}^{\mathrm{T}}\mathsf{R}^{-1}\left(\mathsf{A}\mathbf{x}^{k} - \mathbf{D}\right), \tag{3}$$

where k is the iteration index. The diagonal matrices B and R are calculated from the summation of columns and rows of matrix A, respectively. For each iteration, values for the error propagation and data misfit, $\chi^2 = |\mathsf{A}\mathbf{x} - \mathbf{D}|^2$, are calculated.

4. Results

Helioseismology, as a tool to infer the properties of the solar interior, is based on the fact that different mode sets are sensitive to different layers of the Sun. Hence, by combining these mode sets, it is possible to derive the structure and dynamics of the solar interior. However, these sets are not homogeneous and the number and quality of the modes that are sensitive to the solar radiative interior is significantly lower than those sensitive to the convective zone and the surface layers (as shown in Figure 4). Therefore, the dispersion and the level of uncertainties of the modes that are sensitive to the core are the largest for the entire data set. Another problem arises as we look closely at the uncertainties (see Figure 4): the error level as a function of radius is not strictly monotonic. For a given inner turning radius, the scatter of the errors is rather large and is primarily the consequence of the reduced accuracy of estimates at high frequencies.

Figure 5 shows how consistent both the range in degree and in frequency are when the alternative fitting technique developed for this work is used on data from different instruments. In contrast, the mode sets obtained by the team pipelines, for both GONG and MDI, differ significantly, especially for their frequency spans. The consistency of our fitting technique is shown in Figure 6: this figure shows how both the uncertainties and the data dispersion are reduced when the length of the time series analyzed is increased. The improvement is

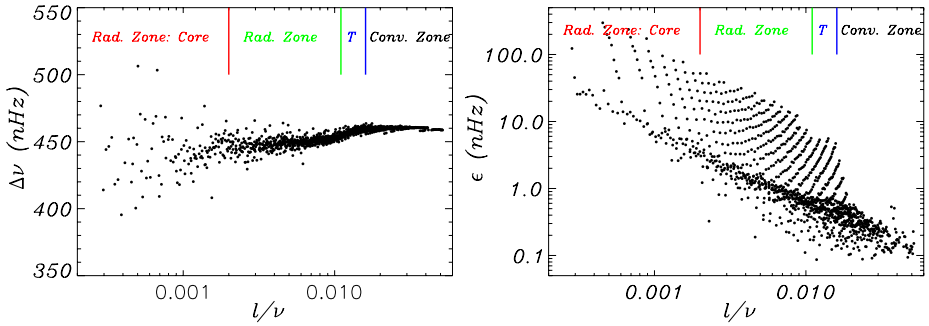


Figure 4 Left panel: observational sectoral frequency splittings (MDI 64 × 72-day long time series) as a function of the $\frac{\ell}{\nu}$ ratio, a proxy for the inner turning radius. For illustrative purposes, the approximate extent of the solar core, radiative zone, tachocline, and convective zone are represented. Right panel: as in the left panel, but for the observational uncertainties of sectoral frequency splittings.

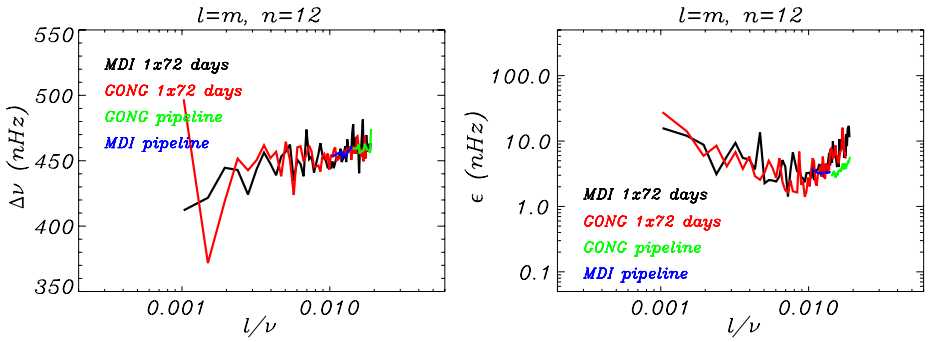


Figure 5 Left panel: observational sectoral frequency splittings as a function of the $\frac{\ell}{\nu}$ ratio, a proxy for the inner turning radius, for $n = 12$ modes obtained by fitting MDI 1 × 72, GONG 1 × 72-day long time series and the MDI and GONG team pipelines. Right panel: as in the left panel, but for the observational uncertainties.

less apparent for the modes that are more sensitive to the solar core and in the data sets corresponding to shorter time series. However, in the case of the 64 × 72-day long mode set, the uncertainties for the modes sensitive to the core are reduced by a factor of four relative to the mode sets obtained from shorter time series.

Hence, uncertainties of the data sensitive to the solar core rotation decrease when longer time series and better fitting technique are used. The level of uncertainties that we need to reach to counteract the low sensitivity of the modes to these regions is illustrated in Figure 7, with test profiles. Two sets are presented:

i) the radiative zone is rotating rigidly, at a rate of 432 nHz, and below $0.12R_\odot$ at rates of 2832, 835, and 132 nHz;
ii) the radiative zone is also rotating rigidly at a rate of 432 nHz, and where below $0.2R_\odot$ the rates are again set to 2832, 835, and 132 nHz.

Out of these six test profiles, only one is substantially and significantly different from the frequency-averaged $\ell = 1$ rotational splittings (*i.e.* averaged over frequencies in the 1.1 to

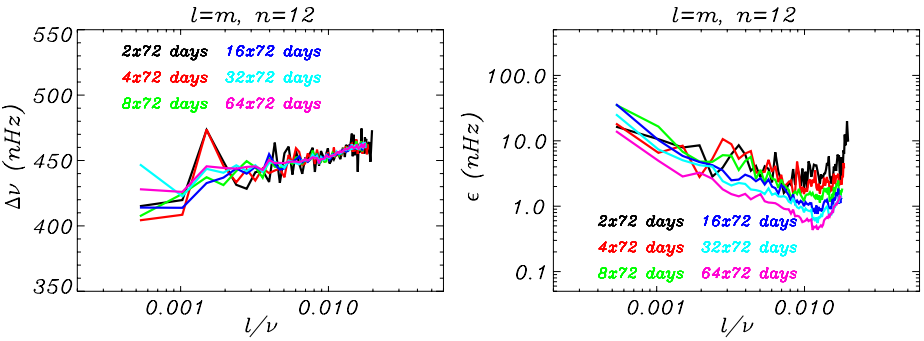


Figure 6 Left panel: observational sectoral-frequency splittings as a function of the $\frac{\ell}{\nu}$ ratio, a proxy for the inner turning radius, for $n = 12$ modes obtained by fitting MDI 2×, 4×, 8×, 16×, 32×, and 64 × 72-day long time series. Right panel: as in the left panel, but for the observational uncertainties.

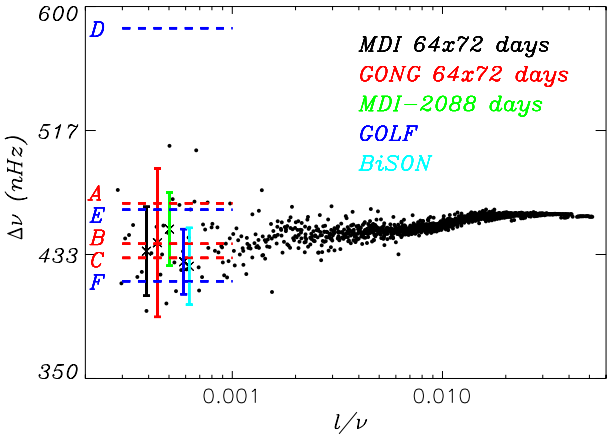


Figure 7 Observational sectoral frequency splittings as a function of the horizontal phase speed for $\ell = 1$ modes (MDI 64 × 72-day long time series). The vertical colored lines represent the frequency-averaged (1.1 to 3.3 mHz) $\ell = 1$ rotational splittings derived from our peak-fitting methodology (MDI and GONG 64 × 72 days), the MDI pipeline (using a 2088-day long time series), GOLF, and BiSON. The red dashed lines represent the average values of the theoretical $\ell = 1$ rotational splittings, if the radiative zone were rotating rigidly at a rate of 432 nHz, with rates below $0.12R_\odot$ of 2832, 835, and 132 nHz (A, B, and C, respectively). The blue dashed lines represent the average values of the theoretical $\ell = 1$ rotational splittings, with rates below $0.2R_\odot$ of 2832, 835, and 132 nHz (D, E and F, respectively).

3.3 mHz range) derived from our peak-fitting methodology (MDI and GONG 64 × 72 days), the MDI pipeline (using a 2088-day long time series), GOLF, and BiSON data sets.

The diagnostic potential of the new global-mode fitting technique when combined with the improved inversion methodology is illustrated in Figures 8, 9, and 10, where we present the time-averaged rotation profiles of the Sun from the surface down to $0.15R_\odot$ that were calculated using either MDI or GONG, and 2×, 4×, 8×, 16×, 32×, and 64 × 72-day long time series. Inversions using recent HMI data (2× and 4 × 72-day long) are also presented, although they are not yet comparable to either MDI or GONG results, since the amount of HMI observations is still significantly smaller.

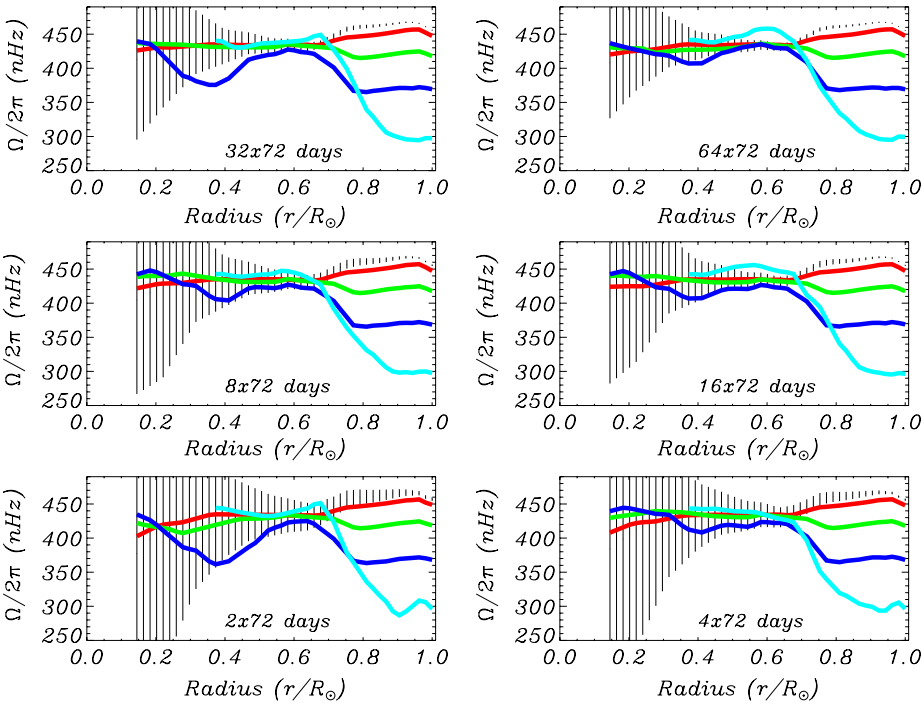


Figure 8 Time-averaged rotational profiles obtained from the inversions of rotational-frequency splittings resulting from fitting MDI 2×, 4×, 8×, 16×, 32×, and 64 × 72-day long time series. Black, red, green, dark-blue, and light-blue lines correspond to the rotational rate at different latitudes, namely 0, 20, 40, 60, and 80°, respectively. Vertical lines represent the error bars for the rotational rate at the Equator.

Both MDI and GONG inversions give similar results, with the largest discrepancies at high latitudes and below $0.40R_\odot$. The most significant difference between the inversions obtained by the same instrument is the reduction of the uncertainties, in particular random noise, when the length of the time series used to fit is increased. This reduction is particularly important in the inner radiative core.

All results are compatible with a radiative zone rotating rigidly at a rate of approximately 431 nHz; however, it is not possible to exclude a faster or slower rotator below $0.2R_\odot$ (*i.e.* up to 600 or down to 300 nHz). Although the radiative zone seems to rotate rigidly, there is a consistent and systematic dip in the rotation profile located at approximately $0.4R_\odot$ and 60° in latitude. This dip is seen in both MDI and GONG results, notwithstanding the actual length of the fitted time series.

This result is intriguing, particularly if we analyze the temporal evolution of the dip, for both MDI and GONG derived profiles, as shown in Figure 11. It was not possible to include the 1× and 2 × 72-day long results, since the quality of the inverted profiles at that depth and latitude is too low. Therefore, we used the 4 × 72-day long data, since its precision and temporal resolution allow us to carry out a temporal evolution analysis with adequate quality of the resulting profiles. Although the dip is certainly at the limit of the resolution of the data and the inversion method, there is a systematic temporal change of the dip. This variation is not found at other latitudes.

The consequences of using different peak-fitting techniques on the inversion results are illustrated in Figure 12. That figure shows the time-averaged rotational rates obtained using

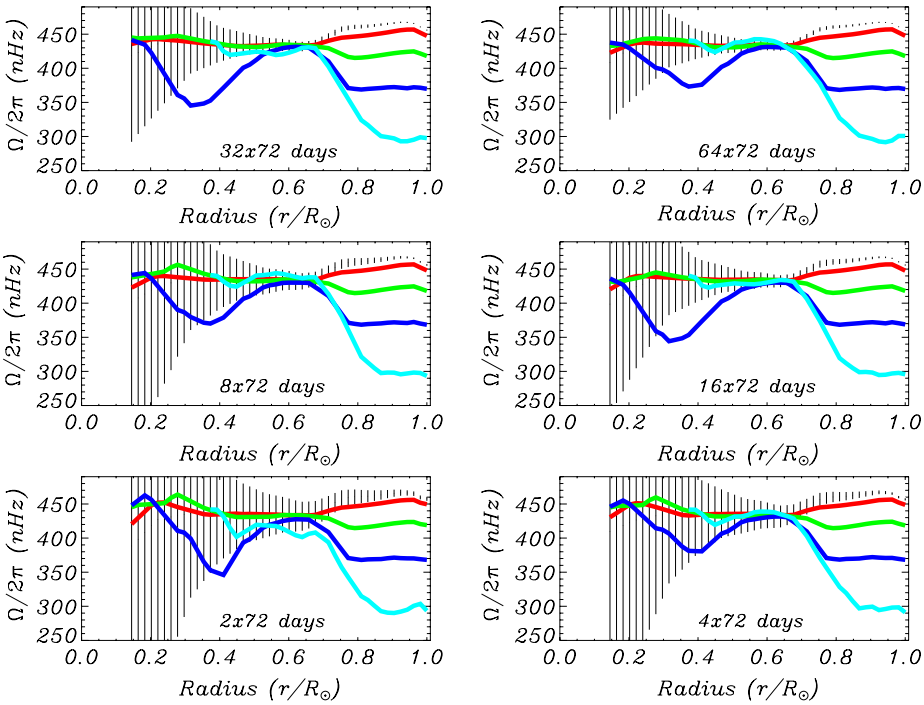


Figure 9 Time-averaged rotational profiles obtained from the inversions of rotational-frequency splittings resulting from fitting GONG 2×, 4×, 8×, 16×, 32×, and 64 × 72-day long time series. Black, red, green, dark-blue, and light-blue lines correspond to the rotational rate at different latitudes, namely 0, 20, 40, 60, and 80°, respectively. Vertical lines represent the error bars for the rotational rate at the Equator.

MDI and GONG 2 × 72-day long alternative fitting method and GONG and MDI respective project pipelines. The lengths of the fitted time series are comparable, however the spherical harmonic degree and frequency ranges of the fitted mode sets differ significantly. In particular, the mode sets obtained by the project pipelines result in rotational profiles that disagree significantly in the spatial extent of the optimal inversion grid and in the inverted rotation rates at high latitudes and in the radiative zone. The mode sets obtained through the alternate technique devised for this work are, in contrast, homogeneous, even though data from different instruments were fitted. Hence systematic differences introduced by different fitting techniques and different mode sets are greatly reduced.

5. Conclusions

We have fitted one solar cycle of MDI and GONG data and the latest HMI data using a new fitting methodology. This method fits individual multiplets, an asymmetric mode profile, incorporates all known instrumental distortion, uses our best estimate of the leakage matrix, and uses an optimal sine multi-tapered spectral estimator. It was applied to time series of varying lengths to study the effect of trading precision for temporal resolution in the inversion results. On the other hand, the improved inversion method that we used is one that estimates the optimal inversion model grid based on the extent of the mode set (over spherical harmonic degree and frequency) and the data uncertainties.

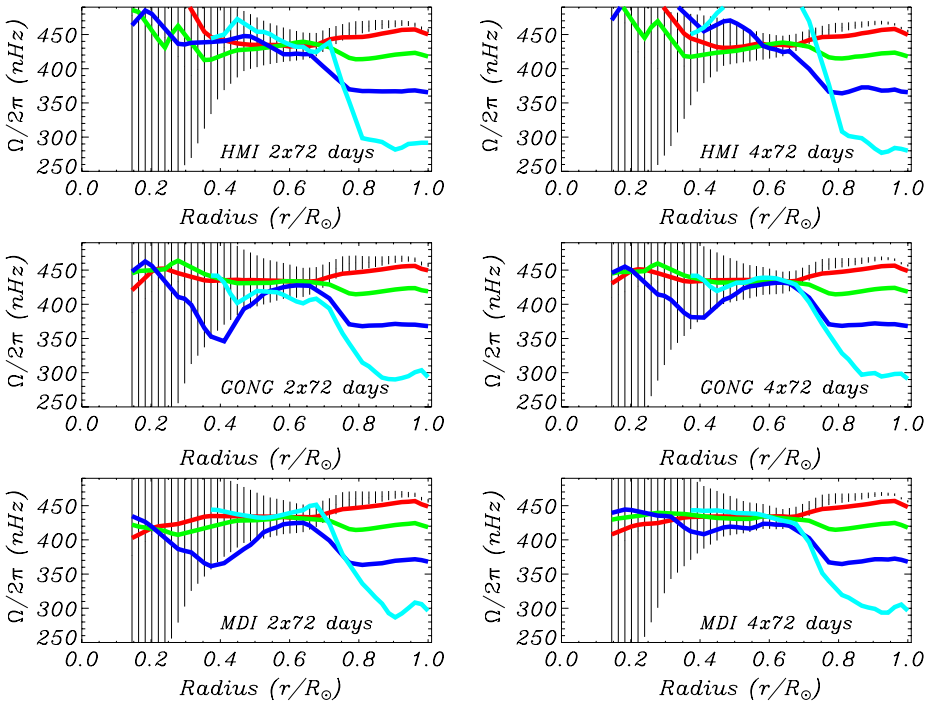


Figure 10 Time-averaged rotational profiles obtained from the inversions of MDI, GONG, and HMI 2 × 72 and 4 × 72-day long sets. Black, red, green, dark-blue, and light-blue lines correspond to the rotational rate at different latitudes, namely 0, 20, 40, 60, and 80°, respectively. Vertical lines represent the error bars for the rotational rate at the Equator.

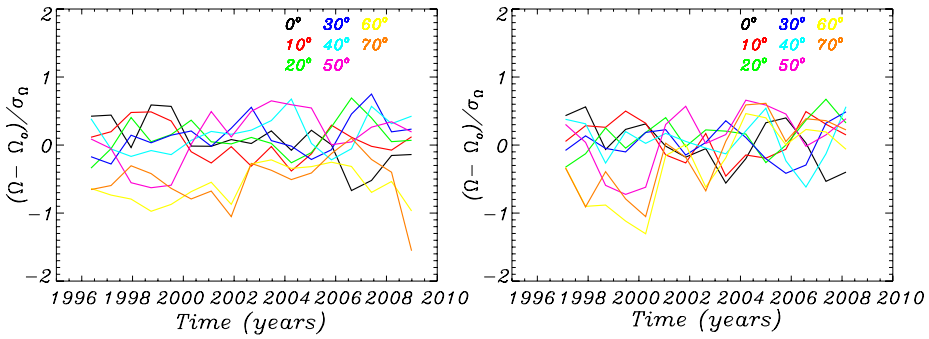


Figure 11 Left panel: temporal evolution of the relative residual rotation rate $[(\Omega - \Omega_o)/\sigma_\Omega]$, where $\Omega_o/2\pi = 432$ nHz, at a depth of $0.4R_\odot$ and for different latitudes obtained from the inversion of the different MDI 4 × 72-day long data sets. Right panel: as in the left panel, but for GONG data.

Our results are summarized in Figure 13, where we present the rotational profiles obtained from inverting frequency splitting derived from fitting time series spanning an entire solar cycle, Cycle 23, for both GONG and MDI observations. These profiles are our best inferences of the rotation in the radiative region, to date. Both results are compatible with a radiative zone rotating rigidly at a rate of approximately 431 nHz; however, it is not possible

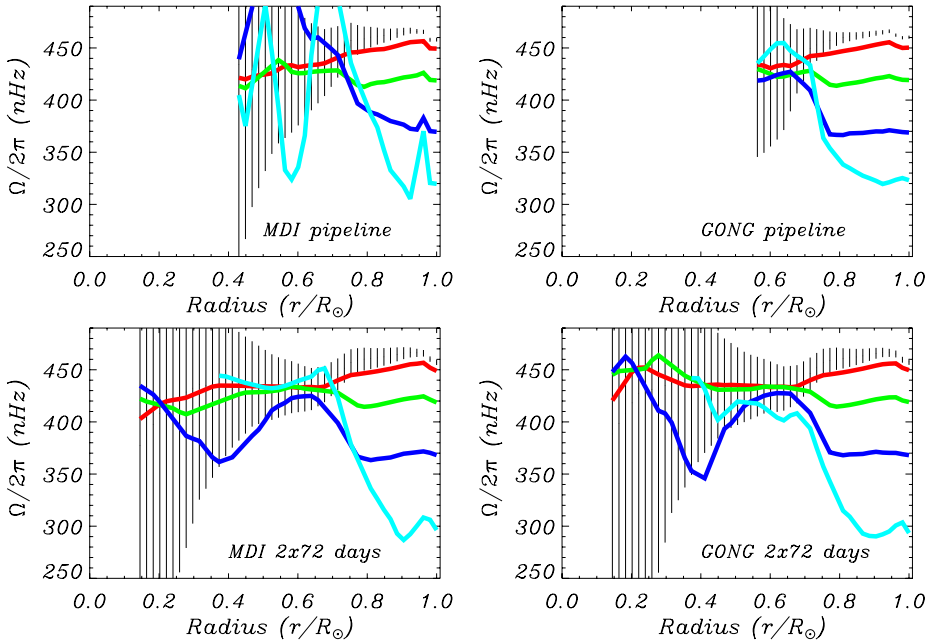


Figure 12 Time-averaged rotational profiles obtained from the inversions of MDI 2×72, GONG 2×72-day long, GONG pipeline, and MDI pipeline. Black, red, green, dark-blue, and light-blue lines correspond to the rotational rate at different latitudes, namely 0, 20, 40, 60, and 80°, respectively. Vertical lines represent the error bars for the rotational rate at the Equator.

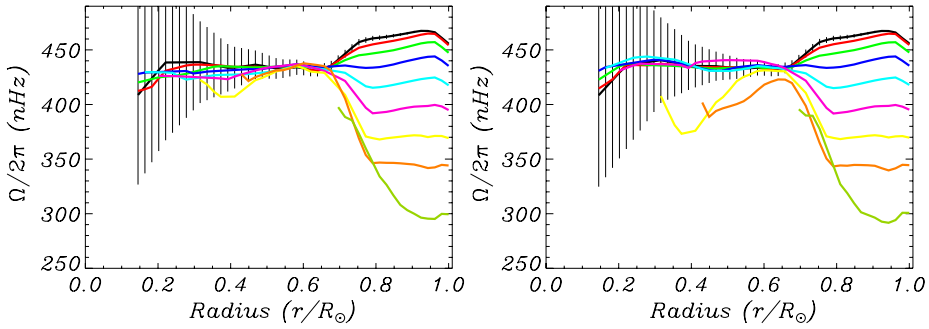


Figure 13 Rotational profiles obtained from the inversions of MDI 64×72 (left panel) and GONG 64×72-day long (right panel). The rotational rates at different latitudes, from the Equator to 80° in steps of 10°, are represented by colored lines. Vertical lines represent the error bars for the rotational rate at the Equator.

to disregard a faster or slower rotator below $0.2R_\odot$ (*i.e.* up to 600 or down to 300 nHz). Although the radiative zone seems to rotate rigidly, there is a consistent and systematic dip in the rotation profile located at around $0.4R_\odot$ and 60° of latitude. This dip appears to evolve with time, although this last result has to be confirmed when additional time series covering Cycle 24 become available.

References

Anderson, E.R., Duvall, T.L. Jr., Jefferies, S.M.: 1990, *Astrophys. J.* **364**, 699. ADS: http://adsabs.harvard.edu/abs/1990ApJ...364..699A.

Broomhall, A.M., Chaplin, W.J., Davies, G.R., Elsworth, Y., Fletcher, S.T., Hale, S.J., Miller, B., New, R.: 2009, *Mon. Not. Roy. Astron. Soc.* **396**, L100. ADS: http://adsabs.harvard.edu/abs/2009MNRAS.396L.100B, doi:10.1111/j.1745-3933.2009.00672.x.

Duez, V., Mathis, S., Turck-Chièze, S.: 2010, *Mon. Not. Roy. Astron. Soc.* **402**, 271. ADS: http://adsabs.harvard.edu/abs/2010MNRAS.402..271D. doi:10.1111/j.1365-2966.2009.15955.x.

Eff-Darwich, A., Pérez-Hernández, F.: 1997, *Astron. Astrophys. Suppl.* **125**, 1. ADS: http://adsabs.harvard.edu/abs/1997A&AS..125..391E, doi:10.1051/aas:1997229.

Eff-Darwich, A., Korzennik, S.G., García, R.A.: 2010, *Astron. Nachr.* **331** 890. ADS: http://adsabs.harvard.edu/abs/2010AN...331..890E, doi:10.1002/asna.201011420.

Eff-Darwich, A., Korzennik, S.G., Jiménez-Reyes, S.J.: 2002, *Astrophys. J.* **573**, 857. ADS: http://adsabs.harvard.edu/abs/2002ApJ...573..857E, doi:10.1086/340747.

Eff-Darwich, A., Korzennik, S.G., Jiménez-Reyes, S.J., García, R.A.: 2008, *Astrophys. J.* **679**, 1636. ADS: http://adsabs.harvard.edu/abs/2008ApJ...679.1636E, doi:10.1086/586724.

Forgács-Dajka, E., Petrovay, K.: 2001, *Solar Phys.* **203**, 195. ADS: http://adsabs.harvard.edu/abs/2001SoPh..203..195F, doi:10.1023/A:1013389631585.

Gabriel, A.H., Grec, G., Charra, J., Robillot, J.M., Roca Cortes, T., Turck-Chièze, S., Bocchia, R., Boumier, P., Cantin, M., Cespédes, E., *et al.*: 1995, *Solar Phys.* **162**, 61. ADS: http://adsabs.harvard.edu/abs/1995SoPh..162...61G, doi:10.1007/BF00733427.

García, R.A., Turck-Chièze, S., Jiménez-Reyes, S.J., Ballot, J., Pallé, P.L., Eff-Darwich, A., Mathur, S., Provost, J.: 2007, *Science* **316**, 1591. ADS: http://adsabs.harvard.edu/abs/2007Sci...316.1591G, doi:10.1126/science.1140598.

García, R.A., Mathur, S., Ballot, J., Eff-Darwich, A., Jiménez-Reyes, S.J., Korzennik, S.G.: 2008, *Solar Phys.* **251**, 119. ADS: http://adsabs.harvard.edu/abs/2008SoPh..251..119G, doi:10.1007/s11207-008-9144-5.

Gough, D.O., McIntyre, M.E.: 1998, *Nature* **394**, 755. ADS: http://adsabs.harvard.edu/abs/1998Natur.394..755G, doi:10.1038/29472.

Harvey, J.W., Hill, F., Hubbard, R.P., Kennedy, J.R., Leibacher, J.W., Pintar, J.A., *et al.*: 1996, *Science* **272**, 1284. ADS: http://adsabs.harvard.edu/abs/1996Sci...272.1284H.

Howe, R.: 2009, *Living Rev. Solar Phys.* **6**, 1. http://solarphysics.livingreviews.org/Articles/lrsp-2009-1/.

Korzennik, S.G.: 2005, *Astrophys. J.* **626**, 585. ADS: http://adsabs.harvard.edu/abs/2005ApJ...626..585K, doi:10.1086/429748.

Korzennik, S.G.: 2008, *J. Phys.* **CS-118**, 012082. ADS: http://adsabs.harvard.edu/abs/2008JPhCS.118a2082K, doi:10.1088/1742-6596/118/1/012082.

Larson, T.P., Schou, J.: 2008, *J. Phys.* **CS-118**, 012083. ADS: http://adsabs.harvard.edu/abs/2008JPhCS.118a2083L, doi:10.1088/1742-6596/118/1/012083.

Pijpers, F.P., Thompson, M.J.: 1994, *Astron. Astrophys.* **281**, 231. ADS: http://adsabs.harvard.edu/abs/1994A&A...281..231P.

Salabert, D., Chaplin, W.J., Elsworth, Y., New, R., Verner, G.A.: 2007, *Astron. Astrophys.* **463**, 1181. ADS: http://adsabs.harvard.edu/abs/2007A&A...463.1181S, doi:10.1051/0004-6361:20066419.

Scherrer, P.H., Bogart, R.S., Bush, R.I., Hoeksema, J.T., Kosovichev, A.G., Schou, J., *et al.*: 1995, *Solar Phys.* **162**, 129. ADS: http://adsabs.harvard.edu/abs/1995SoPh..162..129S, doi:10.1007/BF00733429.

Scherrer, P.H., Schou, J., Bush, R.I., Kosovichev, A.G., Bogart, R.S., Hoeksema, J.T., Liu, Y., Duvall, T.L., Zhao, J., Title, A.M., Schrijver, C.J., Tarbell, T.D., Tomczyk, S.: 2012, *Solar Phys.* **365**, 207. ADS: http://adsabs.harvard.edu/abs/2012SoPh..275..207S, doi:10.1007/s11207-011-9834-2.

Schou, J.: 1992, *On the Analysis of Helioseismic Data*, Ph.D. Thesis, Aarhus University. ADS: http://adsabs.harvard.edu/abs/1992PhDT.....380S.

Starostenko, V.I., Zavorotko, A.N.: 1996, *Comput. Geosci.* **22**, 3. ADS: http://adsabs.harvard.edu/abs/1996CG.....22....3S, doi:10.1016/0098-3004(95)00052-6.

Thompson, M.J., Christensen-Dalsgaard, J., Miesch, M.S., Toomre, J.: 2003, *Ann. Rev. Astron. Astrophys.* **41**, 599. ADS: http://adsabs.harvard.edu/abs/2003ARA&A..41..599T, doi:10.1146/annurev.astro.41.011802.094848.

Turck-Chièze, S., Palacios, A., Marques, J.P., Nghiem, P.A.P.: 2010, *Astrophys. J.* **715**, 1539. ADS: http://adsabs.harvard.edu/abs/2010ApJ...715.1539T, doi:10.1146/annurev.astro.41.011802.094848.

Woodard, M.F.: 1989, *Astrophys. J.* **347**, 1176. ADS: http://adsabs.harvard.edu/abs/1989ApJ...347.1176W.

Zahn, J.P., Brun, A.S., Mathis, S.: 2007, *Astron. Astrophys.* **474**, 145. ADS: http://adsabs.harvard.edu/abs/2007A&A...474..145Z, doi:10.1051/0004-6361:20077653.

DOI 10.1007/978-1-4899-8005-2_4
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0171-x

Latest Results Found with Ring-Diagram Analysis

C.S. Baldner · S. Basu · R.S. Bogart · O. Burtseva · I. González Hernández · D. Haber · F. Hill · R. Howe · K. Jain · R.W. Komm · M.C. Rabello-Soares · S. Tripathy

Received: 1 May 2012 / Accepted: 11 October 2012 / Published online: 13 November 2012

Abstract Ring-diagram analysis is a helioseismic tool useful for studying the near-surface layers of the Sun. It has been employed to study near-surface shear, meridional circulation, flows around sunspots, and thermal structure beneath active regions. We review recent results obtained using ring-diagram analysis, state some of the more important outstanding difficulties in the technique, and point out several extensions to the technique that are just now beginning to bear fruit.

Keywords Helioseismology · Observations

1. Introduction

Ring-diagram analysis (Hill, 1988) is a well-established tool in helioseismology for the study of the near-surface layers of the Sun. It falls into the general category of local helio-

Invited review.

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

C.S. Baldner (✉) · R.S. Bogart
Hansen Experimental Physics Laboratory, Stanford University, Stanford, CA 94305-4085, USA
e-mail: baldner@stanford.edu

S. Basu
Department of Astronomy, Yale University, P.O. Box 208101, New Haven, CT 06520-8101, USA

O. Burtseva · I. González Hernández · F. Hill · K. Jain · R.W. Komm · S. Tripathy
National Solar Observatory, 950 N. Cherry Ave., P.O. Box 26732, Tucson, AZ 85726-6732, USA

D. Haber
JILA, University of Colorado, Boulder, CO 80309-0440, USA

R. Howe
School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

M.C. Rabello-Soares
Department of Physics, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

seismology, which is a class of techniques that measure some local region of the Sun, rather than measuring the whole star at once. These tools include, alongside ring-diagram analysis, time–distance analysis (Duvall *et al.*, 1993), acoustic holography (Lindsey and Braun, 1990), and Fourier–Hankel analysis (Braun, Duvall, and Labonte, 1987). For a general review of local-helioseismic techniques, the reader should consult Gizon and Birch (2005). Here, we discuss recent results in ring-diagram analysis.

Ring-diagram analysis is similar to global-mode analysis in that the measurements are interpreted as resonant modes of the solar interior. They are measured over small patches of the solar surface, so the geometry is approximately plane-parallel, and the waves are treated as plane waves. The modes measured are typically $\ell \gtrsim 150$, which mostly have lower turning radii above $r = 0.95 R_\odot$. The ring diagrams themselves are power spectra made from Fourier transforms in time and space of solar data, and take their name from the characteristic shape that ridges of power take in k-space.

In this review, we focus on recent results in ring-diagram analysis. For a more in-depth discussion of previous results, the reader may consult reviews by Antia and Basu (2007) for a more detailed discussion of the techniques, particularly the inversion techniques, and González Hernández (2008) for a more detailed discussion of the flow results from ring diagrams. In Section 2, we outline the technique with an emphasis on the standard ring-diagram products that are used most commonly in current research. Section 3 discusses recent results, covering ring-diagram mode parameters, flows and dynamics around active regions, and finally determinations of solar structure. Section 4 discusses current problems in ring diagrams as well as current directions of research using ring diagrams and extensions of the technique.

2. Ring Diagrams

In most current research undertaken with ring diagrams, the techniques for construction and analysis of the ring diagrams are very similar, and in this section we cover the salient features. Ring diagrams can, in principle, be constructed using any spatially and temporally resolved set of solar observations. The vast majority of ring-diagram results, however, have been obtained using spatially resolved line-of-sight velocity data, and this remains true today.

Three data sets dominate ring-diagram research today: the *Global Oscillation Network Group* (GONG: Harvey *et al.*, 1996), the *Michelson Doppler Imager* (MDI: Scherrer *et al.*, 1995) on the *Solar and Heliospheric Observatory* (SOHO), and most recently the *Helioseismic and Magnetic Imager* (HMI: Scherrer *et al.*, 2012) on the *Solar Dynamics Observatory* (SDO). All three instruments provide high temporal-cadence, resolved images of the Sun in line-of-sight velocity, commonly referred to as "Dopplergrams," as well as intensity, line-of-sight magnetic-field measurements, and other products at various cadences. GONG data are available more or less continuously from 1996 to the present, but sufficient spatial resolution for ring diagrams was only available after the "GONG+" upgrade in 2001. The standard MDI helioseismic data product also does not have sufficient spatial resolution for ring diagrams. Due to bandwidth constraints, high-resolution data ("full disk" data) are only available with sufficient coverage during "dynamics campaigns", which were run for two or three months every year during the MDI mission. Higher-resolution data from the HMI instrument have been available from April 2010 to the present with virtually continuous coverage.

Ring diagrams are constructed from small patches, typically 15°, of solar data. They may be taken from a constant disk position or, more commonly, tracked with solar rotation. The

data are mapped to a plane using an azimuthal–equidistant projection (or a Postel's projection), which preserves distances along lines that pass through the center of the region, and are then Fourier transformed in space and in time. The resulting power spectrum can be thought of as a two-dimensional k–ω diagram familiar from global helioseismology wrapped around the $k = 0$ axis. At constant temporal frequency, the ridges of power resemble rings, which gives the technique its name.

The size of the tracked region, the length of the track, the position of the region on the solar disk, and the rate (if any) at which the region is tracked can all be chosen depending on the scientific questions involved. Mention should be made here, however, of the synoptic ring-diagram tracking schemes, which are widely used in the literature and available to researchers. "Dense-pack" sets provide coverage over the solar disk using patches with a 15° aperture, tracked at the Snodgrass rate for 1664 minutes (Haber *et al.*, 2000a). The HMI data reduction and analysis pipeline provides a similar set of products with three different aperture sizes: 5°, 15°, and 30°. These areas are tracked at the Carrington rate (Bogart *et al.*, 2011a, 2011b). The Snodgrass rate is a function of latitude while the Carrington rate is a fixed rotation rate. Thus, sets of rings tracked at the Carrington rate maintain their positions relative to each other, but move relative to solar features at different rates depending on their latitude.

Quantitative studies of ring diagrams are undertaken by fitting model spectral profiles to the data. The fits are obtained using a maximum-likelihood method. Historically a number of different models have been used; today two are in common use: one a symmetric profile, the other asymmetric.

The symmetric profile at a fixed frequency common use (Haber *et al.*, 2002) is a Lorentzian with a decaying background term and a fit to the displacement of the ring due to advection from flows:

$$P(\nu, k_x, k_y) = \frac{A}{(\nu - \nu_0 + k_x u_x + k_y u_y)^2 + \Gamma^2} + \frac{b_0}{k^3}. \tag{1}$$

The parameters ν_0, A, and Γ are the peak frequency, amplitude, and width, respectively, and b_0 is the background term. The zonal- and meridional-flow components are u_x and u_y.

The asymmetric profile used in ring-diagram analysis takes its form from Nigam and Kosovichev (1998) and was first used in this form by Basu, Antia, and Tripathy (1999):

$$P(\nu, k_x, k_y) = \frac{\exp[A_0 + (k - k_0)A_1 + A_2(\frac{k_x}{k})^2 + A_3\frac{k_x k_y}{k^2}]S_x}{x^2 + 1} + \frac{\mathrm{e}^{B_1}}{k^3} + \frac{\mathrm{e}^{B_2}}{k^4}, \tag{2}$$

where

$$x = \frac{\nu - ck^p - u_x k_x - u_y k_y}{w_0 + w_1(k - k_0)} \tag{3}$$

and

$$S_x = S^2 + (1 + Sx)^2. \tag{4}$$

This model has 13 parameters; as in Equation (1), the zonal and meridional velocities are u_x and u_y, respectively. The amplitude is given by the parameters A_0 through A_3, the latter three describing the azimuthal variation in power. The width is given by the parameters w_0 and w_1, the background terms are B_1 and B_2, and the dispersion relation is approximated locally as a power law with parameters c and p. The parameter S gives the asymmetry of the profile. Rather than fit a peak frequency, as the Haber *et al.* (2002) model does, this model fits for central k_0. Fits using both models are provided by the HMI pipeline – they are known as rdfitf and rdfitc, respectively. The former is much faster and so is done for all rings

in the synoptic pipeline, the latter is run for all 5° rings and on the Equator and the Central Meridian for 15° and 30° rings.

3. Recent Results

3.1. Ring Diagram Mode Parameters

It has been well established that in the presence of strong surface magnetic fields, such as active regions, helioseismic mode parameters are affected, and ring diagrams show these effects prominently (Haber *et al.*, 2000b; Rajaguru, Basu, and Antia, 2001; Howe *et al.*, 2004). The frequencies, mode widths, and amplitudes (either from Equation (1) or (2)) can be examined directly for clues about the effects of solar activity on acoustic-wave propagation.

It is well known that amplitudes decrease while mode widths increase in the presence of magnetic fields (Rajaguru, Basu, and Antia, 2001). Rabello-Soares, Bogart, and Basu (2008) have reported that the relation between the change in width and mode amplitude was very nearly linear. The correlation between local surface activity and changes in mode parameters is high, but several authors have examined frequency shifts over the course of Solar Cycle 23 for secular trends or changes in the dependence on activity. Tripathy, Jain, and Hill (2011) found that the correlation coefficients between the Magnetic Activity Index (MAI: a measure of the strong unsigned flux in the region sampled by the ring diagram, defined by Basu, Antia, and Bogart, 2004) and mode frequency decreased toward the end of Solar Cycle 23. Burtseva *et al.* (2009) found that these changes were not simply the effect of active regions on the total sample, and showed that the quiet-Sun mode parameters also shifted over the course of the solar cycle. Burtseva *et al.* (2011) found, when looking at low-activity regions using GONG data, that there were long-term variations in amplitude and width that did not appear to be related to activity. Baldner, Bogart, and Basu (2011) have found that the frequency changes with MAI are slightly different depending on the sunspot type. They suggested that if the changes in mode parameters were due to acoustic wave to magnetic wave conversion, this could be explained by differences in field geometry.

3.2. Meridional Circulation

Ring-diagram analysis was originally employed to study surface and near-surface flows, and measurements of the meridional flows were among the earliest ring-diagram results (Hill, 1988, 1989). Meridional circulation plays an important role in flux-transport dynamo models (see reviews by Miesch, 2005; Charbonneau, 2010). Data for the rise and fall of Solar Cycle 23 are now available, as well as for the unusually extended minimum preceding the onset of Cycle 24.

That meridional circulation changes with solar activity has been noted by many authors: from surface measurements (Komm, Howard, and Harvey, 1993; Snodgrass and Dailey, 1996; Hathaway, 1996; Ulrich and Boyden, 2005) and from helioseismic measurements (Giles, 2000; Chou and Dai, 2001). Ring diagrams were used by Basu and Antia (2000, 2003), who found that meridional-flow speed is anti-correlated with solar activity. González Hernández *et al.* (2008) studied the effect of active regions on meridional circulation, following suggestions that the change in flow over the solar cycle might be due to the inflows observed around active regions (*e.g.* Gizon, 2003). Using GONG data for the declining phase of Solar Cycle 23, they compared flow results averaged from all the available data to flow results measured only in quiet Sun. While the measured flows did change,

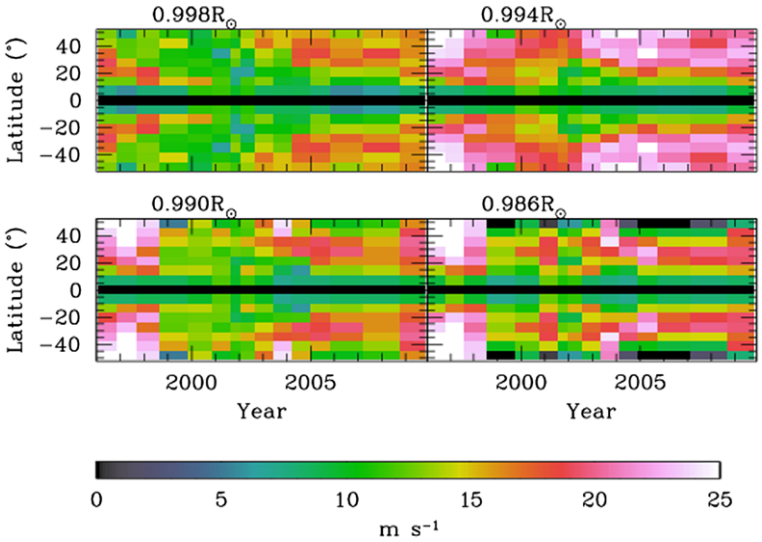


Figure 1 Antisymmetric component of meridional flow at four different depths over the course of Solar Cycle 23. Positive values are poleward flows. The inversions are performed on ring diagrams from MDI data – the temporal resolution is determined by the spacing and duration of MDI dynamics campaigns. Figure from Basu and Antia (2010), reproduced by permission of the AAS.

they found that changes in meridional circulation with activity remained and concluded that active-region inflows were not sufficient to explain the changes in shallow meridional flows.

The nature of the changes in meridional circulation were further looked at in two different studies. González Hernández *et al.* (2010) used ring diagrams from GONG data for the declining phase of Solar Cycle 23 through the onset of Cycle 24, while Basu and Antia (2010) used the sporadically available MDI data over the full solar cycle. Both works confirmed that shallow meridional flows have a component that is anti-correlated with surface activity. González Hernández *et al.* (2010) presented evidence of the appearance of enhanced poleward flows at mid-latitudes appearing prior to the emergence of activity with Solar Cycle 24. Basu and Antia (2010), with an entire solar cycle's worth of data, found that the meridional-flow patterns mimicked the well-known "torsional oscillation" observed in zonal flows (Figure 1).

3.3. Near Surface Dynamics and Solar Activity

While the global solar flows, both zonal and meridional, remain an area of active research as discussed above, one of the advantages of local-helioseismic techniques over global-mode analysis is the ability to spatially resolve flows and structures in longitude as well as latitude and depth. This is of particular interest in the context of solar active regions.

The flows associated with active regions have been studied in some detail using ring diagrams in a series of articles focusing on the signatures of active-region emergence and decay. Komm *et al.* (2004) discussed flow divergence and vorticity measured from ring diagrams, as well as a measurement of vertical velocity from the flow divergence and mass conservation. Komm *et al.* (2008) applied these measurements to a small sample of emerging active regions and found that up-flows were present in emerging-flux regions, while established active regions showed down-flows. Komm, Howe, and Hill (2009) extended this to a much

larger sample and confirmed this result, and Komm, Howe, and Hill (2011) studied a still larger sample. In these works, it was found that up-flows generally accompany the emergence of flux, with the strongest up-flows measured between 10 Mm and 16 Mm below the surface. An increase in down-flows was found to accompany decaying flux. Further, Komm, Howe, and Hill (2011) found that emerging flux was accompanied by an increase in subsurface zonal-flow velocity (Figure 2). Komm, Howe, and Hill (2012) found that emerging flux was associated with an increase in vorticity, while decaying flux was associated with an equivalent decrease.

Hindman, Haber, and Toomre (2009) used high-resolution (2°) ring diagrams, and studied near-surface flows around sunspots inferred from the f-mode. They found that active regions exhibited cyclonic convergent flows at the periphery, with anti-cyclonic divergent flows in the cores. This would seem to be at odds with the measurements of Komm *et al.* (2008), Komm, Howe, and Hill (2009, 2011) who generally found divergence in their much coarser-resolution studies. The difference in resolution between the two studies is too great to conclude an actual discrepancy, however. A more powerful technique, combining the information from ring diagrams of different sizes in a single inversion, has recently been developed by Featherstone, Hindman, and Thompson (2011). They find flows around active regions that are generally consistent with previous high -resolution ring-diagram studies.

The connection between subsurface dynamics and atmospheric activity is another area that has received some attention using ring-diagram analysis. Mason *et al.* (2006) found that regions that produced large flares were associated with high subsurface vorticity. Subsequent work (Komm, Howe, and Hill, 2009; Ferguson *et al.*, 2009; Komm *et al.*, 2011) has upheld this claim. Komm *et al.* (2011), using discriminant analysis to attempt to distinguish between flaring and non-flaring regions, found that the "structure vorticity" (the horizontal gradient of the vorticity) could be used as a discriminant. Studying one active region with an X17.2 flare (AR 10486), Maurya, Ambastha, and Tripathy (2009) found that mode amplitudes increased significantly after the flare. In addition, they found that the vorticity increased toward flare onset, and decreased afterward. Reinard *et al.* (2010) developed a parameter that they called the Normalized Helicity Gradient Variance (NHGV), derived from ring-diagram velocity measurements, which they found could statistically separate pre-flaring regions from non-pre-flaring regions.

3.4. Thermal Structure of the Near-Surface Layers

Basu, Antia, and Bogart (2004) first used ring diagrams to directly probe the thermal structure of active regions. They found, for a small number of active regions, that the sound speed and adiabatic index [Γ_1] do change in the layers below active regions. Both sound speed and adiabatic index are depressed in a shallow layer between approximately $r = 0.99 R_\odot$ and $r = 0.995 R_\odot$ (depths of 3 Mm to 7 Mm), and enhanced in the region between $r = 0.97 R_\odot$ and $r = 0.985 R_\odot$ (depths of 11 Mm to 21 Mm). They reported that the magnitude of these perturbations, both negative and positive, increased with increasing MAI.

Basu, Antia, and Bogart (2007) studied the thermal asphericity of the outer layers of the Sun using untracked ring diagrams. They found substantial changes, varying with latitude, in sound speed and adiabatic index over the course of the solar cycle. Recently, high-degree global-mode analysis has confirmed this result (Rabello-Soares, 2012).

The study of active regions using ring diagrams has been extended by a few authors. Bogart *et al.* (2008) applied the same analysis to more active regions, and generally confirmed the results of Basu, Antia, and Bogart (2004). They found a reasonably tight correlation between MAI and the magnitude of the changes in sound speed and Γ_1. Most recently, Baldner,

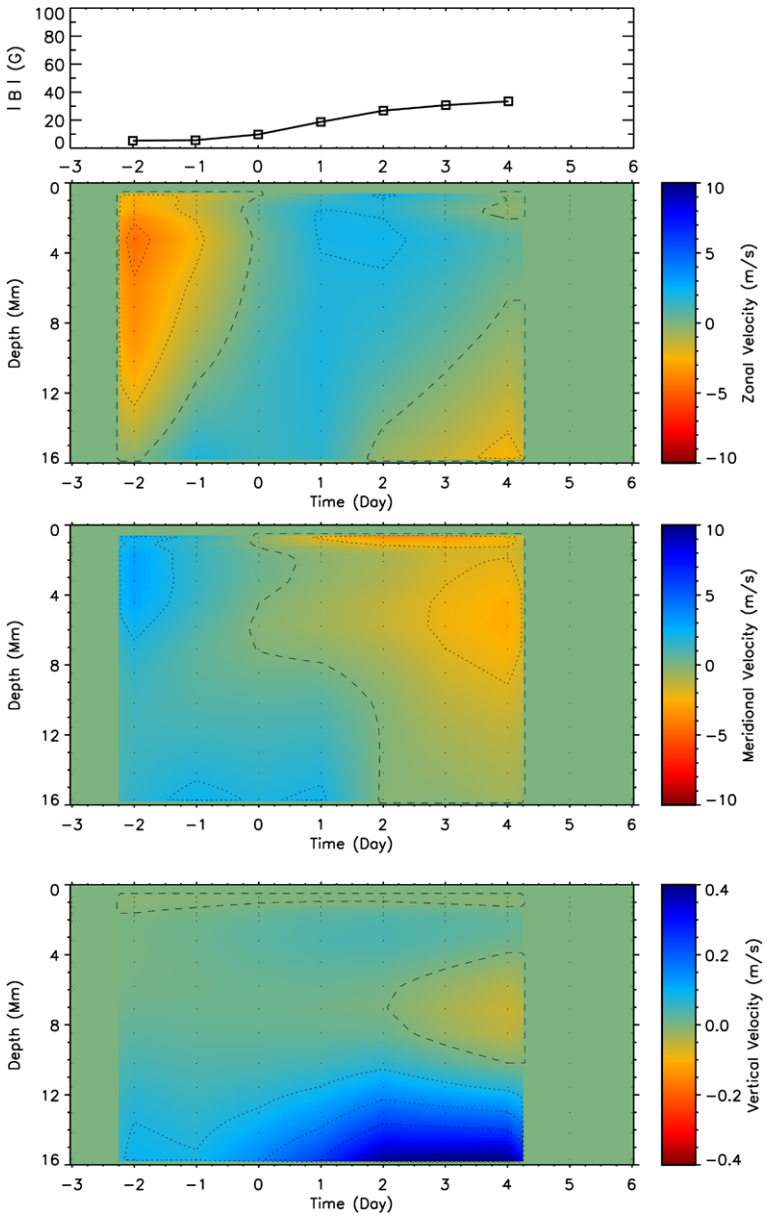


Figure 2 Averaged zonal-velocity residuals and vertical velocities for a number of emerging active regions. A sample of emerging active regions are considered. The average unsigned magnetic flux is shown in the top panel as a function of days before and after emergence. The second panel from the top shows the average difference between the zonal velocity of an emerging active region and the quiet Sun, as a function of time and depth. The third panel from the top shows the corresponding meridional velocity. The bottom panel shows vertical velocity over the same range of times and depths. Figure from Komm, Howe, and Hill (2011) by kind permission Springer Science+Business Media.

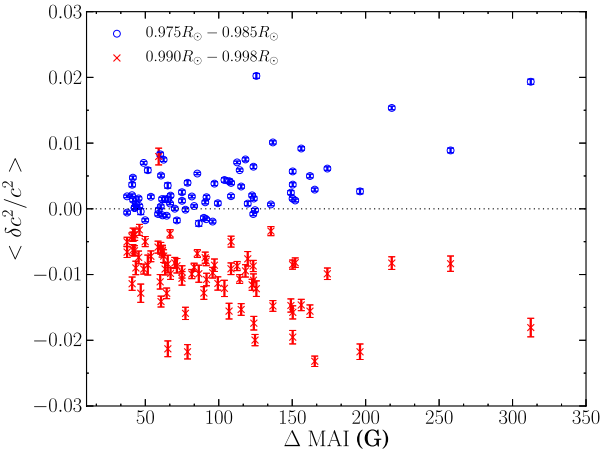


Figure 3 Averages of inversions for c^2 over two depth ranges, plotted as a function of Δ MAI. Blue circles are averages of inverted sound speed between $0.975R_\odot$ and $0.985R_\odot$; red crosses are averages of inverted sound speed between $0.99R_\odot$ and $0.998R_\odot$. Data from Baldner, Bogart, and Basu (2012).

Bogart, and Basu (2012) have looked at a much larger sample of active regions than either of the previous studies. They found the same two-layer structure in sound speed and Γ_1. They found that the magnitudes of the measured changes did tend to increase with increasing MAI (Figure 3), but the correlation was not very good.

The interpretation of these results is difficult. Although the inversion results are reported as changes in *sound speed* (Basu, Antia, and Bogart, 2004; Bogart *et al.*, 2008; Baldner, Bogart, and Basu, 2012), magnetic fields affect wave propagation in ways not accounted for in purely non-magnetic kernels. Lin, Basu, and Li (2009) investigated the effects of this distinction by assuming that the main magnetic contribution to the wave propagation measured with rings was through the magnetic pressure term [$P_{\rm mag} = B^2/8\pi$]. Using a 1D solar model modified to include certain magnetic effects, they found a relation between the measured perturbations in sound speed and adiabatic index, and the associated magnetic field. It must be noted, however, that a number of magnetic effects were neglected, and further work is needed to determine if these assumptions are valid in the regime probed by ring diagrams.

The uncertainty in interpretation inherent in helioseismic measurements of thermal structure in the presence of strong magnetic fields was demonstrated rather dramatically in a recent study of active region AR 9787 (Gizon *et al.*, 2009; Moradi *et al.*, 2010). In that work, a wide variety of analysis techniques were focused on a single active region. Ring diagram analysis was used to measure the perturbation in "sound speed", and time–distance analysis was used to measure the "wave speed" perturbation. The result was an almost perfect disagreement – negative time–distance perturbation where ring-diagram analysis showed positive and *vice versa*. The meaning of these results is not clear. Kosovichev *et al.* (2011) have pointed out that there was a large plage region near the sunspot in AR 9787, and suggested that, given the very large difference in spatial resolution between the two techniques, the ring analysis and time–distance analysis may have been sampling different things. The uniformity in the inversion results from Baldner, Bogart, and Basu (2012) implies that plage regions do not appreciably differ from sunspots examined with ring-diagram analysis. In general, the application of helioseismic techniques in the presence of strong magnetic fields is fraught with difficulty. Magnetic fields affect waves both by changing the thermal structure of the Sun through which the waves propagate, and by changing the physics of the modes themselves. The discrepancy from Gizon *et al.* (2009) and Moradi *et al.* (2010) remains unresolved.

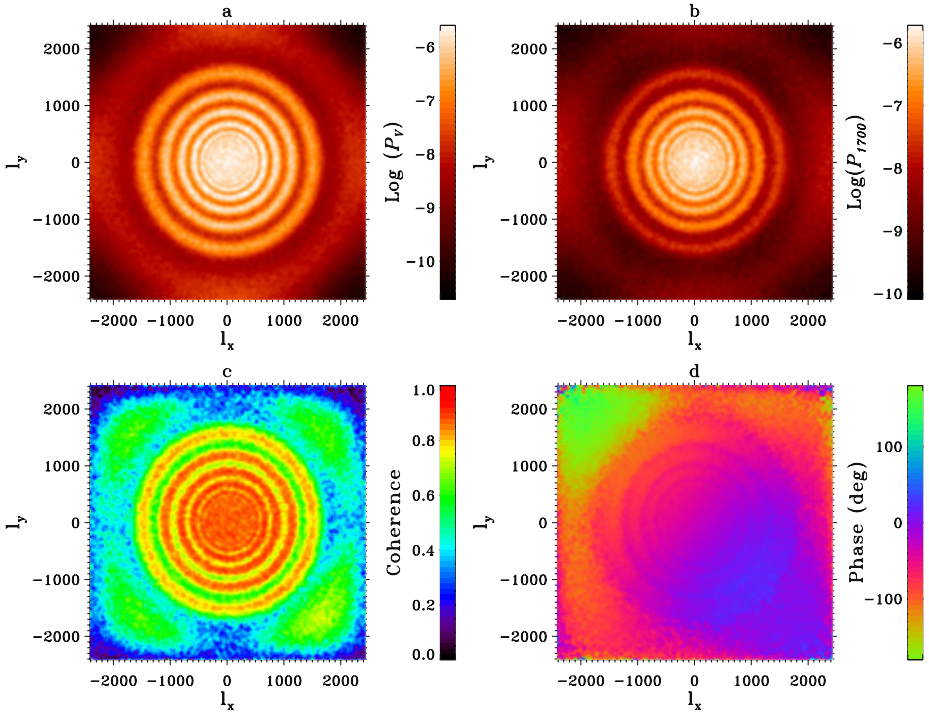


Figure 4 Example of ring diagram with AIA data compared to a ring diagram from HMI data at the same location. Shown is a slice at 5 mHz. An HMI ring from Doppler data (panel a) and AIA 1700 Å data (panel b) are shown. Coherence and phase maps are shown in panels (c) and (d), respectively.

4. Outlook

Ring-diagram analysis is a fairly mature technique in helioseismology. Among the local techniques in current usage it is the oldest, and it is also in many respects the easiest to use, particularly with regards to interpretation. The reader will note that a great deal of the work discussed in the preceding section was novel primarily for the amount of data involved and the statistical inferences made. Ring-diagram analysis has been fairly stable in the past five years. The next five years, however, are likely to see changes.

The most obvious change being assimilated into the field at the present time is the influx of massive quantities of data from the *Solar Dynamics Observatory* (SDO). The SDO/HMI project provides higher spatial resolution, higher temporal resolution, and better duty cycle than any previous experiment. It is extremely well calibrated and characterized. Assuming no unexpected contingencies, the quantity and quality of helioseismic data available for Solar Cycle 24 will be unprecedented.

The photospheric data from HMI are not the only new data available to the helioseismologist from SDO. The *Atmospheric Imaging Assembly* (AIA: Lemen *et al.*, 2012) provides high temporal-cadence high spatial-resolution data of the full solar disk at various ultraviolet wavelengths, imaging the chromosphere at various heights and raising the possibility of helioseismic studies in the solar atmosphere. The feasibility of multilayer helioseismology has already been demonstrated (Howe *et al.*, 2012). Ring-diagram analysis can be applied to this data. In Figure 4 we show example ring power spectra of HMI data compared with

the same region in the 1700 Å band AIA camera. As the AIA instrument was not designed as a helioseismic instrument, some care must be given to data calibration, but the problem is certainly tractable and is currently being addressed by a number of researchers.

One of the principal disadvantages of ring-diagram analysis when compared to time–distance analysis is the limited spatial resolution afforded by the ring-diagram aperture size. The classic techniques in ring-diagram analysis restrict us to one-dimensional inversions for flows or structure. A recently developed technique, however, harnesses tiled ring diagrams with different tile sizes and performs a three-dimensional inversion over the entire set. The work of Featherstone, Hindman, and Thompson (2011), discussed above, achieves both the higher resolution from small-size ring-diagram analysis, and the greater depth sensitivity from larger-size ring-diagram analysis.

Less optimistically, there are signs that ring-diagram analysis has reached the limits of its applicability using existing techniques. One of the prospective advances allowed by HMI data is the extension of useful helioseismic measurements to very large distances from disk center – in particular, achieving measurements at very high latitude. Ring-diagram mode-parameter measurements show as-yet unexplained inhomogeneities across the disk (Bogart *et al.*, 2012), which are likely due to various uncalibrated effects when attempting helioseismology at high inclination angles from the vertical. The effect of systematic errors in parameter estimation has been seen already in the apparent introduction of spurious meridional-flow cells by the differing B_0 angle over the course of a year (Zaatri *et al.*, 2006), and it is likely that uncertainties at the edges of the solar disk will give rise to similar errors. Work is needed, therefore, to extend ring-diagram analysis to higher latitudes.

The problems associated with structure inversions in the presence of high magnetic fields were mentioned in the previous section. This problem is not unique to ring diagrams, of course. In past years, treatments of the effects of magnetic fields have been hobbled by the prohibitive computational cost of treating magnetic fields in a realistic solar environment. Recently, however, as available computational power continues to grow and techniques to mature, a variety of useful simulations have proliferated (see reviews by Nordlund, Stein, and Asplund, 2009; Rempel and Schlichenmaier, 2011). These data are being used by a number of researchers to improve our techniques for inversions of helioseismic data in the presence of strong magnetic fields (*e.g.* Braun *et al.*, 2012).

Acknowledgements C.S.B. is supported by NASA grant NAS5-02139 to Stanford University. O.B., I.G.-H., F.H., K.J., R.W.K., and S.T. acknowledge support from NASA Grant # NNH12AT11I to NSO.

References

Antia, H.M., Basu, S.: 2007, Local helioseismology using ring diagram analysis. *Astron. Nachr.* **328**, 257. doi:10.1002/asna.200610727. ADS:2007AN....328..257A.

Baldner, C.S., Bogart, R.S., Basu, S.: 2011, Evidence for solar frequency dependence on sunspot type. *Astrophys. J. Lett.* **733**, L5. doi:10.1088/2041-8205/733/1/L5. ADS:2011ApJ...733L...5B.

Baldner, C.S., Bogart, R.S., Basu, S.: 2012, The sub-surface structure of a large sample of active regions. *Solar Phys.* doi:10.1007/s11207-012-0148-9.

Basu, S., Antia, H.M.: 2000, Possible solar cycle variations in the convection zone. *Solar Phys.* **192**, 449 – 458. doi:10.1023/A:1005232901528. ADS:2000SoPh..192..449B.

Basu, S., Antia, H.M.: 2003, Changes in solar dynamics from 1995 to 2002. *Astrophys. J.* **585**, 553 – 565. doi:10.1086/346020. ADS:2003ApJ...585..553B.

Basu, S., Antia, H.M.: 2010, Characteristics of solar meridional flows during solar cycle 23. *Astrophys. J.* **717**, 488 – 495. doi:10.1088/0004-637X/717/1/488. ADS:2010ApJ...717..488B.

Basu, S., Antia, H.M., Bogart, R.S.: 2004, Ring-diagram analysis of the structure of solar active regions. *Astrophys. J.* **610**, 1157 – 1168. doi:10.1086/421843. ADS:2004ApJ...610.1157B.

Basu, S., Antia, H.M., Bogart, R.S.: 2007, Structure of the near-surface layers of the Sun: asphericity and time variation. *Astrophys. J.* **654**, 1146 – 1165. doi:10.1086/509251. ADS:2007ApJ...654.1146B.

Basu, S., Antia, H.M., Tripathy, S.C.: 1999, Ring diagram analysis of near-surface flows in the Sun. *Astrophys. J.* **512**, 458 – 470. doi:10.1086/306765. ADS:1999ApJ...512..458B.

Bogart, R.S., Basu, S., Rabello-Soares, M.C., Antia, H.M.: 2008, Probing the subsurface structures of active regions with ring-diagram analysis. *Solar Phys.* **251**, 439 – 451. doi:10.1007/s11207-008-9213-9. ADS:2008SoPh..251..439B.

Bogart, R.S., Baldner, C., Basu, S., Haber, D.A., Rabello-Soares, M.C.: 2011a, HMI ring diagram analysis I. The processing pipeline. *J. Phys.* **CS-271**(1), 012008. doi:10.1088/1742-6596/271/1/012008. ADS:2011JPhCS.271a2008B.

Bogart, R.S., Baldner, C., Basu, S., Haber, D.A., Rabello-Soares, M.C.: 2011b, HMI ring diagram analysis II. Data products. *J. Phys.* **CS-271**(1), 012009. doi:10.1088/1742-6596/271/1/012009. ADS:2011JPhCS.271a2009B.

Bogart, R.S., Baldner, C.S., Basu, S., Gonzalez-Hernandez, I., Haber, D.A., Hill, F., Howe, R., Jain, K., Komm, R.W., Rabello-Soares, M.C., Tripathy, S.: 2012, American Astronomical Society Meeting Abstracts #220, 220, 205.12.

Braun, D.C., Duvall, T.L. Jr., Labonte, B.J.: 1987, Acoustic absorption by sunspots. *Astrophys. J. Lett.* **319**, L27 – L31. doi:10.1086/184949. ADS:1987ApJ...319L..27B.

Braun, D.C., Birch, A.C., Rempel, M., Duvall, T.L.: 2012, Helioseismology of a realistic magnetoconvective sunspot simulation. *Astrophys. J.* **744**, 77. doi:10.1088/0004-637X/744/1/77. ADS:2012ApJ...744...77B.

Burtseva, O., Tripathy, S.C., Hill, F., Kholikov, S., Raouafi, N.-E., Lindsey, C.: 2009, Amplitudes of high-degree p modes in the quiet and active Sun. In: Dikpati, M., Arentoft, T., González Hernández, I., Lindsey, C., Hill, F. (eds.) *Solar-Stellar Dynamos as Revealed by Helio- and Asteroseismology: GONG 2008/SOHO 21* **CS-416**, Astron. Soc. Pac., San Francisco, 293. ADS:2009ASPC..416..293B.

Burtseva, O., Tripathy, S., Howe, R., Jain, K., Hill, F., Bogart, R., Rabello-Soares, M.C.: 2011, Temporal variations of high-degree solar p-modes from GONG and MDI. *J. Phys.* **CS-271**(1), 012012. doi:10.1088/1742-6596/271/1/012012. ADS:2011JPhCS.271a2012B.

Charbonneau, P.: 2010, Dynamo models of the solar cycle. *Living Rev. Solar Phys.* **7**, 3. http://solarphysics.livingreviews.org/Articles/lrsp-2010-3/ (cited on 2 October 2012). ADS: 2010LRSP....7....3C.

Chou, D.-Y., Dai, D.-C.: 2001, Solar cycle variations of subsurface meridional flows in the Sun. *Astrophys. J. Lett.* **559**, L175 – L178. doi:10.1086/323724. ADS:2001ApJ...559L.175C.

Duvall, T.L. Jr., Jefferies, S.M., Harvey, J.W., Pomerantz, M.A.: 1993, Time-distance helioseismology. *Nature* **362**, 430 – 432. doi:10.1038/362430a0. ADS:1993Natur.362..430D.

Featherstone, N.A., Hindman, B.W., Thompson, M.J.: 2011, Ring-analysis flow measurements of sunspot outflows. *J. Phys.* **CS-271**(1), 012002. doi:10.1088/1742-6596/271/1/012002. ADS:2011JPhCS.271a2002F.

Ferguson, R., Komm, R., Hill, F., Barnes, G., Leka, K.D.: 2009, Subsurface flow properties of flaring *versus* flare-quiet active regions. In: Dikpati, M., Arentoft, T., González Hernández, I., Lindsey, C., Hill, F. (eds.) *Solar-Stellar Dynamos as Revealed by Helio- and Asteroseismology: GONG 2008/SOHO 21* **CS-416**, Astron. Soc. Pac., San Francisco, 127. ADS:2009ASPC..416..127F.

Giles, P.M.: 2000, Time–distance measurements of large-scale flows in the solar convection zone. PhD thesis, Stanford University. ADS:2000PhDT.....9G.

Gizon, L.: 2003, Probing flows in the upper solar convection zone. PhD thesis, Stanford University. ADS:2003PhDT.....9G.

Gizon, L., Birch, A.C.: 2005, Local helioseismology. *Living Rev. Solar Phys.* **2**, 6. http://www.livingreviews.org/lrsp-2005-6 (cited on 2 October 2012). ADS:2005LRSP....2....6G.

Gizon, L., Schunker, H., Baldner, C.S., Basu, S., Birch, A.C., Bogart, R.S., Braun, D.C., Cameron, R., Duvall, T.L., Hanasoge, S.M., Jackiewicz, J., Roth, M., Stahn, T., Thompson, M.J., Zharkov, S.: 2009, Helioseismology of sunspots: a case study of NOAA region 9787. *Space Sci. Rev.* **144**, 249 – 273. doi:10.1007/s11214-008-9466-5. ADS:2009SSRv..144..249G.

González Hernández, I.: 2008, Subsurface flows from ring diagram analysis. *J. Phys.* **CS-118**(1), 012034. doi:10.1088/1742-6596/118/1/012034. ADS:2008JPhCS.118a2034G.

González Hernández, I., Kholikov, S., Hill, F., Howe, R., Komm, R.: 2008, Subsurface meridional circulation in the active belts. *Solar Phys.* **252**, 235 – 245. doi:10.1007/s11207-008-9264-y. ADS:2008SoPh..252..235G.

González Hernández, I., Howe, R., Komm, R., Hill, F.: 2010, Meridional circulation during the extended solar minimum: another component of the torsional oscillation? *Astrophys. J. Lett.* **713**, L16 – L20. doi:10.1088/2041-8205/713/1/L16. ADS:2010ApJ...713L..16G.

Haber, D.A., Hindman, B.W., Toomre, J., Bogart, R.S., Thompson, M.J., Hill, F.: 2000a, Solar shear flows deduced from helioseismic dense-pack samplings of ring diagrams. *Solar Phys.* **192**, 335 – 350. ADS:2000SoPh..192..335H.

Haber, D.A., Hindman, B.W., Toomre, J., Bogart, R.S., Thompson, M.J., Hill, F.: 2000b, Solar shear flows deduced from helioseismic dense-pack samplings of ring diagrams. *Solar Phys.* **192**, 335 – 350. ADS:2000SoPh..192..335H.

Haber, D.A., Hindman, B.W., Toomre, J., Bogart, R.S., Larsen, R.M., Hill, F.: 2002, Evolving submerged meridional circulation cells within the upper convection zone revealed by ring-diagram analysis. *Astrophys. J.* **570**, 855 – 864. doi:10.1086/339631. ADS:2002ApJ...570..855H.

Harvey, J.W., Hill, F., Hubbard, R.P., Kennedy, J.R., Leibacher, J.W., Pintar, J.A., Gilman, P.A., Noyes, R.W., Title, A.M., Toomre, J., Ulrich, R.K., Bhatnagar, A., Kennewell, J.A., Marquette, W., Patron, J., Saa, O., Yasukawa, E.: 1996, The Global Oscillation Network Group (GONG) project. *Science* **272**, 1284 – 1286. doi:10.1126/science.272.5266.1284. ADS:1996Sci...272.1284H.

Hathaway, D.H.: 1996, Doppler measurements of the Sun's meridional flow. *Astrophys. J.* **460**, 1027. doi:10.1086/177029. ADS:1996ApJ...460.1027H.

Hill, F.: 1988, Rings and trumpets – three-dimensional power spectra of solar oscillations. *Astrophys. J.* **333**, 996 – 1013. doi:10.1086/166807. ADS:1988ApJ...333..996H.

Hill, F.: 1989, Solar oscillation ring diagrams and large-scale flows. *Astrophys. J. Lett.* **343**, L69 – L71. doi:10.1086/185513. ADS:1989ApJ...343L..69H.

Hindman, B.W., Haber, D.A., Toomre, J.: 2009, Subsurface circulations within active regions. *Astrophys. J.* **698**, 1749 – 1760. doi:10.1088/0004-637X/698/2/1749. ADS:2009ApJ...698.1749H.

Howe, R., Komm, R.W., Hill, F., Haber, D.A., Hindman, B.W.: 2004, Activity-related changes in local solar acoustic mode parameters from Michelson Doppler Imager and Global Oscillations Network Group. *Astrophys. J.* **608**, 562 – 579. doi:10.1086/392525. ADS:2004ApJ...608..562H.

Howe, R., Jain, K., Bogart, R.S., Haber, D.A., Baldner, C.S.: 2012, Two-dimensional power, phase, and coherence spectra of solar dynamics observatory helioseismic observables. *Solar Phys.* doi:10.1007/s11207-012-0097-3. ADS:2012SoPh..tmp..200H.

Komm, R.W., Howard, R.F., Harvey, J.W.: 1993, Meridional flow of small photospheric magnetic features. *Solar Phys.* **147**, 207 – 223. doi:10.1007/BF00690713. ADS:1993SoPh..147..207K.

Komm, R., Howe, R., Hill, F.: 2009, Emerging and decaying magnetic flux and subsurface flows. *Solar Phys.* **258**, 13 – 30. doi:10.1007/s11207-009-9398-6. ADS:2009SoPh..258...13K.

Komm, R., Howe, R., Hill, F.: 2011, Subsurface velocity of emerging and decaying active regions. *Solar Phys.* **268**, 407 – 428. doi:10.1007/s11207-010-9692-3. ADS:2011SoPh..268..407K.

Komm, R., Howe, R., Hill, F.: 2012, Vorticity of subsurface flows of emerging and decaying active regions. *Solar Phys.* **277**, 205–226. doi:10.1007/s11207-011-9920-5. ADS:2011SoPh..tmp..426K.

Komm, R., Corbard, T., Durney, B.R., González Hernández, I., Hill, F., Howe, R., Toner, C.: 2004, Solar subsurface fluid dynamics descriptors derived from Global Oscillation Network Group and Michelson Doppler Imager data. *Astrophys. J.* **605**, 554 – 567. doi:10.1086/382187. ADS:2004ApJ...605..554K.

Komm, R., Morita, S., Howe, R., Hill, F.: 2008, Emerging active regions studied with ring-diagram analysis. *Astrophys. J.* **672**, 1254 – 1265. doi:10.1086/523998. ADS:2008ApJ...672.1254K.

Komm, R., Ferguson, R., Hill, F., Barnes, G., Leka, K.D.: 2011, Subsurface vorticity of flaring *versus* flare-quiet active regions. *Solar Phys.* **268**, 389 – 406. doi:10.1007/s11207-010-9552-1. ADS:2011SoPh..268..389K.

Kosovichev, A.G., Basu, S., Bogart, R., Duvall, T.L. Jr., Gonzalez-Hernandez, I., Haber, D., Hartlep, T., Howe, R., Komm, R., Kholikov, S., Parchevsky, K.V., Tripathy, S., Zhao, J.: 2011, Local helioseismology of sunspot regions: comparison of ring-diagram and time–distance results. *J. Phys.* **CS-271**(1), 012005. doi:10.1088/1742-6596/271/1/012005. ADS:2011JPhCS.271a2005K.

Lemen, J.R., Title, A.M., Akin, D.J., Boerner, P.F., Chou, C., Drake, J.F., Duncan, D.W., Edwards, C.G., Friedlaender, F.M., Heyman, G.F., Hurlburt, N.E., Katz, N.L., Kushner, G.D., Levay, M., Lindgren, R.W., Mathur, D.P., McFeaters, E.L., Mitchell, S., Rehse, R.A., Schrijver, C.J., Springer, L.A., Stern, R.A., Tarbell, T.D., Wuelser, J.-P., Wolfson, C.J., Yanari, C., Bookbinder, J.A., Cheimets, P.N., Caldwell, D., Deluca, E.E., Gates, R., Golub, L., Park, S., Podgorski, W.A., Bush, R.I., Scherrer, P.H., Gummin, M.A., Smith, P., Auker, G., Jerram, P., Pool, P., Soufli, R., Windt, D.L., Beardsley, S., Clapp, M., Lang, J., Waltham, N.: 2012, The Atmospheric Imaging Assembly (AIA) on the Solar Dynamics Observatory (SDO). *Solar Phys.* **275**, 17 – 40. doi:10.1007/s11207-011-9776-8. ADS:2012SoPh..275...17L.

Lin, C.-H., Basu, S., Li, L.: 2009, Interpreting helioseismic structure inversion results of solar active regions. *Solar Phys.* **257**, 37 – 60. doi:10.1007/s11207-009-9332-y. ADS:2009SoPh..257...37L.

Lindsey, C., Braun, D.C.: 1990, Helioseismic imaging of sunspots at their antipodes. *Solar Phys.* **126**, 101 – 115. doi:10.1007/BF00158301. ADS:1990SoPh..126..101L.

Mason, D., Komm, R., Hill, F., Howe, R., Haber, D., Hindman, B.W.: 2006, Flares, magnetic fields, and subsurface vorticity: a survey of GONG and MDI data. *Astrophys. J.* **645**, 1543 – 1553. doi:10.1086/503761. ADS:2006ApJ...645.1543M.

Maurya, R.A., Ambastha, A., Tripathy, S.C.: 2009, Variations in p-mode parameters with changing onset time of a large flare. *Astrophys. J. Lett.* **706**, L235 – L239. doi:10.1088/0004-637X/706/2/L235. ADS:2009ApJ...706L.235M.

Miesch, M.S.: 2005, Large-scale dynamics of the convection zone and tachocline. *Living Rev. Solar Phys.* **2**, 1. http://www.livingreviews.org/lrsp-2005-1 (cited on 2 October 2012). ADS:2005LRSP....2....1M.

Moradi, H., Baldner, C., Birch, A.C., Braun, D.C., Cameron, R.H., Duvall, T.L., Gizon, L., Haber, D., Hanasoge, S.M., Hindman, B.W., Jackiewicz, J., Khomenko, E., Komm, R., Rajaguru, P., Rempel, M., Roth, M., Schlichenmaier, R., Schunker, H., Spruit, H.C., Strassmeier, K.G., Thompson, M.J., Zharkov, S.: 2010, Modeling the subsurface structure of sunspots. *Solar Phys.* **267**, 1 – 62. doi:10.1007/s11207-010-9630-4. ADS:2010SoPh..267....1M.

Nigam, R., Kosovichev, A.G.: 1998, Measuring the Sun's eigenfrequencies from velocity and intensity helioseismic spectra: asymmetrical line profile-fitting formula. *Astrophys. J. Lett.* **505**, L51. doi:10.1086/311594. ADS:1998ApJ...505L..51N.

Nordlund, Å., Stein, R.F., Asplund, M.: 2009, Solar surface convection. *Living Rev. Solar Phys.* **6**, 2. http://www.livingreviews.org/lrsp-2009-2 (cited on 2 October 2012). ADS:2009LRSP....6....2N.

Rabello-Soares, M.C.: 2012, Solar-cycle variation of sound speed near the solar surface. *Astrophys. J.* **745**, 184. doi:10.1088/0004-637X/745/2/184. ADS:2012ApJ...745..184R.

Rabello-Soares, M.C., Bogart, R.S., Basu, S.: 2008, Analysis of the characteristics of solar oscillation modes in active regions. *J. Phys.* **CS-118**(1), 012084. doi:10.1088/1742-6596/118/1/012084. ADS:2008JPhCS.118a2084R.

Rajaguru, S.P., Basu, S., Antia, H.M.: 2001, Ring diagram analysis of the characteristics of solar oscillation modes in active regions. *Astrophys. J.* **563**, 410 – 418. doi:10.1086/323780. ADS:2001ApJ...563..410R.

Reinard, A.A., Henthorn, J., Komm, R., Hill, F.: 2010, Evidence that temporal changes in solar subsurface helicity precede active region flaring. *Astrophys. J. Lett.* **710**, L121 – L125. doi:10.1088/2041-8205/710/2/L121. ADS:2010ApJ...710L.121R.

Rempel, M., Schlichenmaier, R.: 2011, Sunspot modeling: from simplified models to radiative MHD simulations. *Living Rev. Solar Phys.* **8**, 4. http://www.livingreviews.org/lrsp-2011-3 (cited on 2 October 2012). ADS:2011LRSP....8....3R.

Scherrer, P.H., Bogart, R.S., Bush, R.I., Hoeksema, J.T., Kosovichev, A.G., Schou, J., Rosenberg, W., Springer, L., Tarbell, T.D., Title, A., Wolfson, C.J., Zayer, I., MDI Engineering Team: 1995, The solar oscillations investigation – Michelson Doppler Imager. *Solar Phys.* **162**, 129 – 188. doi:10.1007/BF00733429. ADS:1995SoPh..162..129S.

Scherrer, P.H., Schou, J., Bush, R.I., Kosovichev, A.G., Bogart, R.S., Hoeksema, J.T., Liu, Y., Duvall, T.L., Zhao, J., Title, A.M., Schrijver, C.J., Tarbell, T.D., Tomczyk, S.: 2012, The Helioseismic and Magnetic Imager (HMI) investigation for the Solar Dynamics Observatory (SDO). *Solar Phys.* **275**, 207 – 227. doi:10.1007/s11207-011-9834-2. ADS:2012SoPh..275..207S.

Snodgrass, H.B., Dailey, S.B.: 1996, Meridional motions of magnetic features in the solar photosphere. *Solar Phys.* **163**, 21 – 42. doi:10.1007/BF00165454. ADS:1996SoPh..163...21S.

Tripathy, S.C., Jain, K., Hill, F.: 2011, Variation of high-degree mode frequencies during the declining phase of solar cycle 23. *J. Phys.* **CS-271**(1), 012024. doi:10.1088/1742-6596/271/1/012024. ADS:2011JPhCS.271a2024T.

Ulrich, R.K., Boyden, J.E.: 2005, The solar surface toroidal magnetic field. *Astrophys. J. Lett.* **620**, L123 – L127. doi:10.1086/428724. ADS:2005ApJ...620L.123U.

Zaatri, A., Komm, R., González Hernández, I., Howe, R., Corbard, T.: 2006, North south asymmetry of zonal and meridional flows determined from ring diagram analysis of Gong++ data. *Solar Phys.* **236**, 227 – 244. doi:10.1007/s11207-006-0106-5. ADS:2006SoPh..236..227Z.

DOI 10.1007/978-1-4899-8005-2_5
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0010-0

Subsurface Supergranular Vertical Flows as Measured Using Large Distance Separations in Time–Distance Helioseismology

T.L. Duvall Jr. · S.M. Hanasoge

Received: 1 March 2012 / Accepted: 24 April 2012 / Published online: 15 June 2012

Abstract As large-distance rays (say, $10-24^\circ$) approach the solar surface approximately vertically, travel times measured from surface pairs for these large separations are mostly sensitive to vertical flows, at least for shallow flows within a few Mm of the solar surface. All previous analyses of supergranulation have used smaller separations and have been hampered by the difficulty of separating the horizontal and vertical flow components. We find that the large-separation travel times associated with supergranulation cannot be studied using the standard phase-speed filters of time–distance helioseismology. These filters, whose use is based upon a refractive model of the perturbations, reduce the resultant travel-time signal by at least an order of magnitude at some distances. More effective filters are derived. Modeling suggests that the center–annulus travel-time difference $[\delta t_{\mathrm{oi}}]$ in the separation range $\Delta = 10-24^\circ$ is insensitive to the horizontally diverging flow from the centers of the supergranules and should lead to a constant signal from the vertical flow. Our measurement of this quantity, 5.1 ± 0.1 seconds, is constant over the distance range. This magnitude of the signal cannot be caused by the level of upflow at cell centers seen at the photosphere of 10 m s^{-1} extended in depth. It requires the vertical flow to increase with depth. A simple Gaussian model of the increase with depth implies a peak upward flow of 240 m s^{-1} at a depth of 2.3 Mm and a peak horizontal flow of 700 m s^{-1} at a depth of 1.6 Mm.

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

T.L. Duvall Jr. (✉)
Solar Physics Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
e-mail: Thomas.L.Duvall@nasa.gov

S.M. Hanasoge
Max-Planck-Institut für Sonnensystemforschung, Max Planck Straße 2, 37191 Katlenburg-Lindau, Germany
e-mail: hanasoge@mps.mpg.de

S.M. Hanasoge
Department of Geosciences, Princeton University, Princeton, NJ 08544, USA

Keywords Helioseismology, observations · Helioseismology, direct modeling · Interior, convective zone · Supergranulation · Velocity fields, interior

1. Introduction

Supergranulation, first studied more than 50 years ago (Hart, 1954; Leighton, Noyes, and Simon, 1962), continues to be an active area of research. A comprehensive review (Rieutord and Rincon, 2010) details the recent progress. It has proven difficult to measure the vertical flow of supergranules at the photospheric level. The recent measurements of a 10 $\mathrm{m\,s^{-1}}$ upflow at the center of the average cell with a horizontal variation consistent with a simple convection cell (Duvall and Birch, 2010) may have finally settled the issue. It is only through helioseismology that we would hope to measure the supergranulation flows below the photosphere. Duvall *et al.* (1997) first showed that time–distance helioseismic techniques are sensitive to supergranules and that inversions to derive three-dimensional flow fields might be derived. However, Zhao and Kosovichev (2003) showed that their ray-theory inversions could not separate the horizontal and vertical flows for a model flow field. The use of radial-order filters and Born approximation kernels has led to more successful separation (Jackiewicz, Gizon, and Birch, 2008; Švanda *et al.*, 2011). Gizon, Birch, and Spruit (2010) summarize the local helioseismic contributions to supergranulation.

An important quantity in time–distance helioseismology is the arc separation [Δ] between pairs of photospheric locations whose signals are subsequently temporally cross correlated. Previous analyses used Δ less than 5° (61 Mm) (Zhao and Kosovichev, 2003) or 2.4° (29.2 Mm) (Jackiewicz, Gizon, and Birch, 2008; Švanda *et al.*, 2011), as the sensitivity is greater for the horizontal flow, the signal-to-noise ratio is better for small separations, and basically no supergranulation signal could be seen at larger Δ in the travel-time difference maps constructed from the relatively short (8 – 12 hours) time intervals required to study the one-day lifetime supergranules. In the present work, we show that the large separations of $\Delta = 10 - 24°$ yield center–annulus travel-time differences [$\delta t_{\mathrm{oi}} \equiv$ outward-going time minus inward-going time] from the centers of average supergranules that are insensitive to the diverging horizontal flow and hence yield a purely vertical flow response.

In Section 2 the modeling efforts, including flow models that satisfy mass conservation, travel times computed from ray theory, and linear wave simulations through flow fields are presented. In Section 3, the analysis of the data is presented, and in Section 4 the results are discussed.

2. Modeling

2.1. Flow Model

As our observational technique is to make averages about centers of supergranules, the flow model that we take is azimuthally symmetric and decays exponentially away from the cell center to simulate the effect of averaging a stochastic flow field. A time-independent flow is assumed with axisymmetry about the location of the average cell. A Cartesian coordinate system with z positive upwards is defined. Mass conservation is used to constrain the

flow model, leading to the following connection between vertical flow, horizontal flow, and density:

$$\frac{\partial}{\partial z}(\rho v_z) = -\rho \nabla_{\mathbf{h}} \cdot \mathbf{v_h}, \tag{1}$$

where ρ is the z-dependent density, v_z is the z-component of velocity, $\nabla_{\mathbf{h}}$ is the horizontal divergence, and $\mathbf{v_h}$ is the horizontal velocity. It is assumed that horizontal density variations can be ignored, as travel times are only weakly sensitive to density changes in the underlying medium (Hanasoge *et al.*, 2012). The model is assumed to separate into horizontal and vertical functions

$$\mathbf{v_h} \equiv -f(z)\mathbf{g}(x, y), \tag{2}$$

$$v_z \equiv u(z)\nabla_{\mathbf{h}} \cdot \mathbf{g}, \tag{3}$$

where $\mathbf{g}(x, y)$ is a vector in the horizontal plane with no units, $f(z)$ has units of velocity, and $u(z)$ has units of velocity times length. Our method of solution is to first specify the horizontal function $\mathbf{g}(x, y)$ and calculate its horizontal divergence $\nabla_{\mathbf{h}} \cdot \mathbf{g}$. The vertical function $u(z)$ is then specified, and Equation (1) is used to derive $f(z)$. Some straightforward algebra yields

$$f(z) = \frac{\partial u}{\partial z} + u(z)\frac{\partial \ln \rho}{\partial z}. \tag{4}$$

In general, models are considered with the horizontal function $\mathbf{g}(x, y)$ defined by

$$\mathbf{g}(x, y) = \hat{\mathbf{r}} J_1(kr)\, \mathrm{e}^{-r/R}, \tag{5}$$

where $\hat{\mathbf{r}}$ is the outward radial unit vector in a cylindrical coordinate system, J_1 is the Bessel function of order one, k is a wavenumber, r is the horizontal distance from the origin, and R is a decay length. k and R are, in principle, free parameters that will only have the default values $k = \frac{2\pi}{30}$ rad Mm^{-1} and $R = 15$ Mm in this article. This type of horizontal variation was used by Birch and Gizon (2007). This type of horizontal variation describes well the type of averaging of the supergranular field done in this article, where one sees the outward flow as the first positive lobe of the J_1 function and, on average, the inflow from the adjacent supergranular cells as the first negative range of the J_1 function.

In general, models have been considered with a Gaussian z-dependence. v_z is specified at $r = 0$ as a simple Gaussian:

$$v_z(r = 0) = ku(z) = v_0\, \mathrm{e}^{-(z-z_0)^2/2\sigma_z^2}, \tag{6}$$

where z_0 is the location of the peak of the vertical flow, σ_z is the Gaussian sigma, and v_0 is the maximum vertical flow. To ultimately explain the large-distance travel times, the photosphere needs to be in the far tail of the Gaussian. As the upward flow at cell center is 10 m s^{-1} at the photosphere, this implies a considerably larger vertical flow at depth for the average cell. Values for these parameters that approximate the data are $v_0 = 240$ m s^{-1}, $z_0 = -2.3$ Mm, and $\sigma_z = 0.912$ Mm.

2.2. Ray Calculations

In ray theory, the travel-time difference for the two directions of propagation through a flow $\mathbf{v}$ is

$$\delta\tau = -2\int_\Gamma \frac{\mathbf{v} \cdot \mathrm{d}\mathbf{s}}{c^2}, \tag{7}$$

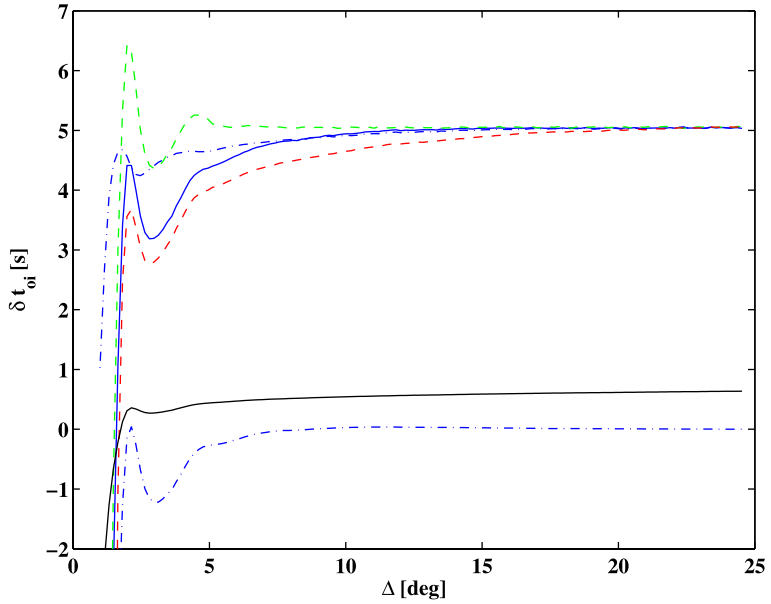


Figure 1 Comparison of various ray models. The blue curves are for the Gaussian model with $z_0 = -2.3$ Mm. The lower dot–dashed blue curve is for the horizontal flow component, the upper dot–dashed blue curve is for the vertical flow component and the solid blue curve is for the sum (upper solid curve). The upper dashed (green) curve is for the sum for $z_0 = -1.15$ Mm model and the lower dashed (red) curve is for the sum of the $z_0 = -3.45$ Mm model. The (lower) solid black curve is the sum for a model with a constant upflow of 10 m s^{-1}.

where Γ is the ray path, c is the sound speed, and d**s** is the element of length in the direction of the ray (Giles, 2000). This equation naturally separates into two terms for the horizontal and vertical flow contributions to the travel time. The ray generation and raw ray integrations are performed using code developed and discussed in detail by Giles (2000). This code was extended to integrate quadrant and annulus surface-focusing integrations for 3D flow and sound-speed perturbation models. In Figure 1 the black curve shows the travel-time difference $[\delta t_{\rm oi}]$ for a flow model with a constant upflow of 10 m s^{-1}. The black curve is for the sum of horizontal and vertical flows. That is, a horizontal flow of the form of Equation (5) consistent through the continuity equation with a 10 m s^{-1} upflow is used. This signal, with a maximum of 0.6 seconds, is much too small to explain the results derived in Section 3.3. This implies that the vertical flow must increase with depth from the photospheric value of 10 m s^{-1}.

Also in Figure 1, the separate horizontal and vertical travel-time contributions and the sum are shown for the Gaussian model mentioned above. For this relatively shallow model, it is seen that in the distance range $10-24^\circ$, the signal is essentially constant and is mostly due to the vertical flow. Towards shorter distances, the horizontal flow contributes an increasing amount. The Gaussian model has three free parameters: v_0, z_0, and σ_z. In a later section the travel-time difference for $10-24^\circ$ will be determined to be 5.1 seconds. This travel-time difference and the 10 m s^{-1} upflow at the photosphere serve to determine two of the three free parameters. The travel time at $\Delta = 24.5^\circ$ is forced to 5.1 seconds. Letting z_0 be the free parameter, three models are plotted in Figure 1 with values of $z_0 = -1.15$ Mm, $z_0 = -2.30$ Mm, and $z_0 = -3.45$ Mm. The shallowest model shows the least contribution

Table 1 Model details. The three models g1, g2, g3 are ones with Gaussian (g) upward flow at cell center peaking at $z0$ Mm with peak flow $v_z(z0)$ m s^{-1} and with Gaussian width σ_z Mm. The upward flow at the top ($z = 0$) is $v_z(0)$ m s^{-1}. The peak horizontal flow at the top is $v_h^{\max}(0)$ m s^{-1} which occurs at radial distance $r^{\max}$ Mm. The peak outward flow [$v_h^{\max}$ m s^{-1}] occurs at $z(v_h^{\max})$ Mm. All of the maxima of v_h occur at the same $r^{\max}$ by construction. The model sim1 is the one for the small features with just a z-flow that are used for the flow perturbation described in Section 2.3.

Name	z_0	σ_z	$v_z(0)$	$v_h^{\max}(0)$	$r^{\max}$	$v_z(z0)$	$v_h^{\max}$	$z(v_h^{\max})$
g1	−3.45	1.39	10	122	7	218	460	−2.44
g2	−2.3	0.912	10	138	7	240	697	−1.62
g3	−1.15	0.433	10	195	7	340	1609	−0.83
sim1	−2.3	0.816	6.3			338		

from the horizontal component in the distance range 5 – 10°. This distance range may supply a way to distinguish the best model while comparing with data.

A summary of the characteristics of these models is contained in Table 1. The maximum vertical flow is at least a factor of 20 larger than the photospheric value. The ratio of the maximum horizontal flow to the photospheric value varies between four and eight. The maxima of the horizontal flow occurs somewhat shallower than the peak vertical flow.

2.3. Simulation Results

The applicability of ray theory to time–distance helioseismology has been called into question (Bogdan, 1997). There are cases for supergranulation studies (Birch and Gizon, 2007, Figure 5) where it works reasonably well (amplitude of ray theory 25 % higher than Born-approximation kernels) and others where it is discrepant by a factor of two (Birch and Gizon, 2007, Figure 6). One case that works extremely well is the comparison of interior rotation determined from global modes and time–distance inversions (Giles, 2000, Figure 6.2) in the radius range $0.89 < r/R_\odot < 0.999$. The difference between these cases seems to be that the ray theory works better when the spatial variation of the perturbation is larger, that is, much larger than the acoustic wavelength. One way to check the ray theory for a particular case is to compare with travel times computed from Born kernels (Birch and Gizon, 2007). Another way to check is to perform simulations on a convectively stabilized solar model with specified flow-velocity perturbations and to propagate acoustic waves through the model (Hanasoge, Duvall, and Couvidat, 2007). Travel times are computed from the simulations and then compared with ray-theory calculations of the travel times using the same flow velocity perturbations. This is the type of checking of the ray theory adopted here.

A global simulation of wave propagation (Hanasoge and Duvall, 2007) with 768 × 384 grid points in longitude and latitude, respectively, is performed over a ten-hour period. The flow perturbation consists of Gaussian features with a constant direction of vertical flow centered at a depth of $z_0 = -2.30$ Mm with $\sigma_z = 0.82$ Mm and with horizontal $\sigma_h = 5.1$ Mm. There are 500 of these features placed at random longitude–latitude pairs at that depth [z_0]. Center–annulus travel-time differences are measured from the simulation and averaged about the known locations. Realization noise is removed to first order by doing a similar simulation with no flow perturbation but with the same source excitations (Werne, Birch, and Julien, 2004; Hanasoge, Duvall, and Couvidat, 2007) and subtracting the resultant $\delta t_{\rm oi}$ to obtain noise-corrected results.

The flow model that is inserted into the simulation is used with Equation (7) to derive ray-theory predictions of the travel-time differences. The results of the ray-theory computations

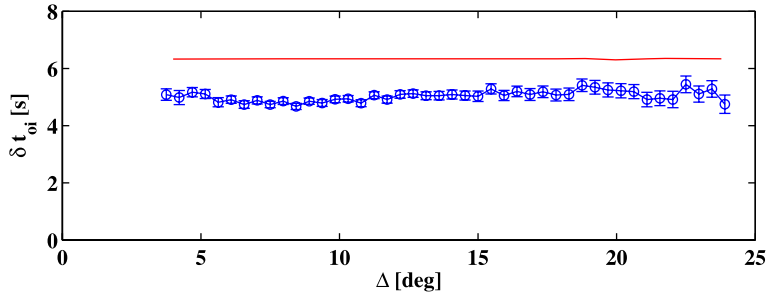


Figure 2 Comparison of the center–annulus travel-time differences [$\delta t_{\rm oi}$] from the linear simulation (blue circles with error bars) with the travel-time difference computed from the ray theory with the same flow perturbations (red solid line). The error bars are computed from the scatter far from the feature locations. No filtering has been done before the travel-time measurements. The average travel-time difference in the range $\Delta = 10-24^\circ$ has been scaled to match the observationally determined mean of 5.1 seconds. The same scaling factor is then used to scale the ray theory results. In the range $\Delta = 10-24^\circ$, the ray theory predicts a travel-time difference too large by 24 %.

and the travel-time difference measurements from the simulation are shown in Figure 2. As the simulation only contains vertical flows, it is expected that there would be little or no variation of the travel-time differences with Δ. The ray theory does predict too large a travel-time difference by 24 %. This excess amplitude is similar to that found by Birch and Gizon (2007). Although the amplitude computed from the ray theory is not precise, the necessity of large vertical flows to generate a travel-time difference of 5.1 seconds at $\Delta = 10-24^\circ$ is confirmed. The peak vertical flow for the normalized case is 338 m s^{-1}. The model parameters are detailed in Table 1.

3. Data Reduction and Analysis

3.1. Reduction

Dopplergrams from the *Helioseismic and Magnetic Imager* (HMI: Schou *et al.*, 2012) onboard the *Solar Dynamics Observatory* (SDO) spacecraft were analyzed for the present work. 32 days (10 June – 11 July 2010) of Dopplergrams were used to derive the final results, although a number of tests subsequently mentioned used only the last three days: 9 – 11 July 2010. This particular time period was used as the Sun was very quiet (sunspot number $RI \approx 15$ for these two months), and it was the final two-month continuous-coverage period for the *Michelson Doppler Imager* (MDI: Scherrer *et al.*, 1995). It might be useful to compare the results from the two instruments, but only the HMI data were used for the present study.

Raw Dopplergrams have a nine-day averaged Dopplergram subtracted as well as the spacecraft velocity which has a significant 24-hour component. These corrected Dopplergrams are remapped onto a coordinate system with equal spacing in latitude and longitude of 0.03° covering a range of 144° in both latitude and longitude. The remapping is achieved using a bilinear interpolation. The remapping in longitude is onto the Carrington system at a central longitude that crosses central meridian at the middle of the twelve-hour interval used. Two twelve-hour intervals are used for each day, covering the first and second halves of the day. Each remapped image is Fourier-filtered and resampled at 0.24° per pixel. Twelve-hour datacubes are constructed from these individual images with 600 × 600 spatial points and

960 temporal points for the 45-second sampling. This procedure works well for the large Δ emphasized in the present study, but would need to be modified to study smaller Δ. Studying smaller Δ is deferred until future work.

The centers of the supergranules are located by a procedure used previously (Duvall and Birch, 2010). The datacubes are spatio-temporally filtered to pass just the solar f-mode oscillations. Center–annulus travel-time differences are computed (Duvall and Gizon, 2000) for the distance range $0.48 - 1.02°$. By using the difference of inward-going times minus outward-going times, the maps are equivalent to maps of the horizontal divergence. The equivalence of these maps to maps of supergranules has been shown before (Duvall and Gizon, 2000; Duvall and Birch, 2010). The travel–time difference maps are smoothed with a Gaussian of $\sigma = 2.9$ Mm to more easily determine the cell-center locations. Local maxima of this signal are picked as the cell centers. Lists of the locations are made. In order to prevent features from being too close together, features are compared in pairs. If a pair of features is closer than 23 Mm, the one with the smaller relative maximum is rejected. As it is desired to use center–annulus separations $\Delta \leq 25°$, locations close to the edge of the maps are not used. On average, 930 features are located for each datacube with a total of 59 549 for the 64 twelve-hour datasets. Overlays of cell-center locations with the divergence maps are shown in Duvall and Birch (2010).

3.2. Filtering

Ray theory has proven invaluable in the development of time–distance helioseismology (Duvall *et al.*, 1993). It led to the basic idea that signals propagating along ray paths between surface locations would lead to correlations between the temporal signals from which travel times could be inferred. As acoustic, or p, modes with the same horizontal phase speed [ω/k: ω angular frequency and k horizontal wavenumber] travel along the same ray path to first order, it was natural to filter the data in ω/k to isolate waves traveling to a certain depth. This type of filter was introduced by Duvall *et al.* (1993) with subsequent development by Kosovichev and Duvall (1997) and Giles (2000). In the spatio-temporal power spectrum of solar oscillations (k–ω diagram), a line from the origin has a constant phase speed ω/k. A range of phase speeds thus corresponds to a pie-shaped wedge. Phase-speed filtering then corresponds to multiplying the Fourier transform of the input datacube by a pie-shaped wedge (appropriately modified for 3D) with some Gaussian tapering in phase speed (Couvidat and Birch, 2006).

The overriding principle of the use of the phase-speed filter is that solar perturbations lead to refractive time changes for signals traveling along a ray path. That this model is deficient was shown nicely in the work of Couvidat and Birch (2006) in which an artificial signal was detected with different amplitude depending on the width [δv] of the pie-shaped wedge. They concluded that the Born-approximation kernels (Birch and Kosovichev, 2000), which are derived based on a single scattering from solar perturbations, are more appropriate. A better way to think of the problem is to consider waves impinging on a perturbation that are subsequently scattered.

An initial analysis was performed using some nominal phase-speed filters and distance ranges detailed in Table 2. Thirteen distance ranges (and hence thirteen different filters) were applied to the six input datacubes for the dates 9 – 11 July 2010. The overall distance range covered for this test is $3.0 - 22.2°$. Center–annulus travel-time differences [δt_{oi}] were obtained for the thirteen distance ranges for each of the six datacubes using a standard surface-focusing time–distance analysis with the Gizon–Birch method of extracting travel times from the correlations (Gizon and Birch, 2004). The travel times are averaged over

Table 2 Nominal phase speed filter parameters. The first column is an identification number. The second column is the range in distance [Δ]. The third column is the mean distance [$\Delta_{\rm mean}$]. The fourth column is the central phase speed. The fifth column is the full width at half maximum (FWHM) of the Gaussian phase speed filter. For the tests in Section 3.2, filters 1 – 13 were used while for the final results in Section 3.3, filters 1 – 14 were used.

Filter	Δ [°]	$\Delta_{\rm mean}$ [°]	v_0 [km s^{-1}]	δv [km s^{-1}]
1	1.56 – 2.28	1.92	32.0	13.2
2	2.28 – 3.00	2.64	41.8	4.2
3	3.00 – 3.96	3.48	46.9	6.0
4	3.96 – 4.92	4.44	54.7	9.5
5	4.92 – 6.12	5.52	64.5	10.1
6	6.12 – 7.80	6.96	75.1	11.1
7	7.56 – 9.24	8.40	84.3	10.4
8	9.12 – 10.7	9.84	93.5	9.4
9	10.92 – 12.60	11.76	104.4	9.7
10	12.36 – 14.52	13.44	114.2	12.5
11	14.28 – 16.44	15.36	125.2	12.3
12	17.16 – 19.32	18.24	141.9	12.6
13	19.08 – 22.20	20.64	156.0	18.5
14	22.08 – 24.00	23.04	170.4	11.8

the distance range and about the locations of the supergranular centers and the signal at the average cell center location is extracted. Uncertainties are estimated from the scatter of the six values. The travel-time differences are plotted *versus* Δ in Figure 3 (black). The variation with Δ does not agree with the models in Figure 1, but peaks near $\Delta = 5°$ and decays approximately to zero near $\Delta = 20°$.

There was some concern that this distance dependence might have something to do with the filtering, so a case was analyzed with no phase-speed filtering. The f-mode was excluded and there was a frequency bandpass filter transmitting $\nu = 1.5 – 6$ mHz. The results of that analysis are also shown in Figure 3 (green). This shows a behavior with Δ much closer to the models in Figure 1 but with considerably more noise and a much larger signal close to five seconds. It was concluded that the phase-speed filters were significantly degrading the signal. To test this further, the filters were broadened by the factors and two, three, four, and five with the results of factors two (blue) and three (red) also shown in Figure 3. More signal is obtained with the larger-width filters, but even with the factor of five, only about 50 % of the signal is obtained and there is still a decay of the signal towards larger Δ.

Because of the approximate constancy of the simulation travel times [$\delta t_{\rm oi}$] with Δ in Section 2.3 and the approximate constancy with Δ for the no phase-speed filter case, it was concluded that the decay of $\delta t_{\rm oi}$ toward larger Δ in the phase-speed filtered results is an artifact of the filtering. Filtering would seem to be desirable because of the large difference in errors for the unfiltered and filtered cases. But a filter is needed that gives a more unbiased result at the large Δ in order that the results not be overly dependent on the modeling. A first attempt at this was to use a set of filters with the full width at half maximum (FWHM) equal to half the phase speed (Birch *et al.*, 2006). The results of these measurements are also shown in Figure 3 (orange). The dependence on Δ is somewhat reduced but there is still some decay at the largest Δ.

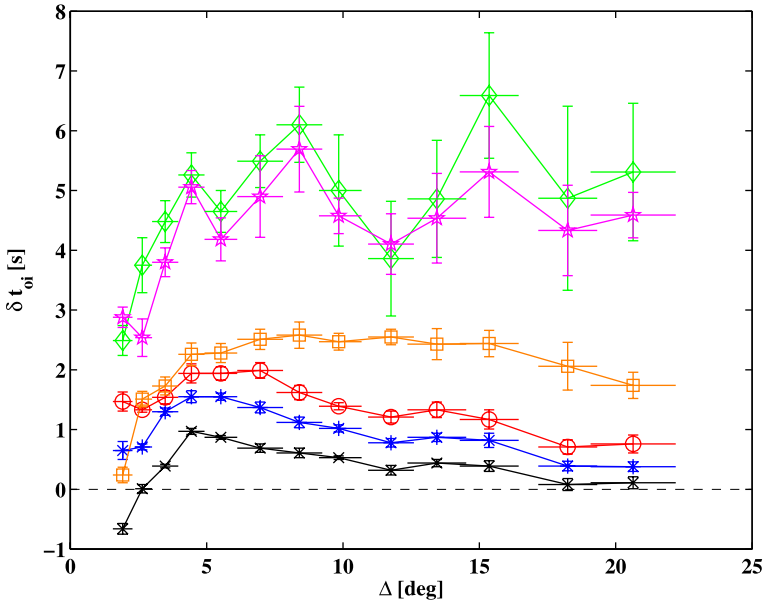


Figure 3 Center–annulus travel-time differences [$\delta t_{\rm oi}$] averaged about the supergranule centers for different filters for the thirteen Δ ranges for three days of HMI data. From the bottom to top (on the right side), the black (cross) points are for the nominal phase-speed filters. The blue (asterisk) points are for the phase-speed filters with the widths doubled. The red (circle) points are for the width tripled. The orange (square) points are for a filter width that is half the nominal phase speed. The magenta (star) curve is for the constant degree-width-filter (Section 3.3) with width $\Gamma_\ell = 400$. The green (diamond) curve is for no phase-speed filter. In all cases the f-mode is excluded via filtering as is signal outside the frequency bandpass $1.5 < \nu < 6$ mHz. These results were obtained using the three days of data 9 – 11 July 2010.

A different way to look at this problem is to consider a p-mode with frequency ν and spherical harmonic degree ℓ impinging on a supergranule. To the extent that the supergranule is static, the mode frequency ν will be maintained and the scattering will spread power in the spectrum away from the nominal value of ℓ. The scattering should spread power over a width in ℓ that depends on the spectrum of supergranulation, which peaks near $\ell = 120$. For similar values of the phase speed (which can also be characterized by ν/ℓ), the spread power should be over a similar width in ℓ (Chou *et al.*, 1996; Woodard, 2002). A filter that is centered at a certain phase speed but whose width is constant with ν and characterized by the FWHM in ℓ ($\equiv \Gamma_\ell$) may be what is needed. A value of $\Gamma_\ell = 400$ should capture most of the supergranular signal. Such a filter was implemented and applied to the thirteen distance ranges used. The results are also shown in Figure 3. The approximate constancy with $\Delta > 5^\circ$ and the larger values than for the normal phase-speed filters suggests that this type of filtering is a useful concept. The filter used has a flat top of width $[\Gamma_\ell/2]$ and is tapered to zero by a cosine bell that goes from 1 to 0 in $\Gamma_\ell/2$.

Some additional tests of this new filter concept were computed. In Figure 4(a) is shown the peak supergranular signal for $\Delta = 19.08 - 22.2^\circ$ as a function of the filter width Γ_ℓ. For small Γ_ℓ, much of the supergranular signal is removed, but as the width is increased more and more of the signal is transmitted. There is clearly a variation in the size of the error bars across Figure 4(a). Another thing to examine is the ratio of the signal in Figure 4(a) to the size of the error bar, which should give a value of the signal-to-noise ratio (S/N) of the filter (Couvidat and Birch, 2006). This is plotted in Figure 4(b). A rather broad peak is

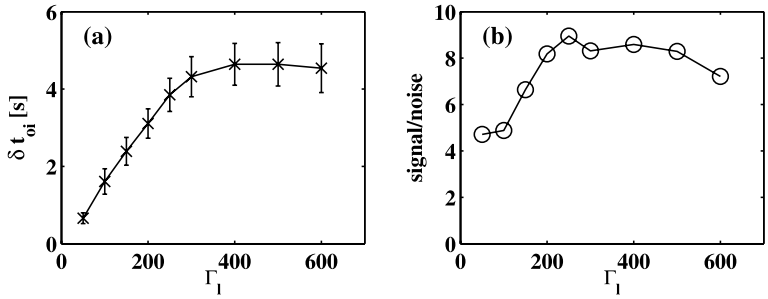


Figure 4 (a) Travel-time difference $[\delta t_{\rm oi}]$ *versus* filter FWHM $[\Gamma_\ell]$. The unfiltered case has 5.3 ± 1.2 seconds. (b) The travel-time difference from (a) divided by the size of the error bar from (a) *versus* the filter FWHM Γ_ℓ. The value for the unfiltered case is 4.6.

seen more or less centered on $\Gamma_\ell = 400$. One noticeable aspect of Figure 4(b) is that the unfiltered case gives a worse S/N (4.6) than the filters near the peak. Also narrow filters have reduced S/N. Based on the flatness of the $\Gamma_\ell = 400$ results with Δ in Figure 3, the extraction of essentially all of the supergranular signal in Figure 4(a), and the value of the S/N in Figure 4(b), $\Gamma_\ell = 400$ has been chosen for further analysis.

3.3. $\Gamma_\ell = 400$ Analysis

The 64 twelve-hour datacubes were analyzed with the fourteen phase-speed filters with $\Gamma_\ell = 400$. Center–annulus cross correlations were computed for the various distance ranges of Table 1. For checking purposes, it would be useful to be able to compute travel times using both the Gabor-wavelet method (Duvall *et al.*, 1997) and the Gizon–Birch method (Gizon and Birch, 2004). However, for the large distances in the present study ($\Delta < 24^\circ$), individual twelve-hour cross correlations have insufficient signal-to-noise ratio to enable the Gabor-wavelet fitting. It was decided to average the temporal correlations spatially about the supergranular centers. The averaging of the 930 (on average) correlations yields a high signal-to-noise ratio correlation that is amenable to the Gabor-wavelet fitting at the largest distances $[\Delta]$. Averaging the resultant $\delta t_{\rm oi}$ for the 64 twelve-hour datacubes enables the use of the envelope-time differences $[t_{\rm env}]$ as well as the normal phase times $[t_{\rm ph}]$. Although most theoretical work applies to phase-time differences, it will be useful to examine the envelope-time differences as they may have independent information.

It was found previously that there is an offset from zero (of unknown origin) of the center–annulus travel-time differences (Duvall and Gizon, 2000). For the small Δ of that study, the value of the offset was 0.16 ± 0.02 seconds. This offset is assumed unrelated to the flows to be measured and needs to be removed from computed travel times. To measure it for the present study, the average maps about the supergranule centers for the 64 12-hour datacubes are computed for the different distance ranges. An example is shown for the largest range of $\Delta = 22.08 - 24^\circ$ in Figure 5(a). An azimuthal average of Figure 5(a) is computed and the result is shown in Figure 5(b). At large radii $[r]$, the signal is approximately constant. This constant, which is assumed to be the offset that needs to be subtracted from the results, is plotted for the various Δ and travel-time methods in Figure 5(c). The Gizon–Birch and Gabor-wavelet phase times yield almost exactly the same results, which is expected for the quiet Sun and the use of the same time window for fitting (20 minutes). The travel-time differences corrected for this offset are plotted in Figure 5(d). Again, the Gizon–Birch and Gabor-wavelet phase times yield almost precisely the same results with an

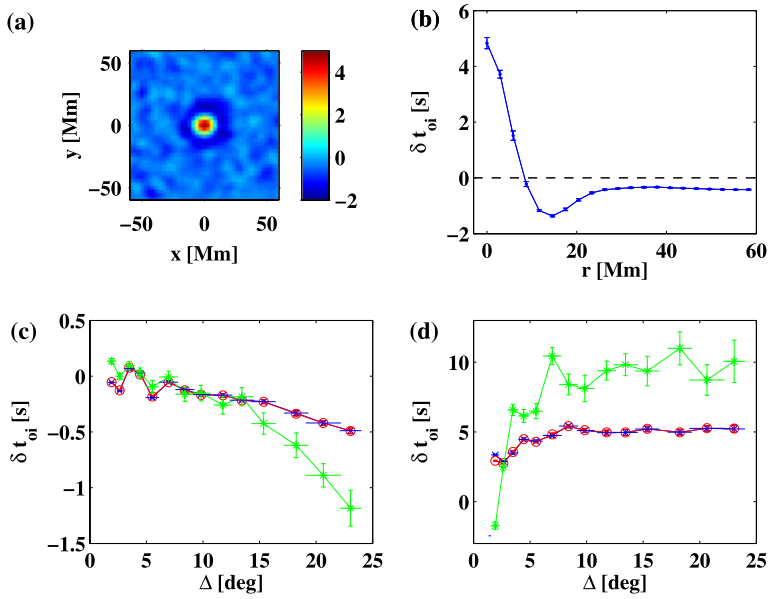


Figure 5 (a) Center–annulus travel-time difference averaged about supergranule centers for the range of $\Delta = 22.08 - 24°$. The scale of the colorbar at right is in seconds. Only the central point corresponding to the supergranular center is used in the present study. (b) Azimuthal average of (a). Note the offset at large radii. This offset is believed to be an artifact which needs to be removed from the results. (c) The offset at $r = 58$ Mm for the different travel-time definitions *versus* Δ. Blue is for the Gabor-wavelet phase time differences. Red is for the Gizon–Birch phase time differences and the green is for the Gabor-wavelet envelope time differences. (d) The resultant travel-time differences averaged for the 64 12-hour datacubes corrected for the offset in (c). The colors are the same as in (c).

average of 5.1 ± 0.1 seconds on the range $\Delta = 10 - 24°$. The Gabor-wavelet envelope times yield a considerably larger value of 9.7 ± 0.3 seconds on $\Delta = 10 - 24°$. In general the group velocity $[\partial\omega/\partial k]$ is about one half of the phase speed $[\omega/k]$ for p-modes, possibly leading to the larger travel time. The systematic error from Figure 5(c) is generally less the 10 % of the signal shown in Figure 5(d).

The Gabor-wavelet phase-time differences are plotted in Figure 6 *versus* the three ray models of Figure 1. The increase of the observed times for $\Delta = 2.5 - 8°$ may be the effect of the horizontal flow. The agreement is a little better for the $z_0 = -2.3$ Mm model.

4. Discussion

The main observational result of this article is that in the distance range $\Delta = 10 - 24°$ that the mean travel-time difference $[\delta t_{\rm oi}]$ at the center of the average supergranule is 5.1 ± 0.1 seconds. How secure is this result? There are no other results to compare with in this distance range. Also, the method of averaging over many supergranules is uncommon (Birch *et al.*, 2006; Duvall and Birch, 2010; Švanda *et al.*, 2011). The largest distance $[\Delta]$ used previously is 5° by Zhao and Kosovichev (2003). The phase-speed filter used in that study would have reduced the travel-time difference $[\delta t_{\rm oi}]$ by a factor of five and so it is difficult to compare. In addition, only inversions were presented and not raw travel-time differences.

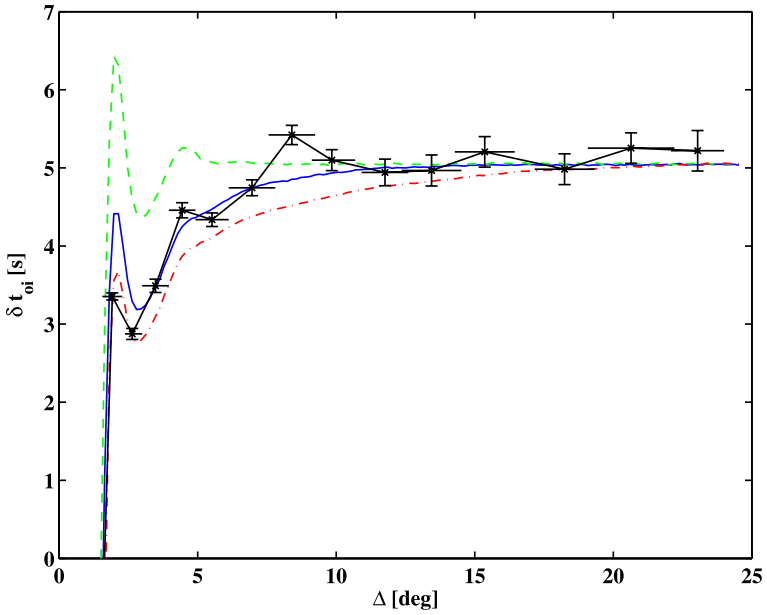


Figure 6 Comparison of the three ray models from Figure 1 with the HMI results (black with symbols and errors) of the $\Gamma_\ell = 400$ filtering and the Gabor-wavelet phase-speed time differences from Figure 5(d). These models are the sum of vertical and horizontal signals for the three Gaussian vertical flows peaking at $z_0 = -3.45$ Mm (red dot–dashed), $z_0 = -2.3$ Mm (blue solid), and $z_0 = -1.15$ Mm (green dashed) specified in Table 1.

One of the most interesting aspects of the present results is the large factor (> 20) between the photospheric vertical flow and the peak vertical flow. This increase of vertical flow then also requires an increase of the horizontal flow from the photosphere to the peak of a factor between 3.7 and 8.3 from the simple Gaussian models g1 – g3. Some simple tests suggest that this horizontal velocity should be detectible by a quadrant time–distance analysis, which will be left for future work. Some previous analyses have detected some larger flows below the surface than at the surface (Duvall *et al.*, 1997). Possibly the subsurface flow is highly variable which has made average properties difficult to ascertain from inversions of individual realizations.

These results are consistent with supergranulation as a shallow phenomenon, with the flow peaking within a few Mm of the photosphere. This would tend to support models in which supergranulation is related to a near-surface phenomenon such as granulation (Rast, 2003). The supergranular wave measurements (Gizon, Duvall, and Schou, 2003; Schou, 2003), and in particular the variation with latitude of the anisotropy of wave amplitude, would suggest a connection with rotation. The near-surface shear layer has been modeled as a possible exciter of supergranulation (Green and Kosovichev, 2006). Hathaway (1982) found that supergranulation as a convective phenomenon generates a near-surface shear layer. A difficulty with the near-surface shear layer as supergranular exciter is that recent simulations of Stein *et al.* (2009) develop a near-surface shear layer but no hint of excess power at supergranular scales is seen.

Acknowledgements The data used here are courtesy of NASA/SDO and the HMI Science Team. We thank the HMI team members for their hard work. This work is supported by NASA SDO and the NASA

SDO Science Center program through grant SCEX22011D awarded to NASA GSFC. S.M.H. acknowledges funding from NASA grant NNX11AB63G.

References

Birch, A., Duvall, T.L. Jr., Gizon, L., Jackiewicz, J.: 2006, *Bull. Am. Astron. Soc.* **38**, 224.

Birch, A.C., Gizon, L.: 2007, *Astron. Nachr.* **328**, 228.

Birch, A.C., Kosovichev, A.G.: 2000, *Solar Phys.* **192**, 193. ADS: 2000SoPh..192..193B. doi:10.1023/A:1005283526062.

Bogdan, T.J.: 1997, *Astrophys. J.* **477**, 475.

Chou, D.-Y., Chou, H.-Y., Hsieh, Y.-C., Chen, C.-K.: 1996, *Astrophys. J.* **459**, 792.

Couvidat, S., Birch, A.C.: 2006, *Solar Phys.* **237**, 229. ADS: 2006SoPh..237..229C. doi:10.1007/s11207-006-0209-z.

Duvall, T.L. Jr., Jefferies, S.M., Harvey, J.W., Pomerantz, M.A.: 1993, *Nature* **362**, 430.

Duvall, T.L. Jr., Kosovichev, A.G., Scherrer, P.H., Bogart, R.S., Bush, R.I., de Forest, C., Hoeksema, J.T., Schou, J., Saba, J.L.R., Tarbell, T.D., Title, A.M., Wolfson, C.J., Milford, P.N.: 1997, *Solar Phys.* **170**, 63. ADS: 1997SoPh..170...63D. doi:10.1023/A:1004907220393.

Duvall, T.L. Jr., Birch, A.C.: 2010, *Astrophys. J. Lett.* **725**, L47.

Duvall, T.L. Jr., Gizon, L.: 2000, *Solar Phys.* **192**, 177. ADS: 2000SoPh..192..177D. doi:10.1023/A:1005239503637.

Giles, P.M.: 2000, Ph.D. Thesis, Stanford Univ.

Gizon, L., Birch, A.C.: 2004, *Astrophys. J.* **614**, 472.

Gizon, L., Duvall, T.L. Jr., Schou, J.: 2003, *Nature* **421**, 43.

Gizon, L., Birch, A.C., Spruit, H.C.: 2010, *Ann. Rev. Astron. Astrophys.* **48**, 289.

Green, C.A., Kosovichev, A.G.: 2006, *Astrophys. J.* **641**, L77.

Hanasoge, S.M., Duvall, T.L. Jr.: 2007, *Astron. Nachr.* **328**, 319.

Hanasoge, S.M., Duvall, T.L. Jr., Couvidat, S.: 2007, *Astrophys. J.* **664**, 1234.

Hanasoge, S.M., Birch, A.C., Gizon, L., Tromp, J.: 2012, *Phys. Rev. Lett.*, submitted.

Hart, A.B.: 1954, *Mon. Not. Roy. Astron. Soc.* **114**, 17.

Hathaway, D.H.: 1982, *Solar Phys.* **77**, 341. ADS: 1982SoPh...77..341H. doi:10.1007/BF00156116.

Jackiewicz, J., Gizon, L., Birch, A.C.: 2008, *Solar Phys.* **251**, 381. ADS: 2008SoPh..251..381J. doi:10.1007/s11207-008-9158-z.

Kosovichev, A.G., Duvall, T.L. Jr.: 1997, In: Pijpers, F.P., Christensen-Dalsgaard, J., Rosenthal, C.S. (eds.) *SCORe'96: Solar Convection and Oscillations and Their Relationship* **225**, Kluwer, Dordrecht, 241.

Leighton, R.B., Noyes, R.W., Simon, G.W.: 1962, *Astrophys. J.* **135**, 474.

Rast, M.P.: 2003, *Astrophys. J.* **597**, 1200.

Rieutord, M., Rincon, F.: 2010, *Living Rev. Solar Phys.* **7**, 2. solarphysics.livingreviews.org/Articles/lrsp-2010-2/.

Scherrer, P.H., Bogart, R.S., Bush, R.I., Hoeksema, J.T., Kosovichev, A.G., Schou, J., Rosenberg, W., Springer, L., Tarbell, T.D., Title, A., Wolfson, C.J., Zayer, I., MDI Engineering Team: 1995, *Solar Phys.* **162**, 129. ADS: 1995SoPh..162..129S. doi:10.1007/BF00733429.

Schou, J.: 2003, *Astrophys. J. Lett.* **596**, L259.

Schou, J., Scherrer, P.H., Bush, R.I., Wachter, R., Couvidat, S., Rabello-Soares, M.C., Bogart, R.S., Hoeksema, J.T., Liu, Y., Duvall, T.L. Jr., Akin, D.J., Allard, B.A., Miles, J.W., Rairden, R., Shine, R.A., Tarbell, T.D., Title, A.M., Wolfson, C.J., Elmore, D.F., Norton, A.A., Tomczyk, S.: 2012, *Solar Phys.* **275**, 229. ADS: 2012SoPh..275..229S. doi:10.1007/s11207-011-9842-2.

Stein, R.F., Georgobiani, D., Schafenberger, W., Nordlund, Å., Benson, D.: 2009, *AIP* **CS-1094**, 764.

Švanda, M., Gizon, L., Hanasoge, S.M., Ustyugov, S.D.: 2011, *Astron. Astrophys.* **530**, A148.

Werne, J., Birch, A., Julien, K.: 2004, In: Danesy, D. (ed.) *SOHO 14 Helio- and Asteroseismology: Towards a Golden Future* **559**, ESA, Noordwijk, 172.

Woodard, M.F.: 2002, *Astrophys. J.* **565**, 634.

Zhao, J., Kosovichev, A.G.: 2003, In: Sawaya-Lacoste, H. (ed.) *GONG+ 2002. Local and Global Helioseismology: The Present and Future* **SP-517**, ESA, Noordwijk, 417.

DOI 10.1007/978-1-4899-8005-2_6
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0073-y

Subsurface Meridional Flow from HMI Using the Ring-Diagram Pipeline

R. Komm · I. González Hernández · F. Hill · R. Bogart · M.C. Rabello-Soares · D. Haber

Received: 29 February 2012 / Accepted: 9 July 2012 / Published online: 31 July 2012

Abstract We have determined the meridional flows in subsurface layers for 18 Carrington rotations (CR 2097 to 2114) analyzing high-resolution Dopplergrams obtained with the *Helioseismic and Magnetic Imager* (HMI) instrument onboard the *Solar Dynamics Observatory* (SDO). We are especially interested in flows at high latitudes up to 75° in order to address the question whether the meridional flow remains poleward or reverses direction (so-called counter cells). The flows have been determined in depth from near-surface layers to about 16 Mm using the HMI ring-diagram pipeline. The measured meridional flows show systematic effects, such as a variation with the B_0-angle and a variation with central meridian distance (CMD). These variations have been taken into account to lead to more reliable flow estimates at high latitudes. The corrected average meridional flow is poleward at most depths and latitudes with a maximum amplitude of about 20 $\mathrm{m\,s^{-1}}$ near 37.5° latitude. The flows are more poleward on the equatorward side of the mean latitude of magnetic activity at 22° and less poleward on the poleward side, which can be interpreted as convergent flows near the mean latitude of activity. The corrected meridional flow is poleward at all depths within $\pm 67.5°$ latitude. The corrected flow is equatorward only at 75° latitude in the southern hemisphere at depths between about 4 and 8 Mm and at 75° latitude in the northern hemisphere only when the B_0 angle is barely large enough to measure flows at this latitude. These counter cells are most likely the remains of an insufficiently corrected B_0-angle variation and not of solar origin. Flow measurements and B_0-angle corrections are difficult at the

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

R. Komm (✉) · I. González Hernández · F. Hill
National Solar Observatory, Tucson, AZ 85719, USA
e-mail: komm@nso.edu

R. Bogart · M.C. Rabello-Soares
HEPL Stanford University, Stanford, CA, USA

D. Haber
JILA University of Colorado, Boulder, CO, USA

highest latitude because these flows are only determined during limited periods when the B_0 angle is sufficiently large.

Keywords Helioseismology, observations · Velocity fields, interior

1. Introduction

We study the average meridional flow in subsurface layers of the Sun using high-resolution Dopplergrams obtained with the *Helioseismic and Magnetic Imager* (HMI) instrument onboard the *Solar Dynamics Observatory* (SDO) spacecraft. The average meridional flow is mainly poleward in each hemisphere with an amplitude of about 20 $\mathrm{m\,s^{-1}}$ at the surface (Bogart, 1987; Howard, 1996; Gizon, Birch, and Spruit, 2010, for reviews). The meridional flow is thus two orders of magnitude smaller than the average solar rotation rate. For this reason, the reliable detection of flows in the north-south direction on the Sun has proved to be a considerable observational challenge especially at high latitudes. The importance of meridional circulation in dynamo theories has been emphasized for many years (Durney, 1974, for example). In flux-transport dynamo models, the structure and the strength of the meridional flow determines the duration of a solar cycle, its rise and fall pattern, and the timing of the reversals of the Sun's polar fields (Dikpati and Charbonneau, 1999; Dikpati, 2004; Dikpati *et al.*, 2010). The meridional flow plays also a role in the evolution of polar magnetic fields (Dikpati, 2011). The poleward meridional flow in the near-surface layers might be linked to the inward gradient of the rotation rate via gyroscopic pumping (Miesch and Hindman, 2011). To improve the understanding of the relationship between dynamo and meridional flow, major topics that need to be addressed are the variation of the meridional flow with the solar cycle, the variation with depth, and the meridional flow at high latitudes. Here, we focus on studying the meridional flow at high latitudes in the upper 16 Mm of the solar convection zone.

The meridional flow in subsurface layers of the Sun have been studied previously with local-helioseismic techniques applied to Dopplergrams obtained with the ground-based *Global Oscillation Network Group* (GONG) and the *Michelson Doppler Imager* (MDI) onboard the *Solar and Heliospheric Observatory* (SOHO). The meridional flow has been studied in the outer few percent of the solar radius analyzing MDI and GONG data with the ring-diagram or the time–distance technique (Giles *et al.*, 1997; Haber *et al.*, 2002; Zhao and Kosovichev, 2004; Gizon and Rempel, 2008; González Hernández *et al.*, 2008; Basu and Antia, 2010). During the solar cycle, the flow amplitude varies by about 5 $\mathrm{m\,s^{-1}}$ near the mean latitude of magnetic activity with large amplitudes during cycle minimum and small ones during maximum.

However, previous studies using local-helioseismic techniques have been limited to a range of about $\pm 60^\circ$ latitude using GONG or MDI Dopplergrams. The spatial resolution of the Dopplergrams combined with the effect of geometric foreshortening has limited the observations to this latitude range. At latitudes of 45° and higher, the measured meridional flow in subsurface layers can change direction from poleward to equatorward during some epochs (Haber *et al.*, 2002). Since the appearance of these counter cells correlates with the annual variation of the B_0 angle, it is generally assumed that their existence is an artifact that needs to be corrected (Zaatri *et al.*, 2006; González Hernández *et al.*, 2006). However, there are theoretical studies that investigate the meridional flow at high latitudes and explore the possible existence of multiple cells in latitude (Dikpati and Gilman, 2012). Mt. Wilson Observatory (MWO) Dopplergrams have been used to study the meridional flow up to 80°

latitude at the solar surface and a reversed circulation pattern is seen in polar regions for three successive solar minima (Ulrich, 2010). A feature-tracking analysis of SOHO observations finds that the meridional flow derived from magnetic features is poleward up to 75° latitude (Hathaway and Rightmire, 2010), while a cross-correlation analysis of MWO magnetograms finds a weakly equatorward meridional flow for latitudes above 60° (Snodgrass and Dailey, 1996). This illustrates how difficult it is to observe the meridional flow at high latitudes even at the solar surface. The high-resolution Dopplergrams obtained with HMI allow us to determine the meridional flow in subsurface layers at high latitudes comparable to those derived from surface observations; the HMI ring pipeline produces flows measured on ring-diagram tiles centered at up to 75° latitude.

Here, we determine the meridional flow averaged over 18 Carrington rotations using the HMI ring-diagram pipeline. We estimate systematic variations, such as a center-to-limb variation of the flows and the more important B_0-angle variation and correct the measured flows accordingly. We find that the corrected meridional flow is poleward at all latitudes and depths with few exceptions. The measured as well as the corrected meridional flows are clearly poleward at all latitudes within the latitude range that has been previously studied with GONG and MDI observations.

2. Data and Analysis

2.1. Ring-Diagram Analysis

We analyze observations obtained during 18 Carrington Rotations (2097 – 2114: 2 May 2010 – 21 September 2011) for which we have high-resolution full-disk Doppler data obtained with the *Helioseismic and Magnetic Imager* (HMI: Schou *et al.*, 2012) instrument onboard the *Solar Dynamics Observatory* (SDO: Pesnell, Thompson, and Chamberlin, 2012) spacecraft. We determine the horizontal components of solar subsurface flows with a ring-diagram analysis using the dense-pack technique (Haber *et al.*, 2002) adapted to HMI data and processed by the HMI ring-diagram pipeline (Bogart *et al.*, 2011a, 2011b). Full documentation on the pipeline analysis modules and associated data products can be found on the web pages of the HMI Ring Diagrams Team (http://hmi.stanford.edu/teams/rings/). Pipeline results are available through the Joint Science Operations Center or JSOC (http://jsoc.stanford.edu).

The full-disk Doppler images are divided into 284 or 281 overlapping tiles with an effective diameter of 15° after apodization. The number and distribution of tiles varies with the B_0 angle (Bogart *et al.*, 2011b) and the centers of the tiles range up to $\pm 75°$ in latitude and central meridian distance (CMD). The tile centers are spaced by 7.5° in latitude, while in CMD the centers are spaced by 7.5° up to a latitude of 60° but sparser at higher latitudes. Throughout this study, we refer to the centers of the ring-diagram tiles when mentioning latitude or CMD values. For example, a tile centered at 75° latitude includes data between 67.5° and 82.5° latitude. The Carrington rotation rate is used as tracking rate. The analysis is carried out on overlapping "days" of 1664-minute time periods and the shift between two days corresponds to exactly 15° of Carrington longitude. We derive daily maps of horizontal velocities at 16 depths from 0.6 to 16 Mm for each tile and combine these daily flow maps to calculate synoptic flow maps for each depth. To create a uniformly spaced grid, we interpolate linearly the values at high latitudes on a grid with centers spaced by 7.5° in CMD. When merging the daily flow maps, we include all tiles at locations within $\pm 60°$ CMD and the synoptic maps are heavily weighted toward the central meridian using a weighting factor

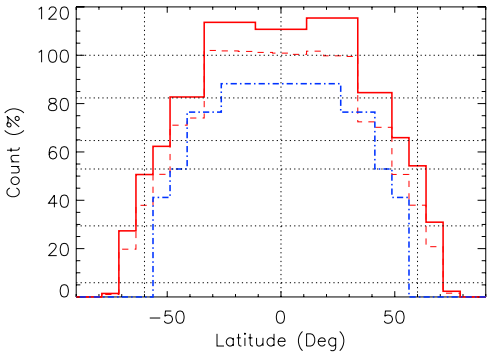


Figure 1 The percentage of tiles that contribute to the average meridional flow at each latitude (red solid line). For each Carrington rotation, there are 24 ring days and 17 positions in CMD covering $\pm 60°$ CMD or 408 data points. This number is used to normalize the curves. Values greater than 100 % indicate that data at greater CMD distance could be included, values less than 100 % indicate less coverage in CMD. The deepest layer of 15.8 Mm has slightly reduced coverage (red dashed line). For comparison, we include the normalized coverage of the dense-pack grid used for GONG and MDI data (blue dot-dashed line), which covers $\pm 52.5°$ in CMD of tile center positions.

of cosine CMD to the fourth power, as in the maps of magnetic activity created at NSO/Kitt Peak.

Figure 1 shows the number of tiles in CMD as a function of latitude that contribute to the meridional flow averaged over all Carrington rotations. For each Carrington rotation, we have 24 ring days and 17 positions in CMD covering $\pm 60°$ CMD or 408 data points. This number is used to normalize the curves. Values greater than one indicate that data at greater CMD distance could be included in the synoptic maps. The number of tiles decreases at high latitudes. However, even at 60° latitude, the measurement includes about half of all possible data points. When compared with MDI or GONG data, which are limited to $\pm 52.5°$ CMD and latitude, HMI data clearly provide better coverage at high latitudes. The coverage is slightly reduced at the greatest depth of 15.8 Mm because the flows are not determined at locations where the number of modes fitted is considered insufficient. This is different from the MDI or GONG ring-diagram pipeline. At 75° latitude, the flow is measured only at the position at central meridian for any ring day if the B_0 angle is greater than 3.625° (Bogart *et al.*, 2011b). Similarly, the maximum distance from disk center changes with latitude and B_0 angle. This accounts for the small differences in Figure 1 between equator and low latitudes within $\pm 30°$ latitude and the tiny north-south difference, since the average B_0 angle is slightly positive.

As a measure of solar activity, we use the NSO SOLIS synoptic maps (http://solis.nso.edu/solis_data.html). The SOLIS synoptic maps have been calibrated to ensure compatibility with Kitt Peak synoptic maps (Jones *et al.*, 2004; Henney and Team, 2007). We convert the magnetogram data to absolute values and bin them into circular areas with 15° diameter centered on a grid with 7.5° spacing in latitude and longitude to match the ring-diagram grid. In this way, we estimate synoptic maps of the unsigned magnetic flux density (in Gauss) on the same spatial scales as the subsurface flow maps. For each Carrington rotation, we calculate the mean latitude of activity in each hemisphere. The mean latitude of activity is $21.7° \pm 1.5°$ in the northern and $-22.8° \pm 0.8°$ in the southern hemisphere averaged over all 18 rotations. It change from 23° to 20° in the northern hemisphere and 23° to 22° in the southern one during the epoch covered.

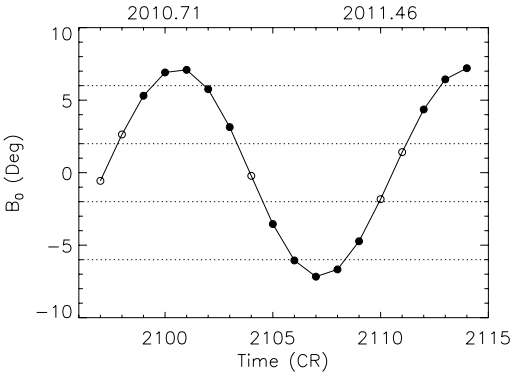


Figure 2 The variation of the B_0 angle during the period used in this study. The time coordinate is given in Carrington rotations (bottom x-axis) and calendar years (top x-axis). When the B_0 angle is large enough (filled circles), the flow can be measured at 75° latitude in the northern hemisphere (positive B_0 angle) or in the southern one (negative B_0 angle).

2.2. Variation with B_0 Angle and CMD

A variety of geometric effects are taken into account during the remapping stage that vary with a period of one year, such as the solar inclination toward the Earth (B_0 angle). However, the processing cannot correct for a loss in spatial resolution introduced by these effects. For example, a large B_0 angle value means that in one solar hemisphere a location at a given solar latitude is located closer to disk center, while the same location is located farther away from disk center in the other hemisphere, which will increase the geometric foreshortening and lower the effective spatial resolution. The effective distance from disk center is proportional to $\cos(\theta - B_0)$ where θ is the latitude. The annual variation of the effective resolution can artificially introduce annual variations in the measured subsurface flows (Zaatri *et al.*, 2006). With the high spatial resolution of the HMI Dopplergrams, we expect that such a B_0-angle induced variation is less pronounced in HMI data and will affect mainly flows at high latitudes. The subsurface flows can be measured up to $\pm 67.5°$ latitude at all times. Flows for any ring day can be determined at 75° latitude but only in one hemisphere or the other when the B_0 angle is 3.625° or larger. For the flow averaged over a synoptic map, we require that at least 25 % of the 24 daily flow maps lead to flows at 75° latitude.

In order to check for annual variations introduced by this effect, we calculate the B_0 angle for the middle of every Carrington rotation and compare its temporal variation with the variation of the measured meridional flows. Figure 2 shows the variation of the average B_0 angle during the 18 Carrington rotations used in this study. The B_0 angle is $+1.1°$ averaged over all 18 Carrington rotations. The maximum and minimum values are comparable to the size of a ring-diagram grid point in latitude; their difference is comparable to the diameter of a ring tile.

We determine the effect of the B_0-angle variation in the next section and correct the meridional flow. For this purpose, we assume that the meridional flow does not vary at a given latitude and that the best estimate of the flow can be determined at the maximum B_0 angle of 7.23° in the northern and $-7.23°$ in the southern hemisphere, which leads to the smallest distance from disk center and thus to the least amount of foreshortening. It has been suggested to use the distance from disk center instead of the B_0 angle to correct the meridional flow (Beckers, 2007). However, this makes no difference to the corrected flows, since the relationship is almost linear between $\cos(\theta - B_0)$ and the B_0 angle. At high latitudes, where the annual variation of the flow is large, the B_0 angle is small compared to the latitude itself, while at low latitudes, the B_0 angle is large compared to the latitude but the annual variation is small.

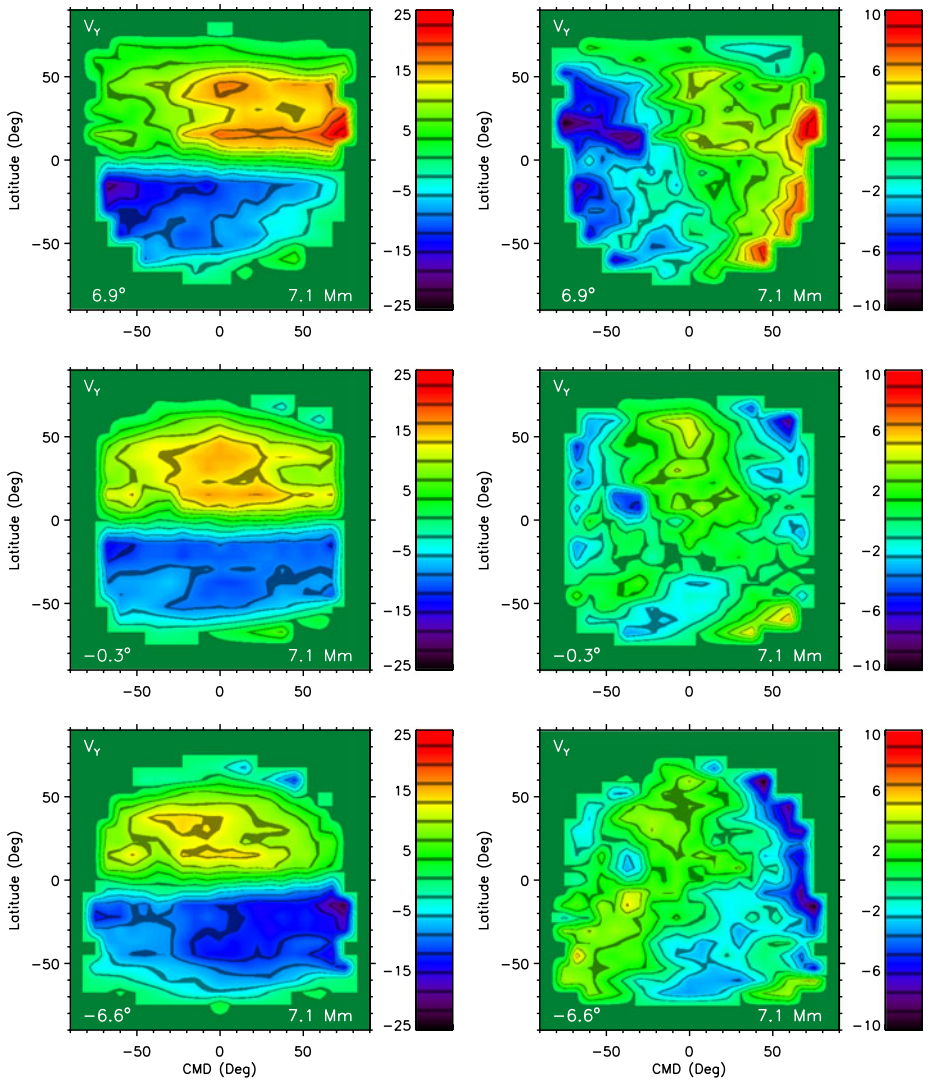


Figure 3 The meridional flow averaged over four Carrington rotations with large positive B_0 angles, four rotations with zero, and three rotations with large negative B_0 angles (top: $B_0 > 6°$; middle: $|B_0| \leq 2°$; bottom: $B_0 < -6°$) at a depth of 7.1 Mm (left) and the residual values after subtracting the mean at each latitude (right). The average B_0 angles are 6.9°, −0.3°, and −6.6° from top to bottom. Positive values indicate meridional flows to the north, negative values indicate flows to the south.

In previous studies (Komm, Howe, and Hill, 2009, 2011), it has been noticed that the meridional flow amplitude at a given latitude varies with disk position or CMD. This CMD variation is considered to be an artifact of unknown origin. We expect that such a variation with CMD has little influence on the meridional flow in a synoptic map simply because the flow values are heavily weighted toward the central meridian. In Figure 3, we show the meridional flow averaged over several Carrington rotations according to the size of the B_0 angle. The variation with disk position is mainly a CMD variation with a maximum at the

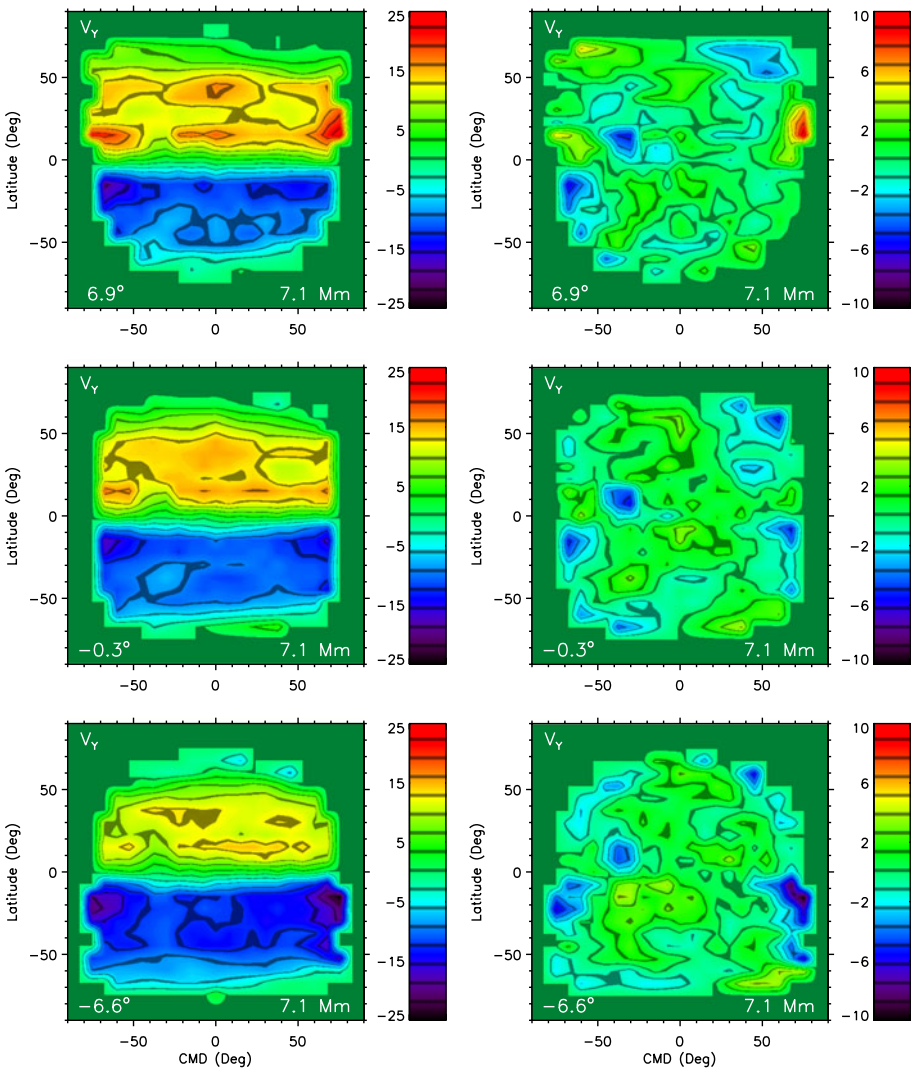


Figure 4 The meridional flow corrected for CMD variations averaged over four Carrington rotations with large positive B_0 angles, four rotations with zero, and three rotations with large negative B_0 angles (top: $B_0 > 6°$; middle: $|B_0| \leq 2°$; bottom: $B_0 < -6°$) at a depth of 7.1 Mm (left) and the residual values after subtracting the mean at each latitude (right). The average B_0 angles are 6.9°, −0.3°, and −6.6° from top to bottom. Positive values indicate meridional flows to the north, negative values indicate flows to the south. (Compare with Figure 3.)

central meridian. In addition, there is an east-west variation that changes sign with the B_0 angle. The meridional flows at other depths show a similar variation.

To compensate for the variation with CMD, we normalize the meridional flow at each latitude with their standard deviation calculated within ±60° CMD, average the normalized flow at each CMD position over ±60° latitude, and fit the average normalized flow with a two-term function representing the CMD variation and the east-west trend. We detrend the

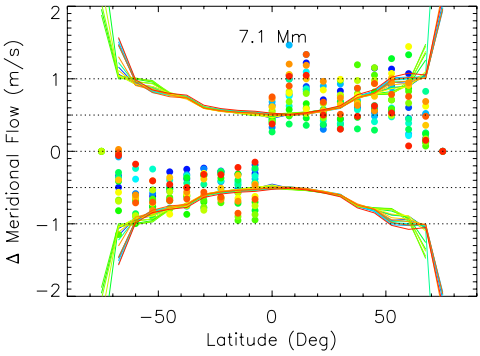


Figure 5 The difference between the meridional flow corrected for the CMD variation and the original values for each one of the 18 Carrington rotations used in this study (full circles). The average error bars of the original data are included for comparison (solid lines). The different colors represent the different Carrington rotations.

velocity data at a given depth by the corresponding fit as a function of CMD (ϕ) and assume that the value at central meridian is the "true" value. The corrected meridional flow is then calculated as

$$v_y^{\rm cor}(\phi,\theta) = v_y^{\rm obs}(\phi,\theta) + \left[a_c(1-\cos\phi) - a_s\sin\phi\right]\left\langle v_y^{\rm obs}(\theta)\right\rangle \tag{1}$$

where $v_y^{\rm cor}$ and $v_y^{\rm obs}$ represent the corrected and observed flow values, $\langle v_y^{\rm obs}\rangle$ is the standard deviation of the observed flow at a given latitude θ, while $\cos\phi$ and $\sin\phi$ represent the CMD variation and the east-west trend. Figure 4 shows the corrected meridional flow at the same depth as Figure 3. The variation with CMD is greatly diminished. This procedure is repeated for every depth and Carrington rotation.

Figure 5 shows the difference between the meridional flow corrected for the CMD variation and the original data for all Carrington rotations studied here. The amount of the correction is comparable to the errors of the original flows as a function of latitude. The same is true for the CMD correction at other depths. This confirms our expectation that this particular effect has little influence on the average meridional flows of synoptic maps. Nevertheless, we use corrected values in this study.

3. Results

Figure 6 shows the meridional flow averaged over all 18 Carrington rotations as a function of latitude. The average meridional flow is poleward in each hemisphere with the exception of small equatorward flows at some depths at 75° latitude. The average flow is well represented by a fit of the derivatives in latitude of the first two even Legendre polynomials which are orthogonal on the solar sphere with zero amplitude at the equator and the poles (Komm, Howard, and Harvey, 1993). The comparison with the Legendre fits shows that near 15° latitude the amplitude of the meridional flow is somewhat larger than the fit and smaller near 30° latitude. Compared with the mean latitude of activity, the average flow is slightly more poleward on its equatorward side and less poleward on its poleward side. This agrees with previous studies that have interpreted this as a convergent flow near the mean latitude of activity (Zhao and Kosovichev, 2004; Komm *et al.*, 2004) which to a large part reflects the presence of active regions (Howard, 1996; Švanda, Kosovichev, and Zhao, 2008). The figure includes also the flow values without correcting for the CMD variation. As expected from the discussion in Section 2.2, the correction for the amplitude variation with disk position does not make much of a difference.

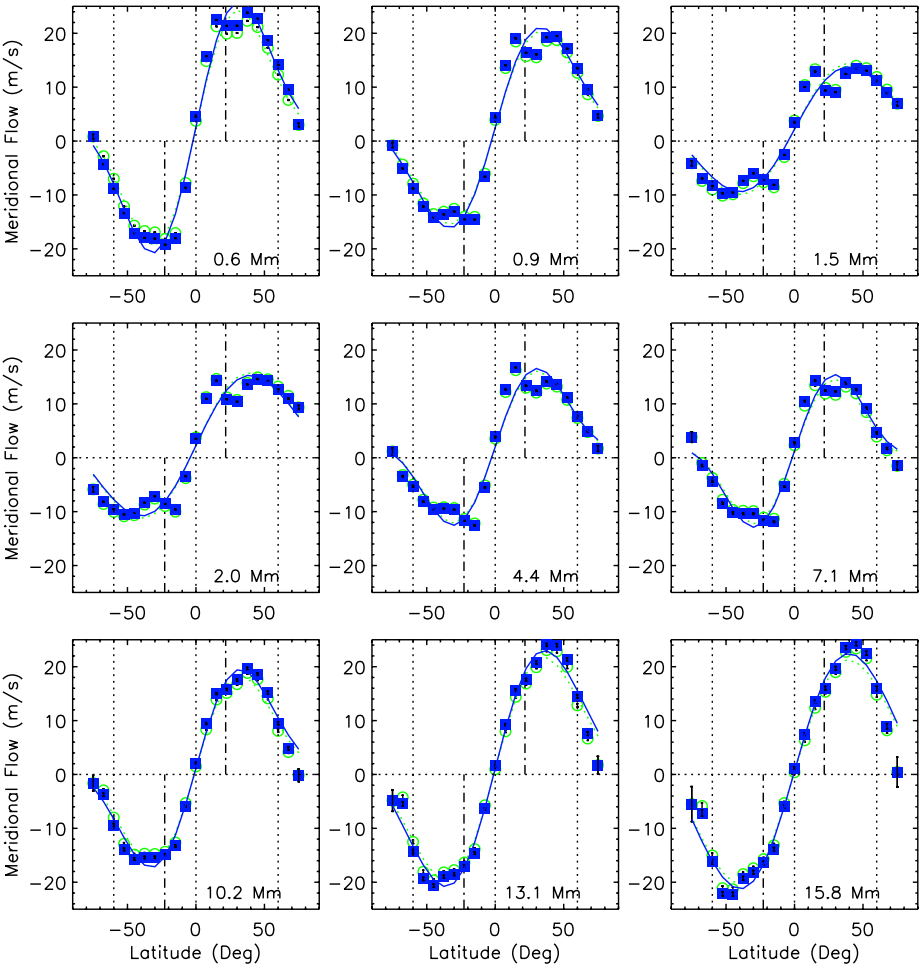


Figure 6 The meridional flow averaged over all 18 Carrington rotations for nine depths from near-surface (top-left) to deeper layers (bottom-right). The blue solid lines indicate a fit of the derivatives in latitude of the first two even Legendre polynomials. Positive values indicate meridional flows to the north, negative values indicate flows to the south. The vertical dot-dashed lines indicate the mean latitude of activity in each hemisphere. The values uncorrected for the CMD variation are included for comparison (green open circles). Small equatorward flows occur at some depths at 75° latitude in both hemispheres.

From previous studies (Zaatri *et al.*, 2006; González Hernández *et al.*, 2006), it is known that the meridional flow varies with the B_0 angle at high latitudes. Figure 7 shows the meridional flow at three latitudes as a function of time for two depths. Each data point represents the flow averaged over one Carrington rotation. At a depth of 7.1 Mm, the meridional flow at 45° and 60° latitude are clearly correlated with the B_0 angle in both hemispheres, while it is not so obvious at a depth of 11.6 Mm. The meridional flow at 75° latitude is only measured when the B_0 angle is sufficiently large and is determined in the northern hemisphere during eight Carrington rotations and during only five in the southern hemisphere.

Figure 8 shows that the relationship is nearly linear between meridional flow and B_0 angle for the northern (filled circles) and the southern hemisphere (open circles). The sign was

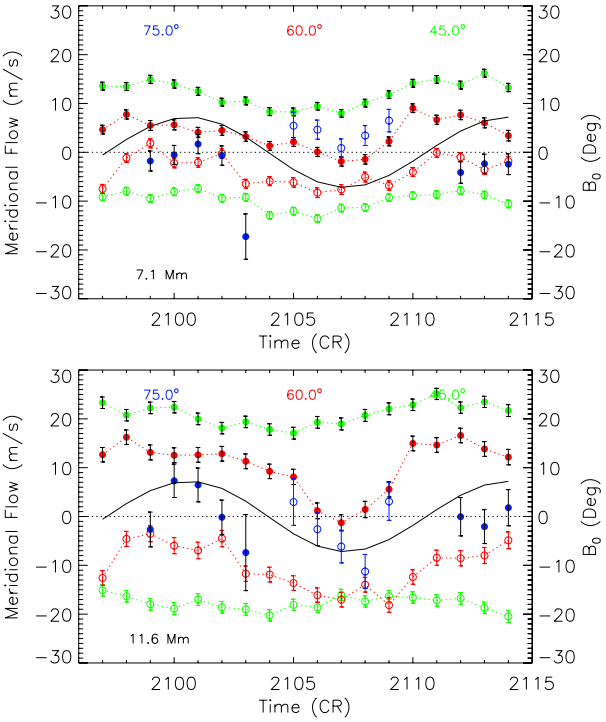


Figure 7 The temporal variation of the meridional flow at a depth of 7.1 Mm (top) and 11.6 Mm (bottom) for 45°, 60° and 75° latitude (filled circles: northern hemisphere; open circles: southern hemisphere). Each data point represents the flow values averaged over one Carrington rotation. Positive values indicate meridional flows to the north, negative values indicate flows to the south. The solid line indicates the B_0 angle in degree.

reversed for flow and B_0-angle values of the southern hemisphere ($-B_0, -v_y[\theta < 0]$). This means that positive flows indicate poleward meridional flows and negative values indicate equatorward meridional flows independent of hemisphere. A positive B_0 angle thus means in this figure that the distance from disk center is smaller at a given latitude compared to a zero B_0 angle and a negative value means a greater distance from disk center. A negative B_0 angle leads to a reduced effective spatial resolution. The influence of the B_0 angle is greatest in the deeper layers as indicated by the steeper slopes. At large negative B_0 angles, the relationship might not be simply linear at the deepest layers. At less than 45° latitude, the slopes are much closer to zero indicating that the meridional flow does not vary much with the B_0 angle at these latitudes. The slopes are positive in most cases; the flow values are large when the B_0 angle is large. This is expected from previous studies, where the B_0 angle introduced a counter cell. But at some depths and latitude combinations the slopes are negative, which implies that the effect of the B_0 angle is to artificially enhance the meridional flow.

Figure 9 shows the error of the meridional flow as a function of the B_0 angle. As expected, the error increases when the B_0-angle values are negative. As in Figure 8, a negative B_0 angle means that the hemisphere is tilted away from the observer and the effective spatial resolution is reduced. The variation with B_0 angle is small for error values at less than 60° latitude. The difference in the errors between $\pm 7.2°$ is generally less than 20 % at these latitudes. However, the error varies much more with the B_0 angle at the two most poleward latitudes. At 67.5° latitude, the error increases rapidly for $B_0 < -3°$ up to a factor of two in the northern hemisphere and similarly for $B_0 > 3°$ in the southern hemisphere. At these B_0-angle values, the error increases faster than quadratic; the relationship is clearly non-linear. At 75° latitude, the error increases rapidly for $B_0 < 5°$ up to a factor of three. When

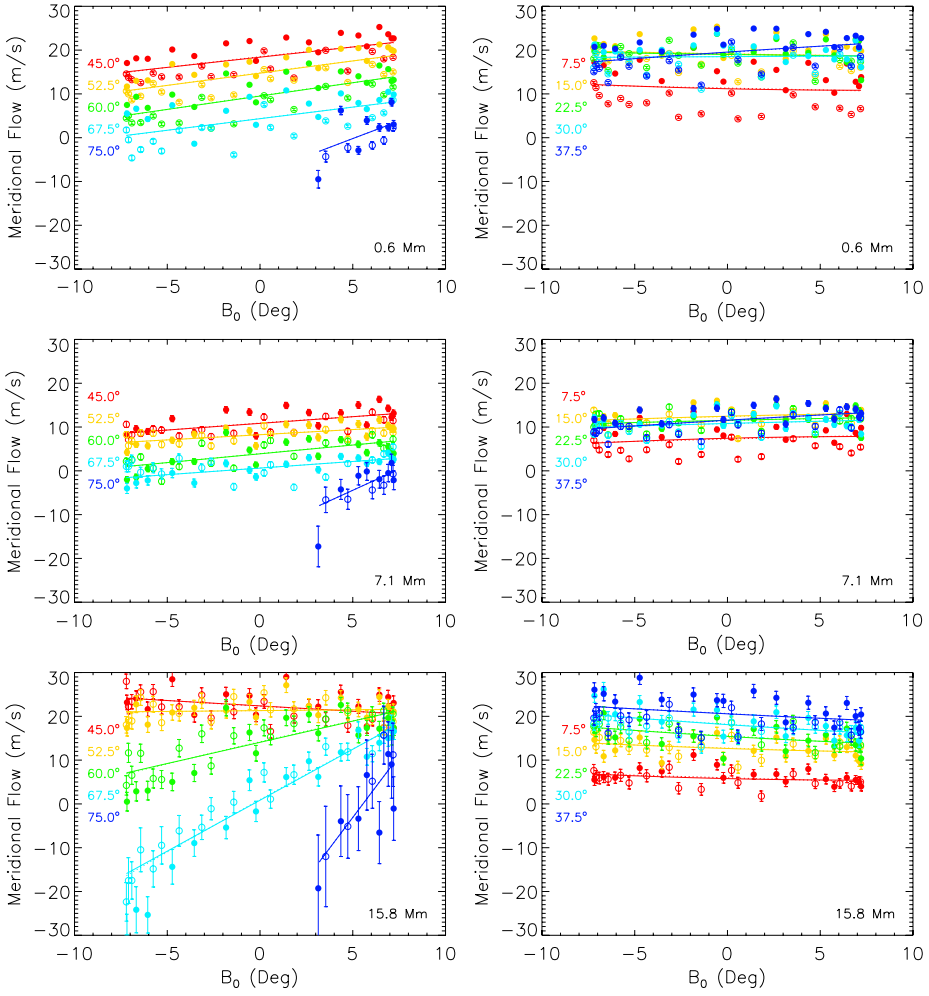


Figure 8 The average meridional flow as a function of B_0 angle for 18 Carrington rotations at latitudes of 45° and higher (left) and at lower latitudes (right) at three depths (top: 0.6 Mm; 2nd: 7.1 Mm; bottom: 15.8 Mm) for the northern (filled circles) and the southern hemisphere (open circles). The sign was reversed for flow and B_0-angle values of the southern hemisphere ($-B_0, -v_y[\theta < 0]$). Positive flows thus indicate poleward meridional flows and negative values indicate equatorward flows. The solid lines indicate a linear regression with the B_0 angle. The regression with $\cos(\theta - B_0)$ is included for comparison (dotted line).

compared with Figure 8, this is the range of B_0-angle values where the flow values decrease more rapidly than linear with decreasing B_0 angle. This might imply that there is a lower limit of B_0 angles below which the meridional flow cannot be reliably determined at the two highest latitudes (67.5° and 75°). When we have more data available, we will determine this relationship more accurately and derive a B_0-angle threshold for these latitudes.

We use the linear regression between B_0 angle and meridional flow at each depth and latitude, as shown in Figure 8, to correct the meridional flow values assuming that the values at the largest B_0 angle of 7.23° represent the best flow estimates and the values at $-7.23°$

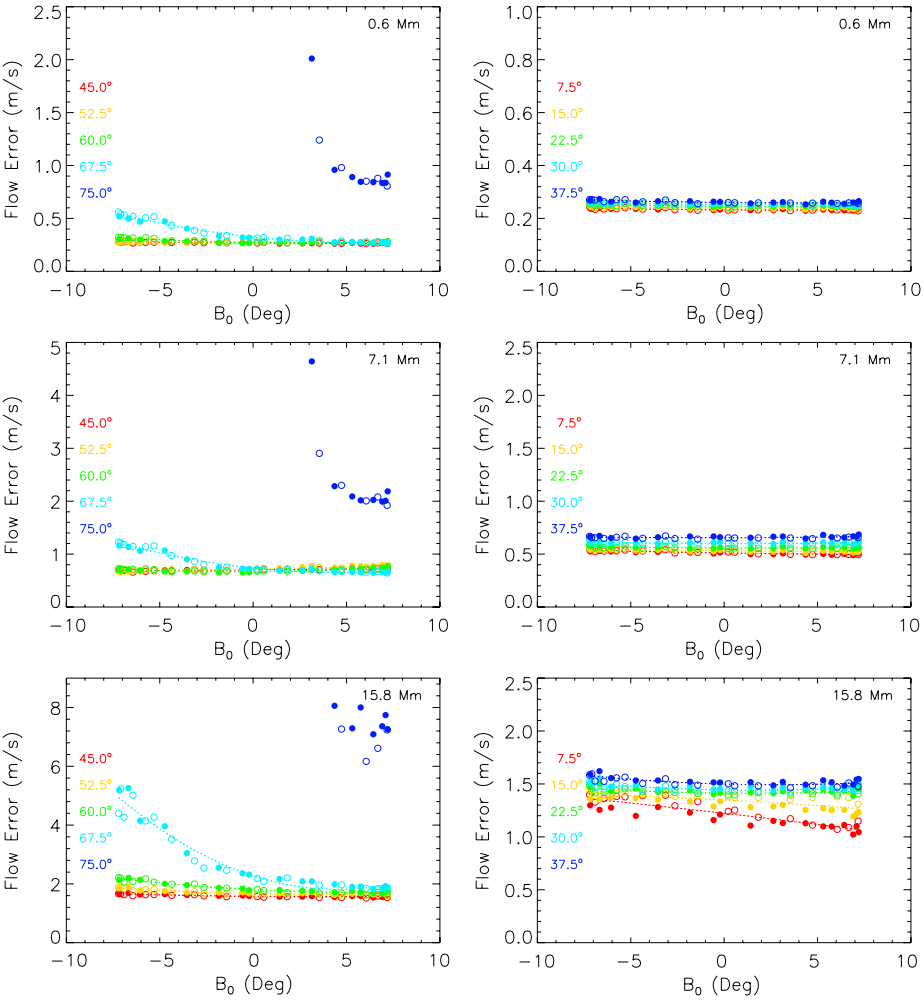


Figure 9 The average error of the meridional flow as a function of B_0 angle at different latitudes for 18 Carrington rotations at latitudes of 45° and higher (left) and at lower latitudes (right) at three depths (top: 0.6 Mm; 2nd: 7.1 Mm; bottom: 15.8 Mm) for latitudes of 45° and higher for the northern (filled circles) and the southern hemisphere (open circles). The sign was reversed for the B_0-angle values of the southern hemisphere ($-B_0, \sigma_y[\theta < 0]$).

represent the worst estimates. At each depth, the corrected meridional flow as a function of latitude, θ, and time, t, is then calculated as

$$v_y^{\rm cor}(\theta, t) = v_y^{\rm obs}(\theta, t) + \left[B_0^{\rm max} - B_0(t)\right] a_0(\theta). \tag{2}$$

Figure 10 shows the same as Figure 7 for the corrected values. The meridional flows at a given latitude still vary with time, but the annual variation is greatly reduced compared to the uncorrected values. The meridional flow at 60° latitude is now poleward at all times compared to the uncorrected values shown in Figure 7. The corrected meridional flow at 75° latitude is now predominantly poleward in the northern hemisphere at all depths and poleward in the southern hemisphere in deeper layers. In the southern hemisphere, the corrected

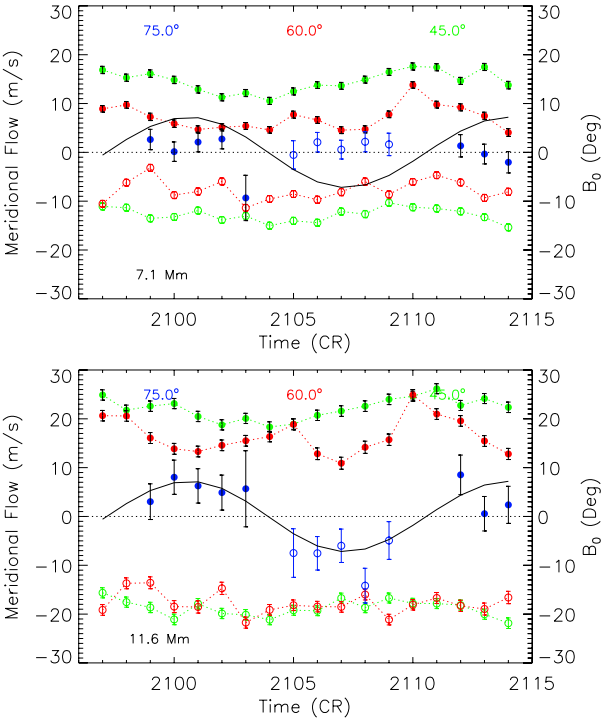


Figure 10 The temporal variation of the meridional flow at a depth of 7.1 Mm (top) and 11.6 Mm (bottom) for 45°, 60° and 75° latitudes corrected for the variation with the B_0 angle (filled circles: northern hemisphere; open circles: southern hemisphere). Each data point represents the flow averaged over one Carrington rotation. Positive values indicate meridional flows to the north, negative values indicate flows to the south. The solid line indicates the B_0-angle variation in degree. (Compare with Figure 7.)

meridional flow at 75° is less equatorward than before at a depth of 7.1 Mm, but the flow is still slightly equatorward.

Now, we can study the meridional flow averaged over all Carrington rotations corrected for the B_0-angle variation. Figures 11 and 12 show the average meridional flow as a function of latitude and depth. In Figure 11, the amplitudes of the corrected flows are slightly larger than the uncorrected ones with few exceptions. The corrected flows are poleward at all latitudes within $\pm 67.5°$ at all depths. Even at 75° latitude, the flows are poleward at most depths in both hemispheres. The meridional flows in the southern hemisphere at 4 and 7 Mm are the exception; at these depths there is a counter cell at $-75°$ latitude. For comparison, we include the extrapolation of the meridional flow to a B_0 angle of 7.23°, which is symmetric in hemisphere. This comparison shows that there is a small asymmetry with slightly larger meridional flow in the northern hemisphere than in the southern one. In this context, we like to note that the average B_0 angle is slightly positive; the northern hemisphere is slightly better resolved than the southern one.

We can speculate what the meridional flow might be like near the poles and extrapolate the measurements to 82.5° and 90° using a least-squares quadratic fit. If we assume that the meridional flow is zero at the poles, the flow gradually decreases in amplitude toward the poles without reversing direction except in the southern hemisphere between 4.4 and 7.1 Mm as discussed above. Without this constraint, the extrapolation leads frequently to large flows at the poles, which is unrealistic since the measured flows are averages over 18 Carrington rotations. In this unrealistic case, counter cells are common and the flows reverse at 78° latitude except near 2 Mm.

Figure 12 shows the corrected average meridional flow as a function of latitude at all depths. The two most noticeable items are that at depths between about 4 and 8 Mm the flow

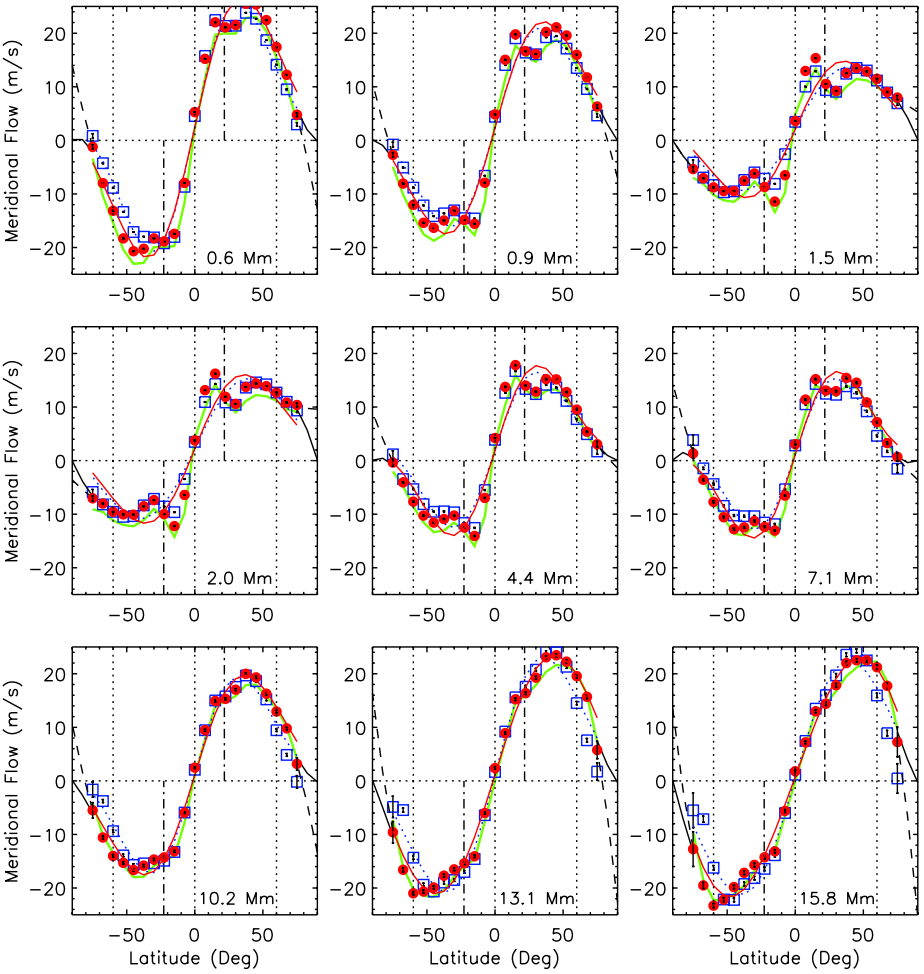


Figure 11 The meridional flow averaged over all 18 Carrington rotations for nine depths from near-surface (top-left) to deeper layers (bottom-right) corrected for the variation with the B_0 angle (red solid circles) and without correction (blue open squares). The solid red lines indicate a fit of the derivatives in latitude of the first two even Legendre polynomials. Positive values indicate meridional flows to the north, negative values indicate flows to the south. The B_0 angle is 1.1° on average. The thick green line indicates the flow extrapolated to a B_0 angle of 7.23°. The vertical dot-dashed lines indicate the mean latitude of activity in each hemisphere. We include extrapolated flow values (solid black line: assuming zero velocity at the poles; dashed black line: without constraint) using a least-squares quadratic fit. (Compare with Figure 6.)

amplitudes are smaller and decrease more rapidly with increasing latitude than at shallower or deeper depths. This is also the depth range where the meridional flow is equatorward at 75° latitude in the southern hemisphere. Within this depth range, the meridional flow pattern appears to be shifted (compressed) toward the equator. The maximum amplitude (of the Legendre polynomial fit) is located at 30° latitude compared to 37.5° at shallower depths and 45° at depths greater than 12 Mm. The average maximum amplitude is 22.2 ± 4.3 m s^{-1} within 1 Mm of the surface, 17.0 ± 0.7 m s^{-1} between 1 and 9 Mm, and 23.4 ± 0.3 m s^{-1} at depths greater than 12 Mm. The corresponding errors increase with latitude and depth.

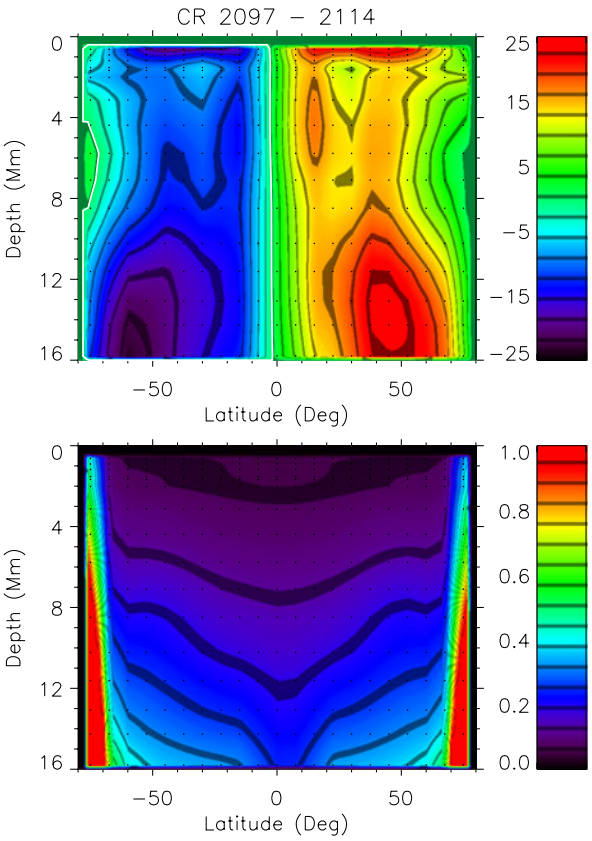


Figure 12 Top: the meridional flow (in $\mathrm{m\,s^{-1}}$) averaged over all 18 Carrington rotations as a function of latitude and depth corrected for the variation with the B_0 angle. Positive values indicate meridional flows to the north, negative values indicate flows to the south. Bottom: the corresponding error values from 18 Carrington rotations with 24 daily measurements each. (Compare with Figure 11.)

Errors are small near the equator at the surface and large near the poles in deep layers. The smallest errors at a given depth occur at small positive latitudes. This slight offset is not too surprising since the average B_0 angle is slightly positive.

Finally, we focus on the variation with latitude and time of the meridional flow. Figures 13 to 15 show the temporal variation at three representative depths. Figure 13 shows the meridional flow close to the surface at a depth of 0.6 Mm. The measured flows vary with the B_0 angle which is noticeable as an annual variation of the contour lines in latitude in addition to the presence of counter cells at the highest latitudes. Equatorward flows appear in the southern hemisphere when the B_0 angle is positive and in the northern one when it is negative. In addition, there are two counter cells in the northern hemisphere at 75° when the B_0 angle is positive but small. Similar counter cells exist in the southern hemisphere. The corrected meridional flow shows only one counter cell; it is in the northern hemisphere at 75° and small positive B_0 angle. The contour lines are now more or less parallel to the equator. The corrected meridional flow at this depth does not change much during the 18 Carrington rotations. The errors are nearly constant with time and they are large at latitudes poleward of 60° and small at other latitudes.

Figure 14 shows the meridional flow at a depth of 7.1 Mm. This is within the depth range where the meridional flow pattern appears shifted (compressed) toward the equator as shown in Figure 12. The measured flow shows strong extended counter cells at high latitudes in both hemispheres, which are more pronounced than in the near-surface flows

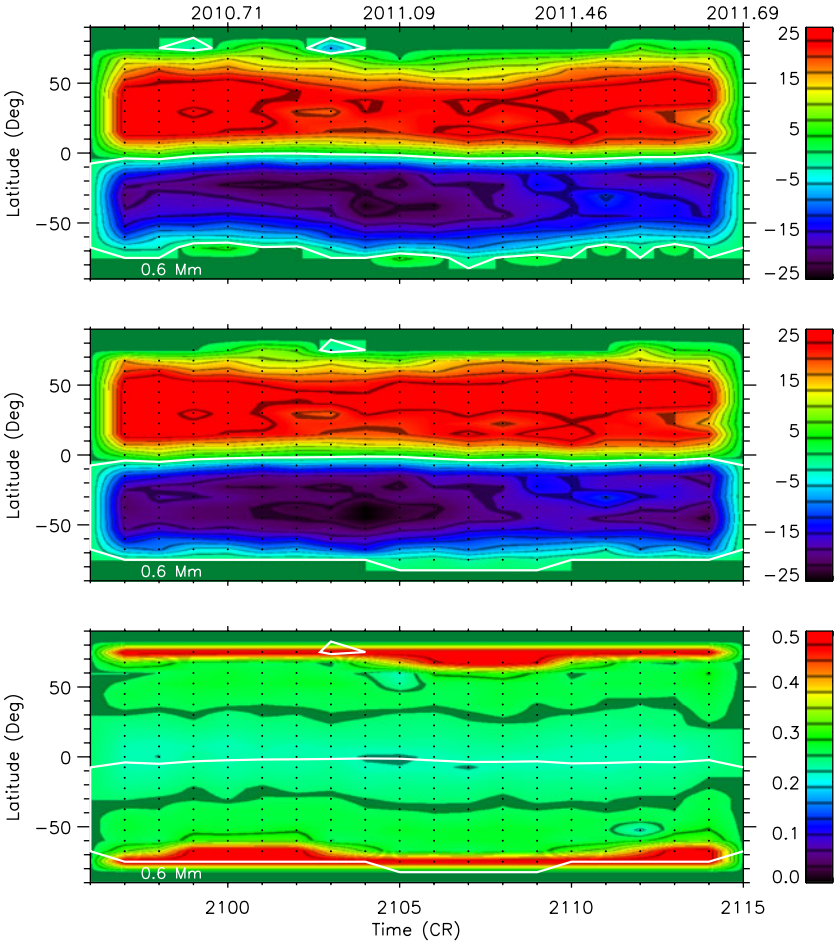


Figure 13 Top: the temporal variation of the meridional flow for CR 2097 – 2114 near the surface at 0.6 Mm. Positive values indicate meridional flows to the north, negative values indicate flows to the south. The white contour line indicates zero flow. Middle: the meridional flow corrected for the B_0-angle variation. Bottom: the corresponding error values. The time coordinate is given in Carrington rotations (bottom x-axis) and calendar years (top x-axis).

shown in Figure 13. In the corrected meridional flow, a few counter cells remain at 75° in both hemispheres, while all except one in the southern hemisphere are gone at 67.5° latitude. The contour lines are more or less parallel to the equator in the corrected flow. The errors are small at low latitudes equatorward of 30° and increase at higher latitudes. The errors are nearly constant with time at most latitudes except near 50° where the error amplitude varies with the B_0 angle.

As Figure 15 shows, the meridional flow at 15.8 Mm shows just as pronounced counter cells as the flows at 7.1 Mm. However, the corrected flow is poleward at all latitudes and epochs except at the end of the data set in the northern hemisphere. The error values are large at high latitudes in the northern hemisphere when the B_0 angle is negative and in the southern hemisphere when the B_0 angle is positive. The smallest errors near the equator also

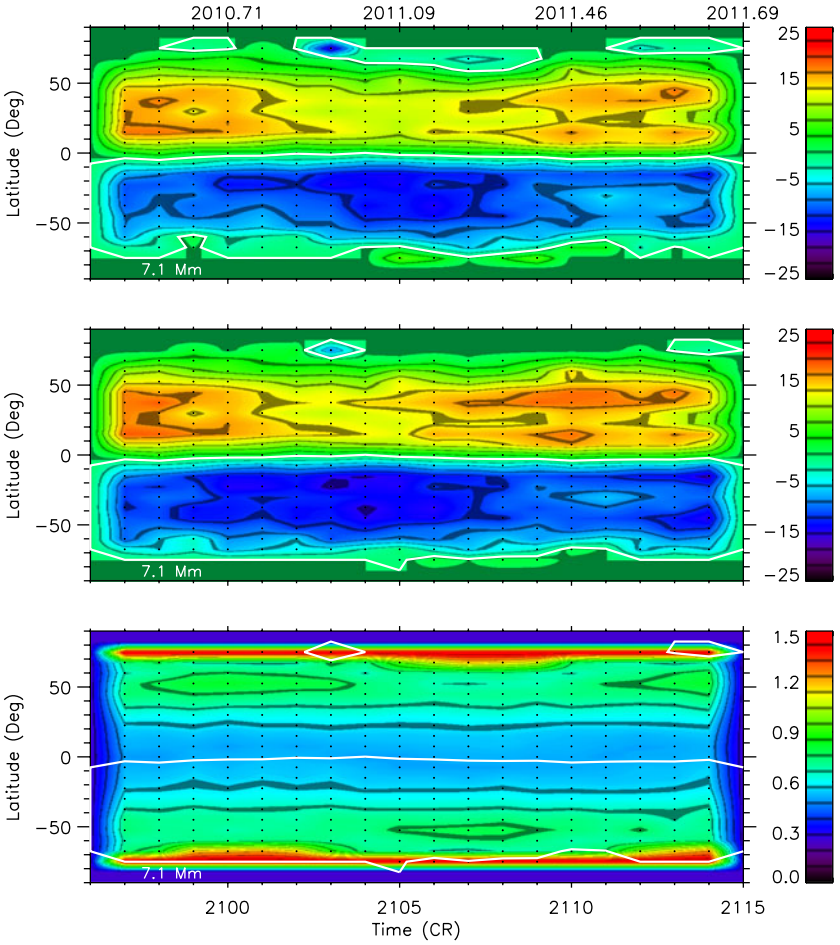


Figure 14 Top: the temporal variation of the meridional flow for CR 2097 – 2114 at 7.1 Mm. Positive values indicate meridional flows to the north, negative values indicate flows to the south. The white contour line indicates zero flow. Middle: the meridional flow corrected for the B_0-angle variation. Bottom: the corresponding error values. The time coordinate is given in Carrington rotations (bottom x-axis) and calendar years (top x-axis). (Compare with Figure 13.)

show this variation with the minimum at northern latitudes when the B_0 angle is positive and at southern latitudes when the B_0 angle is negative.

When comparing the corrected meridional flows at these three depths, it is noticeable that the meridional flow at 7.1 Mm has smaller amplitudes than at deeper or shallower layers. There are some epochs with very small flow amplitudes at 7.1 Mm; the amplitude seems to vary on an annual time scale. As noticed in Figure 12, the flow amplitude at 7.1 Mm decreases much faster toward zero with increasing latitude than the flow amplitude at the two other depths.

We can speculate whether the meridional flow reverses direction near the poles. Figure 16 shows the corrected meridional flow extrapolated to the poles at three depths assuming that the flow is zero at the poles. This is a reasonable assumption since the flow values are averages over one rotation. Small equatorward flows are common in the extrapolated values in

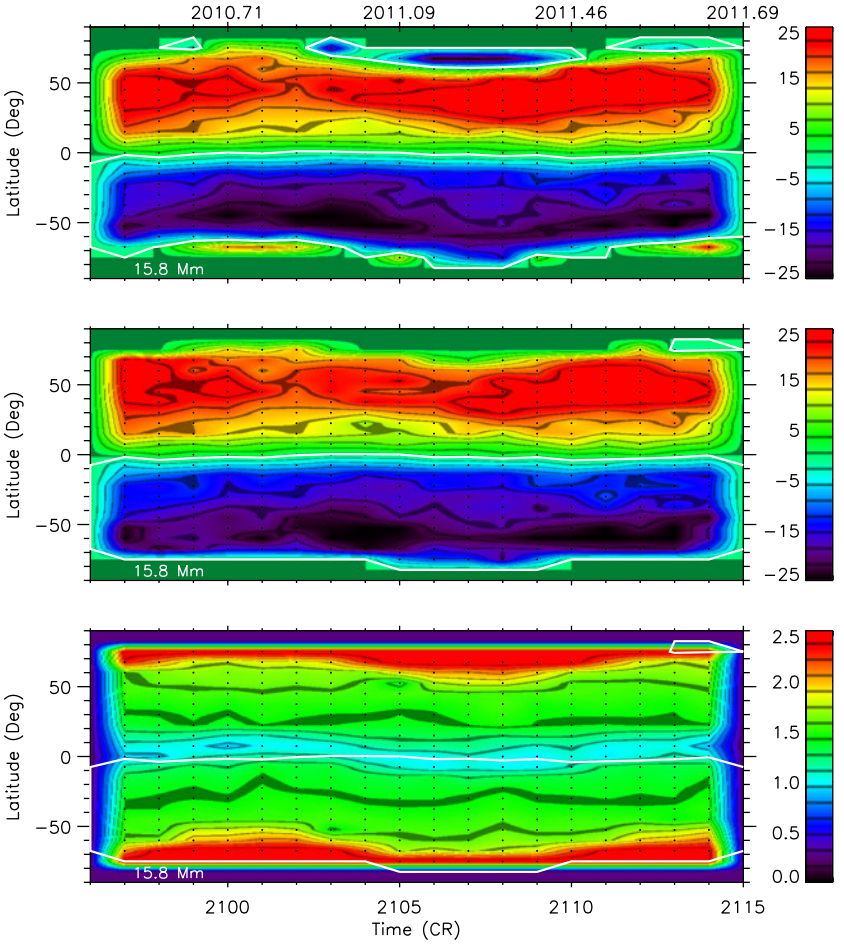


Figure 15 Top: the temporal variation of the meridional flow for CR 2097 – 2114 at 15.8 Mm. Positive values indicate meridional flows to the north, negative values indicate flows to the south. The white contour line indicates zero flow. Middle: the meridional flow corrected for the B_0-angle variation. Bottom: the corresponding error values. The time coordinate is given in Carrington rotations (bottom x-axis) and calendar years (top x-axis). (Compare with Figures 13 and 14.)

both hemispheres at all depths. However, they are generally more pronounced during epochs when the flow errors are large at high latitudes (see Figures 13 – 15 bottom panels) and less pronounced when the flows are measured at 75° latitude. This little exercise indicates that the observations do not rule out the existence of counter cells close to the poles (poleward of 75°).

4. Summary and Conclusions

We have determined the average meridional flows in subsurface layers. For this purpose, we have analyzed high-resolution HMI Dopplergrams with the HMI ring-diagram pipeline and determined meridional horizontal flows within the outer 16 Mm of the convection zone. We

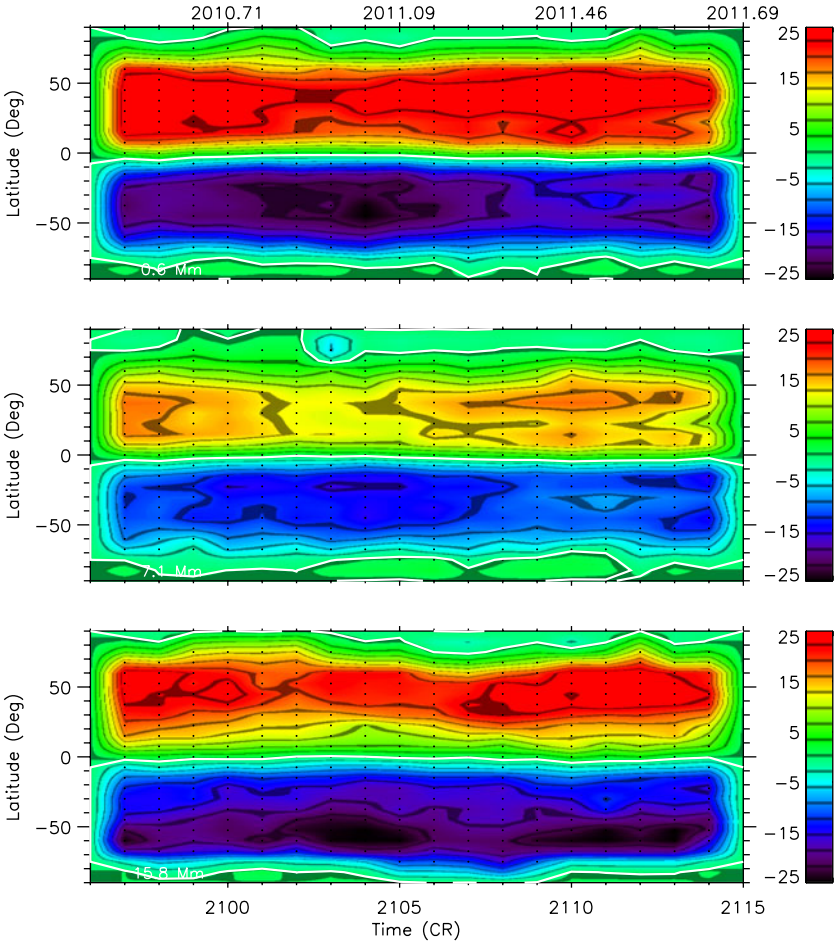


Figure 16 The corrected temporal variation of the meridional flow for CR 2097 – 2114 at 0.6 Mm (top), 7.1 Mm (middle), and 15.8 Mm (bottom). Positive values indicate meridional flows to the north, negative values indicate flows to the south. The white contour line indicates zero flow.The flow values have been extrapolated with a least-squares quadratic fit to higher latitudes assuming that the meridional flow is zero at the poles when averaged over a Carrington rotation. The time coordinate is given in Carrington rotations (bottom x-axis) and calendar years (top x-axis). (Compare with middle panels of Figures 13 to 15.)

have derived the meridional flows as a function of latitude for 18 Carrington rotations. We have calculated the grand average and studied the temporal variation.

The high-resolution HMI Dopplergrams allow us to study the meridional flow at higher latitudes than previously with GONG Dopplergrams or MDI Dynamics Program data. The ring-diagram tiles have an effective diameter of 15° after apodization and are centered on a grid with centers spaced 7.5° in latitude similar to the dense-pack analysis of GONG and MDI data. We find that the average meridional flow is poleward in both hemispheres at least up to $\pm 52.5°$ latitude, which is the range covered by analyses of GONG and MDI data. At 60° and higher, there are occasional equatorward flows. For example, the meridional flow at a depth of 7 Mm is equatorward at 60° in the northern hemisphere during Carrington rotation 2107 and in the southern one during CR 2099. At these poleward latitudes, the

meridional flows clearly vary with the B_0 angle. Such annual variations have been noticed in previous studies of GONG and MDI data (Zaatri *et al.*, 2006; González Hernández *et al.*, 2006, 2008). A non-zero B_0 angle will either increase or decrease the effective distance from disk center at a given latitude and thus enhance or diminish the geometric foreshortening at a given latitude. This variation affects the spatial resolution in the north-south direction and can lead to an artificially enhanced or diminished meridional flow. Equatorward flows or counter cells have been first observed in MDI data in deeper layers (Haber *et al.*, 2002). In GONG data, however, they appear at shallow layers near the surface. Given the much higher spatial resolution of HMI Dopplergrams, it is not surprising that the measured meridional flow in this study shows no equatorward flow at any depth at latitudes studied with GONG and MDI data. This confirms the conclusion from the previous studies that these counter cells are artifacts due to the varying B_0 angle.

Here, we are especially interested in flows at high latitudes between 60° and 75° latitude, which have not been studied with GONG or MDI data using local-helioseismic techniques. The measured meridional flows show an annual variation and are occasionally equatorward. To correct for the variation with the B_0 angle, we perform a linear regression analysis and assume that the best flow estimate is determined at a given latitude in the northern hemisphere when the B_0 angle reaches its maximum and in the southern hemisphere with the B_0 angle at a minimum. Most episodic counter cells disappear when we correct the meridional flow in this way for the B_0-angle variation. In the northern hemisphere, the corrected data show equatorward flows only at 75° latitude with very small amplitudes when the B_0 angle is small. These counter cells are most likely the remains of an insufficiently corrected B_0-angle variation and not of solar origin.

In the southern hemisphere, the corrected meridional flow is equatorward at 75° latitude at depths between 4 and 8 Mm. In this context, it is interesting to note that an analysis of MWO Dopplergrams shows counter cells at latitudes poleward of 60° in both hemispheres during and after solar minimum (Ulrich, 2010). The equatorward flow at $-75°$ in the HMI data could thus be a remnant of such a pattern. Especially, if one considers the extrapolated values that suggest a flow reversal near 78° latitude at most depths in both hemispheres. The extrapolated flow reversal occurs at higher latitudes than in the MWO data, which would be consistent since the HMI data used here cover a later epoch than the MWO data and the counter cells in the MWO data always disappear during the rising phase of the next cycles.

However, it cannot be ruled out that the counter cells in the southern hemisphere are artifacts as well. We have only five Carrington rotations with sufficiently large negative B_0-angle values to measure flows at 75° in the southern hemisphere compared with eight rotations with large positive B_0-angle values. On the other hand, the flow at 75° in the southern hemisphere is consistently equatorward. We obviously need a longer time series to settle the question whether this is an artifact or of solar origin.

The corrected average meridional flow is poleward at most depths and latitudes with a maximum amplitude of about 20 $\mathrm{m\,s^{-1}}$ near 37.5° latitude. This is comparable to the meridional flow at the solar surface derived from MWO Doppler measurements (Ulrich, 2010) and to the subsurface meridional flow derived from a ring-diagram analysis of GONG and MDI data (Haber *et al.*, 2002; Komm *et al.*, 2004; González Hernández *et al.*, 2008; Basu and Antia, 2010). Tracking of magnetic features at the solar surface leads to meridional flows with about 40 % smaller amplitudes (Hathaway and Rightmire, 2010; Komm, Howard, and Harvey, 1993; Snodgrass and Dailey, 1996).

The measured meridional flow also varies with disk position (Komm, Howe, and Hill, 2011). We have taken the approach to measure the variation with disk position and to simply subtract it, since it is an artifact of unknown origin. The measured variation is mainly a CMD

variation with a maximum at the central meridian and an additional east-west trend when the B_0 angle is non-zero. This correction procedure leads to small corrections of the average flow values. Zhao *et al.* (2012) have taken a different approach and tried to deduce a north-south variation of the meridional flow assuming that it is the same as the east-west variation of the zonal flow. They suggest this procedure because it leads to similar flows for different HMI observables when analyzing 10 days of HMI data with a time–distance technique. The resulting flows are much smaller and more in line with values derived from tracking of magnetic features. Clearly, systematics due to disk position need further study.

The subsurface meridional flows are more poleward on the equatorward side of the mean latitude of magnetic activity and less poleward on the poleward side, which can be interpreted as convergent flows near the mean latitude of activity. This also agrees with previous studies of the meridional flow in subsurface layers (Haber *et al.*, 2002; Komm *et al.*, 2004; González Hernández *et al.*, 2008). The meridional flow profile appears to be shifted toward the equator at depths between about 4 and 8 Mm compared to the flow profiles at other depths. The maximum amplitude of a Legendre polynomial fit is located at lower latitudes. In the same depth range, the flow amplitudes are smaller than at other depths. This increase of the amplitude of the meridional flow with depth seems to be counter-intuitive since the density increases by an order of magnitude between 7 and 14 Mm. However, previous studies using GONG and MDI data have also shown that the amplitude of the meridional flow does not decrease with increasing depth in these layers (Haber *et al.*, 2002; Komm *et al.*, 2004; González Hernández *et al.*, 2006). This variation with depth might be related to the existence and dynamics of the near-surface shear layer and deserves further study. In the near future, we will perform a quantitative comparison of the meridional flow derived from HMI and GONG data when we have sufficiently long data sets from both covering the same epoch.

Acknowledgements The data used here are courtesy of NASA/SDO and the HMI Science Team. We thank the HMI team members for their hard work. SOLIS VSM data used here are produced cooperatively by NSF/NSO and NASA/GSFC. We thank the referee for helpful suggestions. The ring-fitting analysis is based on algorithms developed by Haber, Hindman, and Larsen with support from NASA and Stanford University. This work was supported by NASA grant NNX11AQ57G to the National Solar Observatory.

References

Basu, S., Antia, H.M.: 2010, *Astrophys. J.* **717**, 488.
Beckers, J.M.: 2007, *Solar Phys.* **240**, 3. doi:10.1007/s11207-006-0286-z.
Bogart, R.S.: 1987, *Solar Phys.* **110**, 23. doi:10.1007/BF00148199.
Bogart, R.S., Baldner, C., Basu, S., Haber, D.A., Rabello-Soares, M.C.: 2011a, *J. Phys. Conf. Ser.* **271**, 012008.
Bogart, R.S., Baldner, C., Basu, S., Haber, D.A., Rabello-Soares, M.C.: 2011b, *J. Phys. Conf. Ser.* **271**, 012009.
Dikpati, M., Gilman, P.A.: 2012, *Astrophys. J.* **746**, 65.
Dikpati, M.: 2011, *Astrophys. J.* **733**, 90.
Dikpati, M., Gilman, P.A., de Toma, G., Ulrich, R.K.: 2010, *Geophys. Res. Lett.* **37**, 14107.
Dikpati, M.: 2004, In: Danesy, D. (ed.) *Helio- and Asteroseismology: Towards a Golden Future*, **SP-559**, ESA, Noordwijk, 233.
Dikpati, M., Charbonneau, P.: 1999, *Astrophys. J.* **518**, 508.
Durney, B.R.: 1974, *Astrophys. J.* **190**, 211.
Giles, P.M., Duvall, T.L., Scherrer, P.H., Bogart, R.S.: 1997, *Nature* **390**, 52.
Gizon, L., Birch, A.C., Spruit, H.C.: 2010, *Annu. Rev. Astron. Astrophys.* **48**, 289.
Gizon, L., Rempel, M.: 2008, *Solar Phys.* **251**, 241. doi:10.1007/s11207-008-9162-3.
González Hernández, I., Komm, R., Hill, F., Howe, R., Corbard, T., Haber, D.A.: 2006, *Astrophys. J.* **638**, 576.

González Hernández, I., Kholikov, S., Hill, F., Howe, R., Komm, R.: 2008, *Solar Phys.* **252**, 235. doi:10.1007/s11207-008-9264-y.
Haber, D.A., Hindman, B.W., Toomre, J., Bogart, R.S., Larsen, R.M., Hill, F.: 2002, *Astrophys. J.* **570**, 885.
Hathaway, D.H., Rightmire, L.: 2010, *Science* **327**, 1350.
Henney, C.J., SOLIS Team: 2007, In: *AGU Fall Meeting Abstracts*, A1104.
Howard, R.F.: 1996, *Annu. Rev. Astron. Astrophys.* **34**, 75.
Jones, H.P., Harvey, J.W., Henney, C.J., Keller, C.U., Malanushenko, O.M.: 2004, *Bull. Am. Astron. Soc.* **36**, 709.
Komm, R.W., Corbard, T., Durney, B.R., González Hernández, I., Hill, F., Howe, R., Toner, C.: 2004, *Astrophys. J.* **605**, 554.
Komm, R.W., Howe, R., Hill, F.: 2009, *Solar Phys.* **258**, 13. doi:10.1007/s11207-009-9398-6.
Komm, R.W., Howe, R., Hill, F.: 2011, *Solar Phys.* **268**, 407. doi:10.1007/s11207-010-9692-3.
Komm, R.W., Howard, R.F., Harvey, J.W.: 1993, *Solar Phys.* **147**, 207. doi:10.1007/BF00690713.
Miesch, M.S., Hindman, B.W.: 2011, *Astrophys. J.* **743**, 79.
Pesnell, W.D., Thompson, B.T., Chamberlin, P.C.: 2012, *Solar Phys.* **275**, 3. doi:10.1007/s11207-011-9841-3.
Schou, J., Scherrer, P.H., Bush, R.I., Wachter, R., Couvidat, S., Rabello-Soares, M.C., *et al.*: 2012, *Solar Phys.* **275**, 229. doi:10.1007/s11207-011-9842-2.
Snodgrass, H.B., Dailey, S.B.: 1996, *Solar Phys.* **163**, 21. doi:10.1007/BF00165454.
Švanda, M., Kosovichev, A.G., Zhao, J.: 2008, *Astrophys. J.* **680**, L161.
Ulrich, R.K.: 2010, *Astrophys. J.* **725**, 658.
Zaatri, A., Komm, R., González Hernández, I., Howe, R., Corbard, T.: 2006, *Solar Phys.* **236**, 227. doi:10.1007/s11207-006-0106-5.
Zhao, J., Kosovichev, A.G.: 2004, *Astrophys. J.* **603**, 776.
Zhao, J., Nagashima, K., Bogart, R.S., Kosovichev, A.G., Duvall, T.L. Jr.: 2012, *Astrophys. J. Lett.* **749**, L5.

DOI 10.1007/978-1-4899-8005-2_7
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0180-9

Properties of High-Frequency Wave Power Halos Around Active Regions: An Analysis of Multi-height Data from HMI and AIA Onboard SDO

S.P. Rajaguru · S. Couvidat · Xudong Sun · K. Hayashi · H. Schunker

Received: 4 May 2012 / Accepted: 22 October 2012 / Published online: 22 November 2012

Abstract We study properties of waves of frequencies above the photospheric acoustic cut-off of ≈ 5.3 mHz, around four active regions, through spatial maps of their power estimated using data from the *Helioseismic and Magnetic Imager* (HMI) and *Atmospheric Imaging Assembly* (AIA) onboard the *Solar Dynamics Observatory* (SDO). The wavelength channels 1600 Å and 1700 Å from AIA are now known to capture clear oscillation signals due to helioseismic *p*-modes as well as waves propagating up through to the chromosphere. Here we study in detail, in comparison with HMI Doppler data, properties of the power maps, especially the so-called "acoustic halos" seen around active regions, as a function of wave frequencies, inclination, and strength of magnetic field (derived from the vector-field observations by HMI), and observation height. We infer possible signatures of (magneto)acoustic wave refraction from the observation-height-dependent changes, and hence due to changing magnetic strength and geometry, in the dependences of power maps on the photospheric magnetic quantities. We discuss the implications for theories of *p*-mode absorption and mode conversions by the magnetic field.

Keywords Active regions, magnetic fields · Chromosphere, active · Helioseismology, observations · Magnetic fields, photosphere · Magnetohydrodynamics · Sunspots, magnetic fields · Waves, acoustic, magnetohydrodynamic, modes, propagation

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

S.P. Rajaguru (✉)
Indian Institute of Astrophysics, Bangalore 34, India
e-mail: rajaguru@iiap.res.in

S. Couvidat · X. Sun · K. Hayashi
Hansen Experimental Physics Lab, Stanford University, Stanford CA, USA

S. Couvidat
e-mail: couvidat@stanford.edu

H. Schunker
Max-Planck Institute for Solar System Research, Katlenburg-Lindau, Germany
e-mail: schunker@mps.mpg.de

1. Introduction

Enhanced power of high-frequency waves surrounding strong-magnetic-field structures such as sunspots and plages is one of several intriguing wave-dynamical phenomena observed in the solar atmosphere. This excess power known as "acoustic halo", first observed in the early 1990s at photospheric (Brown *et al.*, 1992) as well as chromospheric (Braun *et al.*, 1992; Toner and LaBonte, 1993) heights, is at frequencies above the photospheric acoustic cut-off of ≈ 5.3 mHz, in the range of $5.5-7$ mHz, and over regions of weak- to intermediate-strength ($50-250$ G) photospheric magnetic field. A good number of observational studies (Hindman and Brown, 1998; Thomas and Stanchfield, 2000; Jain and Haber, 2002; Finsterle *et al.*, 2004; Moretti *et al.*, 2007; Nagashima *et al.*, 2007) since then have brought out additional features, and we refer the reader to Khomenko and Collados (2009) for a succinct summary of them as known prior to 2009. On the theoretical side, no single model describing all of the observed features has been achieved yet, although there have been several focused efforts (Kuridze *et al.*, 2008; Hanasoge, 2008; Shelyag *et al.*, 2009; Hanasoge, 2009; Khomenko and Collados, 2009). However, a large number of studies centered around modeling acoustic wave–magnetic-field interactions over heights from the photosphere to chromosphere, with relevance to high-frequency power excess observed around sunspots, have been carried out (Rosenthal *et al.*, 2002; Bogdan *et al.*, 2003; Bogdan and Judge, 2006; Cally, 2006; Schunker and Cally, 2006; Khomenko and Collados, 2006, 2008; Jacoutot *et al.*, 2008; Khomenko *et al.*, 2009; Vigeesh, Hasan, and Steiner, 2009; Khomenko and Cally, 2012). A central theme of all the above theoretical studies, except that of Jacoutot *et al.* (2008), has been the conversion of acoustic wave modes (from below the photosphere) into magnetoacoustic wave modes (the fast and slow waves) at the magnetic canopy defined by the plasma $\beta = 1$ layer. Enhanced acoustic emission by magnetically modified convection over weak- and intermediate-field regions, suggested as one possible mechanism by Brown *et al.* (1992) and further advocated by Jain and Haber (2002), was found to be viable, through 3D numerical simulations of magneto-convection, by Jacoutot *et al.* (2008). It should perhaps be noted here that Hindman and Brown (1998) suggested some form of field-aligned incompressible wave motions as agents for excess power over magnetic regions, implied by their finding, from SOHO/MDI observations, of visibility of these halos in photospheric Doppler velocities but not in continuum intensities. This suggestion, however, seems to contradict chromospheric-intensity observations (Braun *et al.*, 1992; Moretti *et al.*, 2007), which show that the halos at these heights, in terms of their dependence on magnetic-field strength as well as frequency behavior, are clearly related to the photospheric halos; Hindman and Brown's reasoning that the chromospheric Ca K intensities observed by Braun *et al.* (1992) have large cross-talk from Doppler shifts are, however, contradicted by the simultaneous velocity and intensity observations made by Moretti *et al.* (2007).

In a recent study, Schunker and Braun (2011) have brought out a few new properties, *viz.*

i) the largest excess power in halos is at horizontal magnetic-field locations, in particular, at locations between opposite-polarity regions,
ii) the larger the magnetic-field strength the higher the frequency of peak power, and
iii) the modal ridges over halo regions exhibit a shift towards higher wavenumbers at constant frequencies.

Although none of the proposed theoretical explanations or mechanisms causing the power halos are able to match all of the observed properties, and hence provide an acceptable theory, the mechanism based on MHD fast-mode refraction in canopy-like structure

of strong expanding magnetic field studied by Khomenko and Collados (2009) appears to match some major observed features. This theory also predicts certain other observable features that we will address here.

From what has been learned so far, from observations as well as theoretical studies, it is clear that transport and conversion of energy between magneto-acoustic wave modes, which are driven by acoustic waves and convection from below the photosphere and mediated by the structured magnetic field in the overlying atmosphere, provide a plausible approach for identifying the exact mechanism. A crucial diagnostic of such wave processes requires probing several heights in the atmosphere simultaneously with magnetic-field information. The instruments HMI and AIA onboard SDO, with photospheric Doppler and vector magnetic-field information from the former and the upper-photospheric and lower-chromospheric UV emissions in the 1700 Å and 1600 Å wavelength channels of the latter, provide some interesting possibilities for such studies. We exploit this opportunity, and make a detailed analysis of high-frequency power halos around four different active regions, over at least five different heights from the photosphere to chromosphere.

2. Data and Analysis Method

We use data from the *Helioseismic and Magnetic Imager* (HMI: Scherrer *et al.*, 2012) and *Atmospheric Imaging Assembly* (AIA: Lemen *et al.*, 2012) onboard the *Solar Dynamics Observatory* (SDO): photospheric Doppler velocity [v], continuum intensity [I_c], line-core intensity [I_{co}], and disambiguated vector magnetic field [B_x, B_y, and B_z] derived from HMI observations, and chromospheric UV emissions observed by AIA in the wavelength channels 1700 Å and 1600 Å, which we denote I_{uv1} and I_{uv2}, respectively. The intensities I_{uv1} and I_{uv2} are now known to capture clear oscillation signals due to helioseismic p-modes as well as propagating waves in the atmosphere (Howe *et al.*, 2010; Hill *et al.*, 2011). The photospheric observations by HMI are in the form of filtergram images captured from across the magnetically sensitive line Fe I 6173.34 Å using two different cameras: the Doppler camera uses six images each of the left- and right-circular polarization components (Stokes $I + V$, and $I - V$) to measure v, I_c, the line depth [I_{ld}], and line width with a cadence of 45 seconds (Scherrer *et al.*, 2012), while the second vector-field camera records the full Stokes vector [I, Q, U, V] in six combinations over six wavelengths (*i.e.*, a total of 36 filtergrams) with a cadence of 135 seconds (Hoeksema *et al.*, 2013). It should perhaps be pointed out that I_c, I_{co}, and v from HMI form at three different heights spread over $z = 0-300$ km above the continuum optical depth $\tau_c = 1$ ($z = 0$ km) level (Norton *et al.*, 2006): I_c is from about $z = 0$ km, v corresponds to an average height of about $z = 140$ km (Fleck, Couvidat, and Straus, 2011), and the line-core intensity calculated as $I_{co} = I_c - I_{ld}$ corresponds to the top layer, at about $z = 280-300$ km, of the line formation region (Norton *et al.*, 2006; Fleck, Couvidat, and Straus, 2011). In addition, exploiting the availability of individual filtergram (level 1) images from HMI, we derive a Doppler velocity [v_{50}] from only the outer pair of filtergrams [I_0 and I_5], which are measured at + 172 mÅ and − 172 mÅ, respectively, from the rest-frame line-center wavelength of 6173.34 Å: $v_{50} = k_{50}(I_5 - I_0)/(I_5 + I_0)$, where the calibration constant k_{50} is derived from the spacecraft velocity [OBSVR] known to very good accuracy. Filtergrams I_0 and I_5 sample the line wings and hence a height corresponding to the lower end within the line formation region (Norton *et al.*, 2006), *i.e.* at $z = 20$ km. We should note here that the positions of I_0 and I_5, and hence the height level that they sample, depend on Zeeman broadening and the line shifts due to the spacecraft velocity: in fields stronger than about 1000 G and when OBSVR exceeds about 2 km s^{-1}, these wing

filtergrams do not sample the blue and red wings symmetrically and this not only invalidates the above height identification, but also corrupts the Doppler-velocity estimate. Since we focus only on the power excess, which is seen over weak and intermediate field strengths not exceeding about 800 G (see Figures 5 – 10), use of v_{50} in our analysis is not subject to the above uncertainties (on height of formation and Doppler amplitudes). The AIA 1700 Å and 1600 Å intensities [$I_{\rm uv1}$ and $I_{\rm uv2}$] form at average heights of 360 km and 430 km (Fossum and Carlsson, 2005), respectively. Thus, we have velocities and intensities from at least five different heights ranging from $z = 0$ to 430 km.

Our chosen target regions of study are the four active regions, NOAA 11092, 11161, 11243, and 11306, each of spatial size covered in 512×512 pixels, with a sampling rate of 0.06 degrees (heliographic) per pixel, and tracked at the Carrington rotation rate for about 14 hours each during their central meridian passage dates of 3 August 2010, 11 February 2011, 3 July 2011, and 2 October 2011, respectively. Both the HMI and AIA data cubes for the above regions are temporally and spatially aligned through the tracking and remapping (Postel) routines of the Stanford SDO data pipeline systems. The vector magnetic field over the four regions are determined using the HMI vector-field pipeline, which does the following: Stokes parameters derived from filtergrams observed over a 12-minute interval are inverted using a Milne–Eddington-based algorithm, the *Very Fast Inversion of the Stokes Vector* (VFISV: Borrero *et al.*, 2011); the 180° azimuthal ambiguity in transverse field is resolved by an improved version of the *minimum energy algorithm* (Metcalf, 1994; Metcalf *et al.*, 2006; Leka *et al.*, 2009). A detailed description of the production and relevant characteristics and outstanding questions regarding HMI vector-field data reduction are also described by Hoeksema *et al.* (2013). The vector-field maps over the target regions are tracked and remapped the same way as for the other data.

From the disambiguated vector-field maps [B_x, B_y, B_z] we determine the field inclination [γ] defined as $\tan(\gamma) = B_z/B_{\rm h}$, where $B_{\rm h} = \sqrt{B_x^2 + B_y^2}$, similarly to Schunker and Braun (2011). Hence, γ ranges from $-90°$ to $+90°$ with $\gamma = 0°$ when the field is purely horizontal, $\gamma < 0°$ when $B_z < 0$ and $\gamma > 0°$ when $B_z > 0$. Figure 1 displays images of $I_{\rm c}$, absolute total field strength $B = \sqrt{B_x^2 + B_y^2 + B_z^2}$, and γ; the latter two are averages over the period of observation used in this work.

Using the 14-hour long time series of images prepared as above, we calculate maps of power summed over a 1-mHz band of frequencies centered every 0.25 mHz in the frequency range of 2 – 10.75 mHz, and normalize them with average power estimated over a quiet-Sun patch identified on each date (*i.e.* on each target region, see Figure 1) separately. Figures 2 – 4 display the power maps obtained from the different observables, which correspond to different heights in the atmosphere as explained above. In the following sections we present results of our analyses of these power maps as a function of B and γ.

3. Magnetic Field and High-Frequency Power Enhancements

Almost all previous observational studies of high-frequency power enhancements around active regions have analysed them as a function of LOS magnetic field, except for Schunker and Braun (2011), who used magnetic-field inclinations derived from potential-field extrapolations of MDI-LOS magnetograms. Here, we use more direct measurements of the vector field from HMI to analyse the power enhancements around sunspots and active regions as a function of absolute total field [$B = |\mathbf{B}|$] and inclination [γ]. Since we are interested in studying wave power around sunspots, not within them, we mask out areas within sunspots:

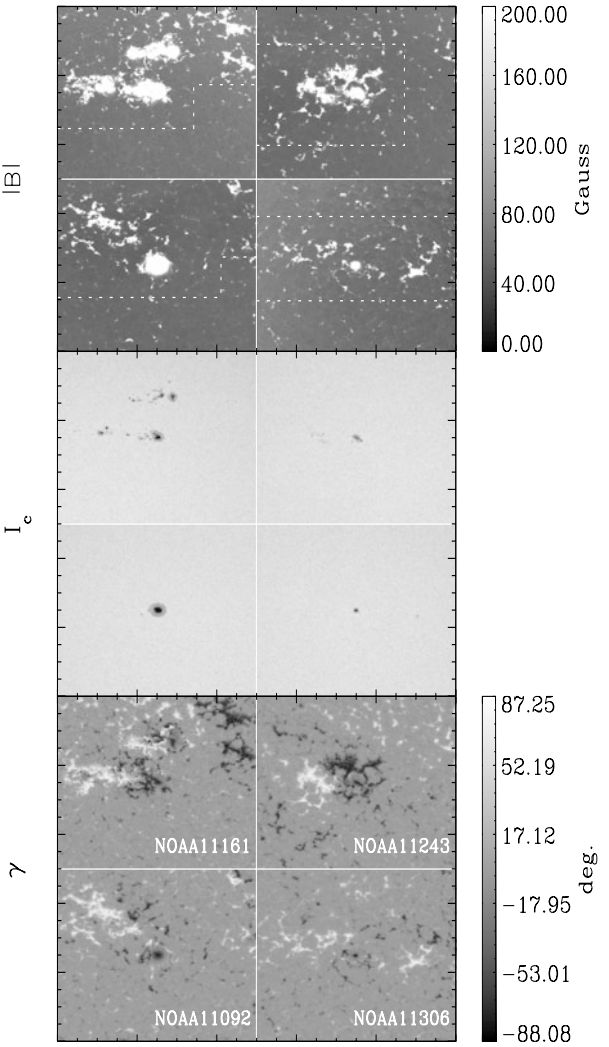


Figure 1 Average total magnetic field [B] (top panel, pixels above 200 G are saturated in the grey scale), a snapshot of continuum intensity [I_c] (middle) and average magnetic-field inclination [γ] (bottom) of the four active regions studied. Averages of B and γ shown here are over the length of time (14 hours) used for power map estimation. Areas covered between white dashed lines and boundary axes in the top panel, in each region, are the largely non-magnetic ones used for estimation of quiet-Sun power of waves studied here. Each active region here covers a square area of 373×373 Mm^2.

pixels falling within the outer-penumbral boundary defined by $0.92I_c$ are excluded from the analyses. This eliminates regions with B stronger than about 850 G, and hence we do not study the highly suppressed power seen within sunspots as well as other features such as three-minute umbral oscillations and penumbral p-mode power due to wave propagation caused by the reduced acoustic cut-off frequency in inclined fields (Rajaguru *et al.* 2007, 2010).

Normalized power maps, *i.e.* power in quiet-Sun units, are averaged over 10 G bins in B and 4° bins in γ. For ease of analysis and appreciation of major features in the variation of halo power *versus* B and γ, we produce two sets of figures: the first set (Figures 5, 7, and 9) shows power *versus* B, with pixels grouped in three different ranges of $|\gamma|$, *viz.* nearly horizontal field with $|\gamma| < 16°$, inclined fields with $16° < |\gamma| < 60°$, and nearly vertical fields with $|\gamma| > 60°$; the second set shows the (Figures 6, 8, 10) dependences of power on γ, with pixels grouped in three different ranges of B, *viz.* $B < 100$ G, $100 < B < 200$ G, and

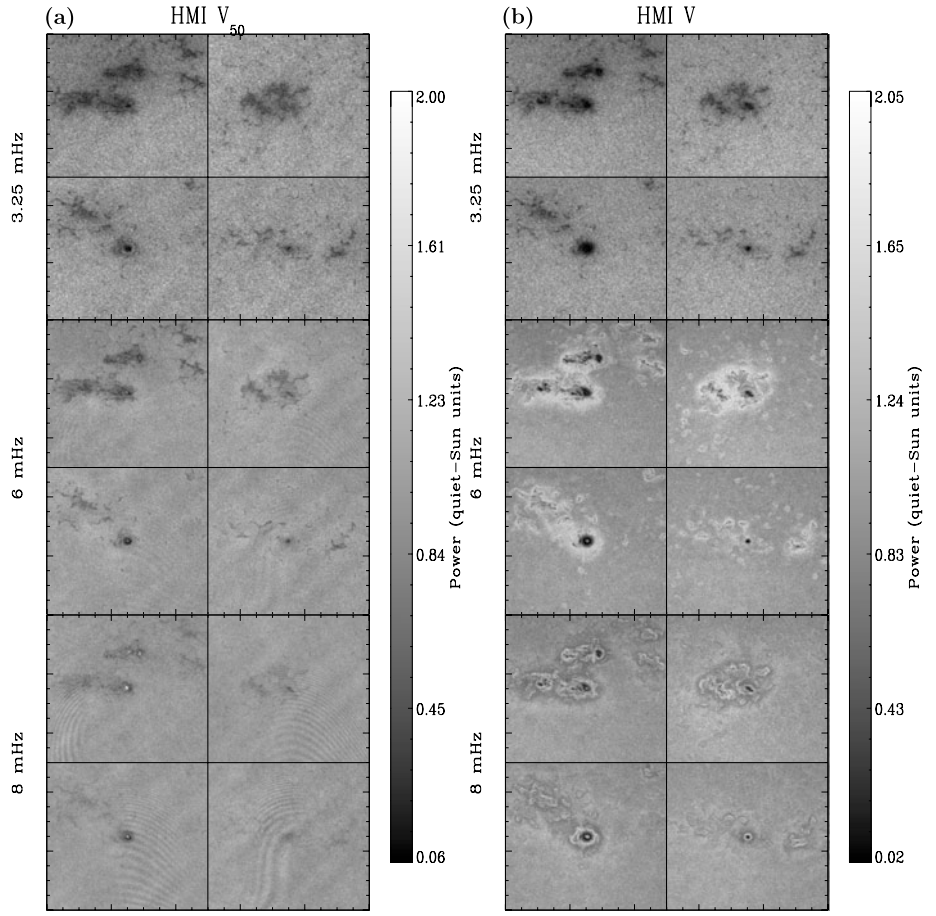


Figure 2 Normalized, with respect to quiet Sun, power over the four active regions (arranged as in Figure 1), each covering a square area of 373 × 373 Mm^2, at three representative frequencies: 3.25, 6, and 8 mHz (top, middle, and bottom panels, respectively). Left part (a), of the figure is obtained from v_{50} and the right part (b), is from v.

$200 < B < 450$ G. We note here that B, in the average over the period of observation, has a noise background of about 40 G in magnitude, and hence this value is a rough minimum for B. We further note that, for results in Figures 5 – 9, using a fairly wide band of 1 mHz in ν centered every 0.25 mHz (*i.e.* overlapping ranges along the y-direction in the figures) and the process of averaging over $B = 10$ G or $\gamma = 4^\circ$ bins tend to produce vertical stripes, and these are especially prominent in Figure 9; however, such stripes do not affect significantly the physical signatures studied and discussed in this article.

We add a caveat here that the analysis of power halos in terms of photospheric values for B or γ do not capture accurately the positions of peak power in $B - \nu$ or $\gamma - \nu$ space for low values of B and $|\gamma|$, especially for observables representing the higher layers. For example, it is obvious from the spatial power maps in Figures 2 – 4 that enhanced power does appear over very weak- or non-magnetic pixels (*i.e.*, those with minimum values for B or $|\gamma|$) around sunspots or between opposite-polarity regions. In this situation, the process of averaging pixels falling within certain small bin sizes of B leads to incorrect B-dependence

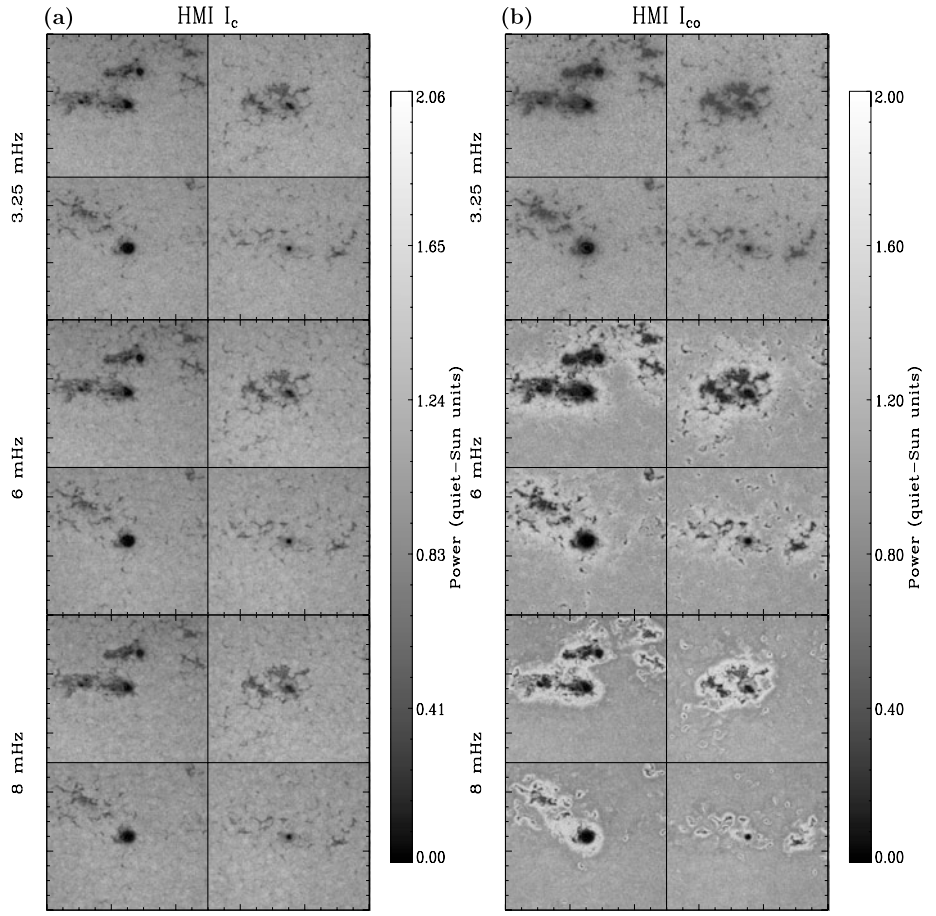


Figure 3 Normalized power over the four active regions (arranged as in Figure 1) at three representative frequencies, 3.25, 6, and 8 mHz (top, middle, and bottom panels, respectively) obtained from observables HMI I_c [(a)] and I_{co} [(b)]. Each active region covers a square area of 373×373 Mm^2.

of peak power over the low-end values of B: far away quiet-Sun pixels with no excess power at all get mixed with those that are non-magnetic but with enhanced power; this should, however, not affect the variations seen, say, above about $B = 50$ G, or $|\gamma| >$ a few degrees. A more accurate and physically meaningful association of wave power observed at a given height requires B and γ values for this height from appropriate magnetic-field extrapolation modeling. For our analysis here, we focus on several other interesting properties of power maps that are not affected by the above finer issues, and we defer the above sophistication to a future analysis.

3.1. Photospheric Behavior of Power Halos

The lowest height observables I_c and v_{50}, as seen from Figures 2a, 3a, and 5 – 8, do not show any appreciable high-frequency halos: while I_c shows no excess power at all, v_{50} shows a very small excess, less than 1 %. The power maps from v_{50}, especially at $\nu = 6$ and 8 mHz, are modulated by some large- and small-scale fringes, which are of instrumental

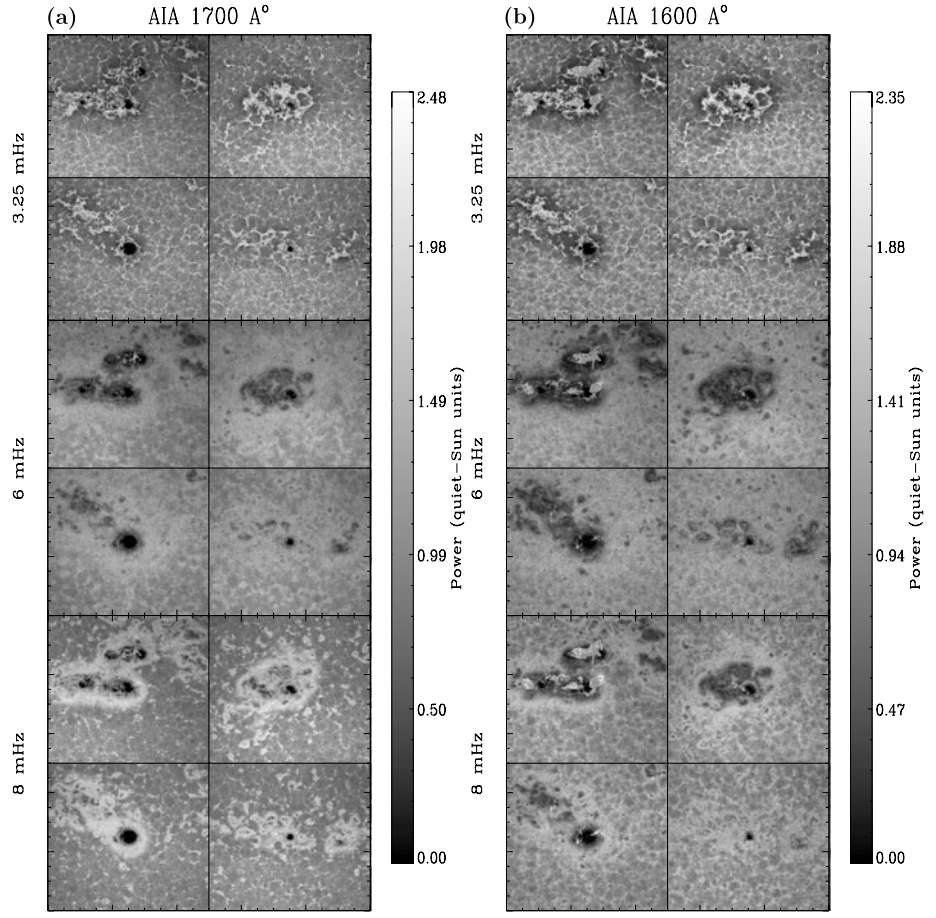


Figure 4 Normalized power over the four active regions (arranged as in Figure 1) at three representative frequencies, 3.25, 6, and 8 mHz (top, middle, and bottom panels, respectively) obtained from observables AIA 1700 Å [$I_{\rm uv1}$, (a)] and 1600 Å [$I_{\rm uv2}$, (b)]. Each active region covers a square area of 373 × 373 Mm^2.

origin: the procedure of estimating v_{50} from I_0 and I_5, calibrating and tracking it is different from the standard HMI pipeline procedure for v, and these seem to cause the CCD flat-field leak into the final velocity estimates. We have so far been unsuccessful in removing them. These fringes of instrumental origin, as can be clearly seen in Figure 2a, appear all over and hence do not cause any confusion in the identification and analyses of halos in and around the magnetic regions. Our main result from v_{50}, *viz.* that the power halos have very small magnitude, appears sound: since fringes modulate the power distribution in equal and opposite magnitudes they do not contribute any net artificial signals in our analysis of real solar signals.

It is interesting to note that there is a broad and uniform distribution as a function of wave frequency [ν] (Figures 5 – 8) above about 5 mHz as well as a function of γ, of power seen in v_{50} and I_c over weak ($B < 100$ G) field and quiet-Sun areas. This feature indicates a more or less uniform injection of high-frequency wave power at the lowest photospheric heights. These high-frequency waves arrive directly from acoustic sources associated with

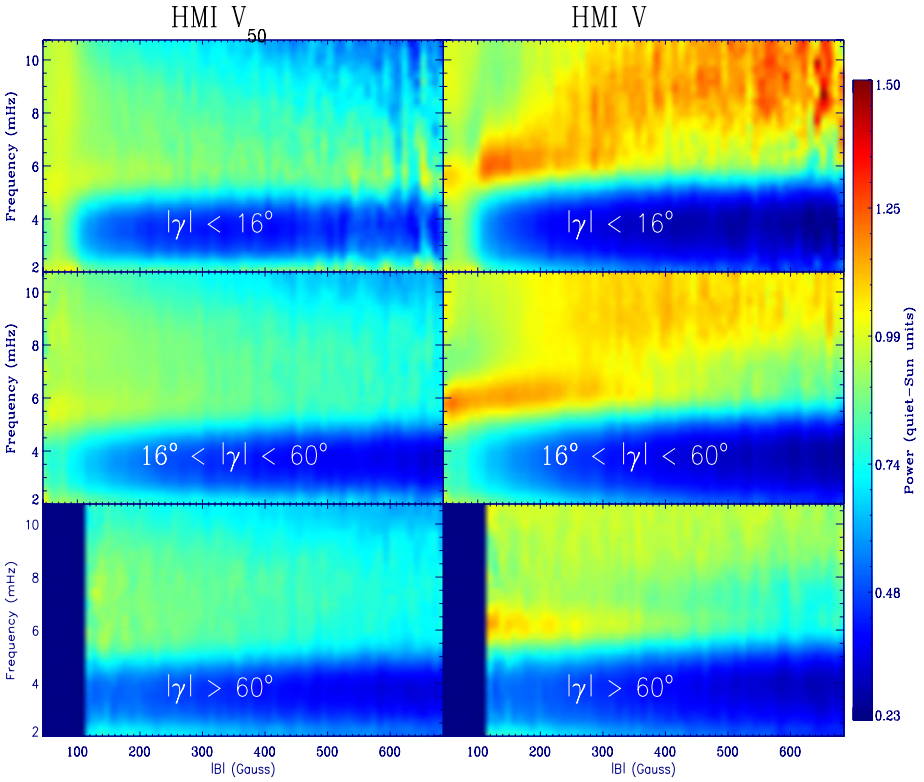


Figure 5 Power, averaged over three different ranges of γ (as marked in the panels), as a function of total magnetic-field strength [B], for all the four active regions combined. Left column is from observable HMI v_{50} and the right from HMI v. Each active region covers a square area of 373×373 Mm^2.

the convective turbulence just below the photosphere, and provide an uniform background of wave energy above the acoustic cut-off, in agreement with the standard picture of propagation through the atmosphere of waves with such frequencies. However, as can be seen in wave-power measurements from v (Figures 2b, 5, and 6), which corresponds to a height of about 140 km, the magnetic field profoundly modifies wave propagation between heights $z = 0$ and 140 km and causes power excess, relative to non-magnetic areas at the same height level, depending on B and γ: for the familiar and well-observed halo in the ν range of 5.5 – 7 mHz, power over weak fields ($B \leq 100$ G) is a slightly increasing function of $|\gamma|$ (see the top-right and middle-right panels in Figure 6 and 5, respectively) and over the intermediate-strength ($100 < B < 300$ G) fields it decreases in amplitude as $|\gamma|$ increases (from top to bottom panels in the right of Figure 5), *i.e.* the more horizontal the field the larger is the power excess. Overall, the maximum excess is in the latter of the above, and this property has been the most known one from earlier studies. In addition, as B increases, between 40 and 250 G, the excess power slowly shifts to higher frequencies, and this is the behavior seen and reported by Schunker and Braun (2011). We associate the weak B (≤ 100 G) inclined field ($|\gamma| > 16°$) behavior to the small-scale flux elements of plages and network flux tubes.

An interesting new feature, revealed by the higher-cadence HMI velocity data for the first time, is a branch of the power halo in the frequency range 7 mHz and above with peak values

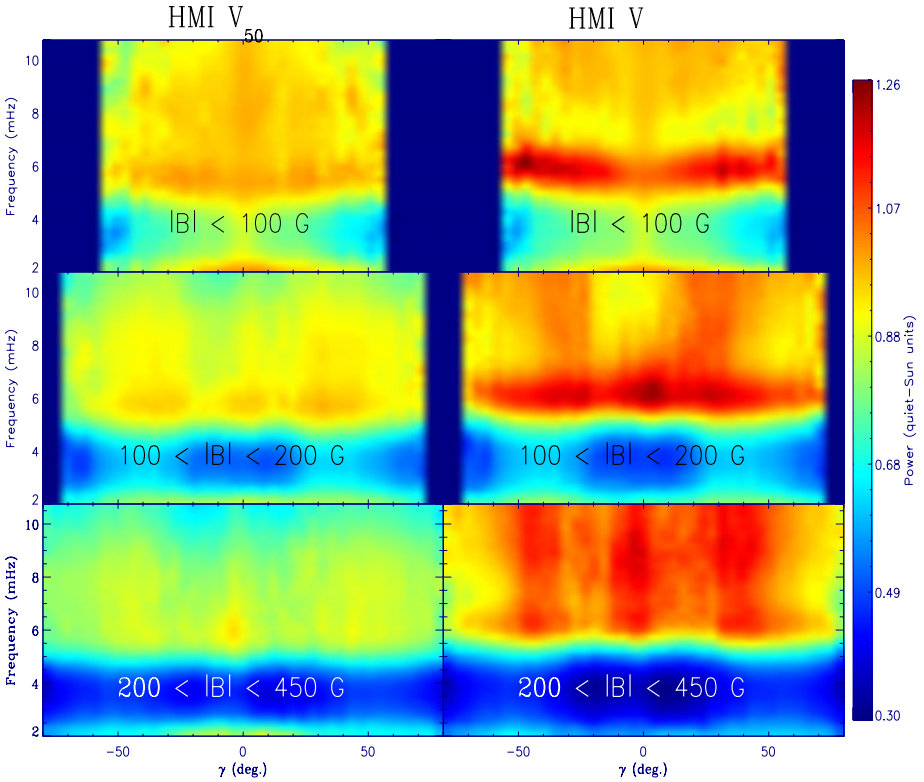


Figure 6 Power, averaged over three different ranges of B (as marked in the panels), as a function of γ for all the four active regions combined. Left column is from observable HMI v_{50} and the right from HMI v. Each active region covers a square area of 373×373 Mm^2.

at about 8 mHz. This higher-ν halo connects to the 6-mHz halo between 200 and 300 G, and its frequencies depend on B more sharply in the above B range (see Figures 5 and 6, right column). Here again, the halo is prominent when $|\gamma| < 16°$ but at $B > 250$ G, *i.e.* in stronger horizontal fields at the base regions of canopy surrounding large structures such as sunspots; these compact halos are distinct in the spatial maps (Figure 2b, bottom panel). However, as seen in Figure 6, power above 7 mHz exhibits slightly more complex γ- and B-dependence in inclined-field areas ($20° < |\gamma| < 60°$) when $B > 100$ G: the excess power shifts to less inclined fields as B increases, and these are seen as two pillars of power (corresponding to the two polarities with same γ, *i.e.* at $\pm\gamma$) that shift to larger $|\gamma|$ as B increases. We identify this latter $|\gamma| > 20°$ halos as those surrounding the small-scale flux elements, which possibly do not fan out too strongly to produce large horizontal fields at heights where v is measured. If we associate the locations of power halos, in height [z] and B, to layers where the plasma $\beta \approx 1$ (see Section 4), then the above features (in Figures 5 and 6) would indicate that these layers coincide with large horizontal-field locations surrounding sunspots while the same would occur over similar height (*i.e.* about the same strength) but inclined fields ($20° < |\gamma| < 60°$) of small-scale flux tubes. This, in turn, would imply a slightly slower rate of expansion of small-scale flux tubes than the sunspot fields. Thus, it is clear that the existence of power halos at $\nu > 7$ mHz is dictated largely by $\beta = 1$ transitions (*i.e.* by B) through the atmosphere, rather than by γ.

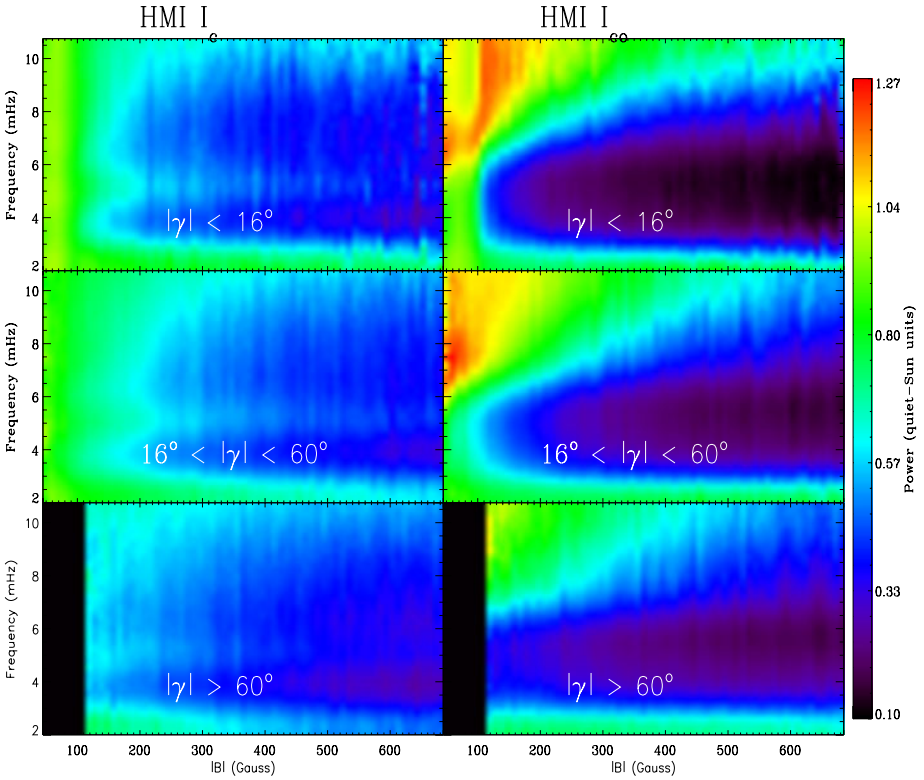


Figure 7 Power, averaged over three different ranges of γ (as marked in the panels), as a function of total magnetic-field strength [B], for all the four active regions combined, from observables HMI I_c (left column) and HMI I_{co} (right column). Each active region covers a square area of 373×373 Mm^2.

Finally, as an interesting and important feature that possibly relates to wave refraction and associated horizontal propagation around the $\beta = 1$ layer (Khomenko and Collados, 2009), we identify the reduced-power region between the outer, much extended, weak enhancement (over $B < 100$ G region, top-right panel of Figure 6) and the strong compact halo (closer to fields of $B = 200$ G and above) discussed above. This curious reduced-power region in spatial maps (Figure 2b, bottom panel) corresponds, in the B *vs* ν maps of Figure 5 (top two right-side panels), to a region that starts at (ν, B) $\approx$ (7 mHz, 40 G) and extends across a curved region over B in the range of 100 – 200 G and higher ν. The values of power in this region, as light-green shade in Figure 5, are less than those obtained from v_{50} over the same $\nu - B$ region (seen as yellowish region over the same location in the left panel of Figure 5). We reason that these features could be signatures of a refracting fast magneto-acoustic wave traveling from high β layers below (Khomenko and Collados, 2009), corresponding to one or both of the following two scenarios:

i) The intermediate $B = 100 - 200$ G field corresponds to $\beta = 1$ height well above 140 km from where v signals arise, and hence no excess power at this latter height, whereas this height difference is close to zero over the stronger-field compact halo region closer to the spots and strong-field structures; the outer much-extended weak 8-mHz halo over the weak field region ($B < 100$ G) could be due to the fast-to-slow-mode conversion resulting in propagation along the field lines that curve back towards the photosphere.

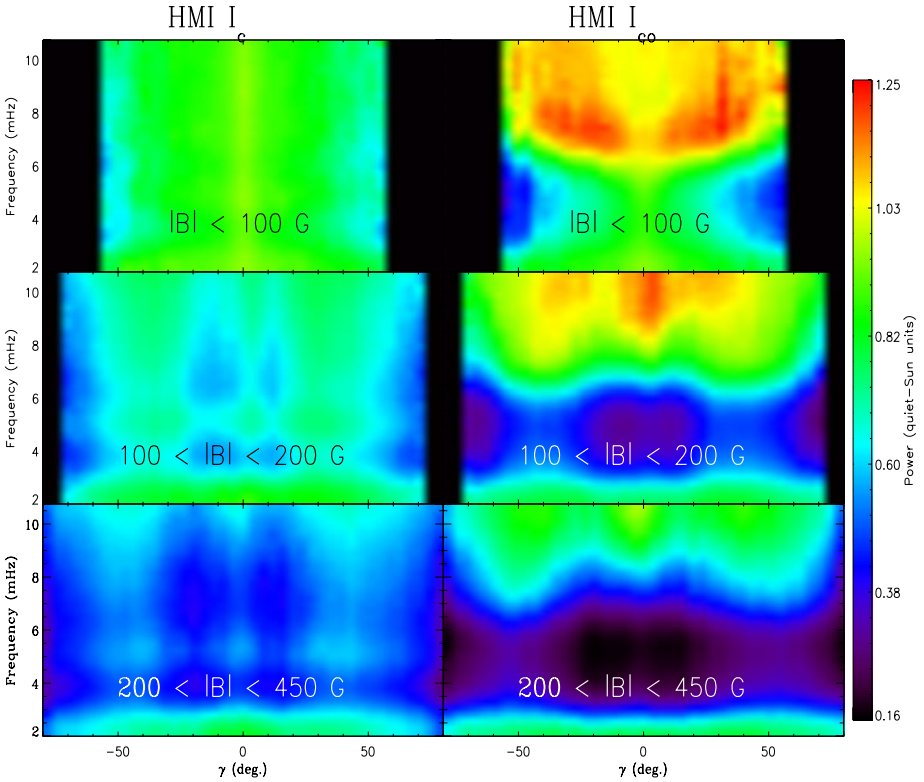


Figure 8 Power, averaged over three different ranges of B (as marked in the panels), as a function of γ for all the four active regions combined. Left column is from observable HMI I_c and the right from HMI I_{co}.

ii) If the Doppler v indeed captures waves from around the $B = 100-200$ G, $\beta = 1$ layer [note that the estimate of 140 km height refers only to the peak position of the response function for v for the HMI line (Fleck, Couvidat, and Straus, 2011), contributions to which come from layers as high as 300 km (Norton *et al.*, 2006)], then the reduction in power could arise directly from the small or vanishing vertical (which is close to the LOS for the regions studied here) velocities due to horizontal propagation in the refraction region of the fast wave.

Increasing fast-mode speed with height leads to eventual reflection toward the lower-height stronger-field base of the canopy causing the compact halos immediately surrounding the spot (bottom panel of Figure 2b). If the waves are mainly compressive, then intensity observations should not exhibit such reductions. What we show and discuss below using power maps from I_{co}, which forms closest in height to v, and from I_{uv1} and I_{uv2} (see Section 3.2) appears to support scenario ii).

The above-observed features may agree qualitatively with the theory and simulations advanced by Khomenko and Collados (2009) that show magneto-acoustic wave refractions as possible causes of power enhancements, and also agree well with such wave refraction studied by Nutto *et al.* (2012) using 2D MHD simulations. We discuss further these aspects and explanations provided by the theory and simulations referred to above, based on our observed features and results, in Section 4.

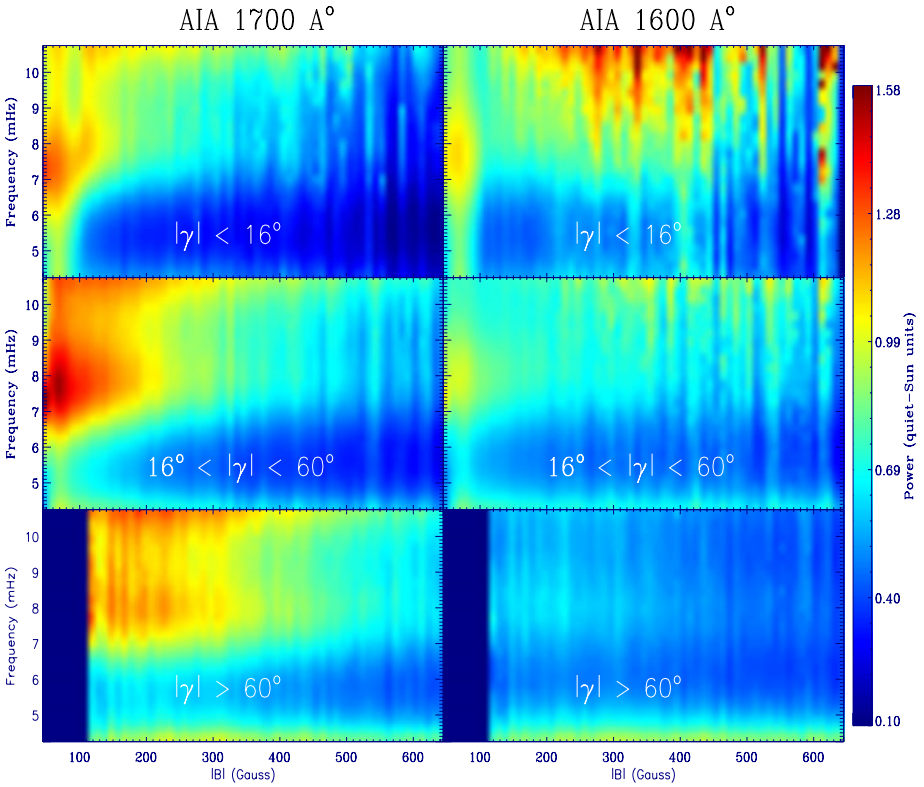


Figure 9 Power, averaged over three different ranges of γ (as marked in the panels), as a function of total magnetic-field strength [B], for all the four active regions combined. Left column is from the observable $I_{\rm uv1}$ and the right from $I_{\rm uv2}$.

Power maps derived from $I_{\rm c}$ and $I_{\rm co}$ are shown in Figures 3b, 7, and 8. Consistent with earlier results, $I_{\rm c}$ does not show any high-frequency power excess. In $I_{\rm co}$ maps, the twin-halo structure, in ν, as a whole shifts towards higher ν (clearer in Figure 8): the halo at lower values of B and ν now centers about 7 mHz (in the range 6 – 8 mHz), while that at higher ν and B peaks beyond 8 mHz. Excess power at 6 mHz is now small, typically less than 10 % (middle panel of Figure 3b). However, both the halos taper down sharply as B increases, with the 6 – 8 mHz halo power decreasing faster than that above 8 mHz, and no excess power is seen beyond about 250 G. It is worth noting that, in contrast to that from v, the 6 – 8 mHz halo peaks at inclined ($16° < |\gamma| < 50°$) lower B (< 100 G) region; in addition, the excess power migrates to higher frequencies as γ (in the above range) increases for weak fields (the wine-glass like structure in the top-right panel of Figure 8). Interestingly, at those $\nu - B$ locations where a reduction in power is observed in maps from v, *i.e.* between 100 and 200 G and above about 8 mHz (in Figure 5), we now see bright halos as measured from $I_{\rm co}$ (Figure 3b, bottom right panel, and Figure 7). Although it is formed around 150 km above the mean height for v, it is likely that these high-ν waves seen in $I_{\rm co}$, as well as those seen in $I_{\rm uv1}$ and $I_{\rm uv2}$ from AIA discussed below, are the same ones seen in v. This gives further credence to our earlier inference on wave refraction and attendant horizontal propagation made from v power maps.

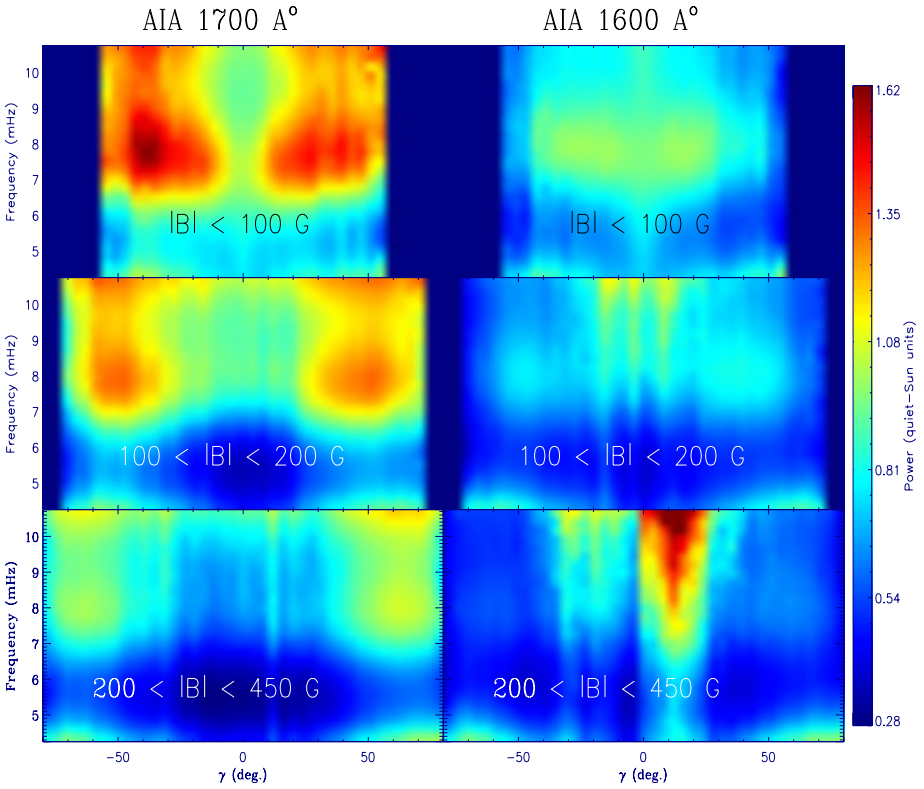


Figure 10 Power, averaged over three different ranges of B (as marked in the panels), as a function of γ for all the four active regions combined. Left column is from observable $I_{\rm uv1}$ and the right from $I_{\rm uv2}$.

3.2. Upper-Photospheric and Lower-Chromospheric Behavior of Power Halos

Power maps estimated from the upper-photospheric and lower-chromospheric UV emissions, $I_{\rm uv1}$ and $I_{\rm uv2}$, are shown in Figures 4, 9, and 10. Note that the network region and plage flux tubes show prominent five-minute oscillations (as seen in the top two panels of Figure 4), and these oscillations extend further down to about 1.75 mHz in ν, forming the so-called long-period network oscillations, known from the early 1990s (Lites, Rutten, and Kalkofen, 1993). We do not discuss or analyse their properties here. Similar to that seen from $I_{\rm co}$, high-frequency power maps from $I_{\rm uv1}$ have the whole ν-pattern of twin halos shifting upwards in ν, with the lower end at about 6 mHz exhibiting a weak, but much outward extended, halos surrounding sunspots and strong flux elements. The ν-extent of the weaker-B halo (see Figures 9 and 10) is now broader than that seen in $I_{\rm co}$, covering the range of 6 – 9 mHz, with the peak excess power about 8 mHz. However, $I_{\rm uv2}$ does not show the secondary $\nu > 9$ mHz halo seen both in $I_{\rm co}$ and $I_{\rm uv1}$; the apparently noisy and large-amplitude excess seen over high-B (300 G and above) in Figures 9 and 10 is due to flare-activity-induced enhancements, especially in NOAA 11161, seen to fall over the sunspots and nearby strong-field pixels (as seen in Figure 4b, middle and bottom panels). The overall properties of the 6 – 9 mHz halo seen in $I_{\rm uv1}$ and $I_{\rm uv2}$ are in agreement with several earlier results on acoustic halos in the chromosphere, using the Ca II K emissions (Braun *et al.*, 1992; Toner and LaBonte, 1993; Ladenkov *et al.*, 2002) and Na D_2 line (5890 Å) and K D_1 line

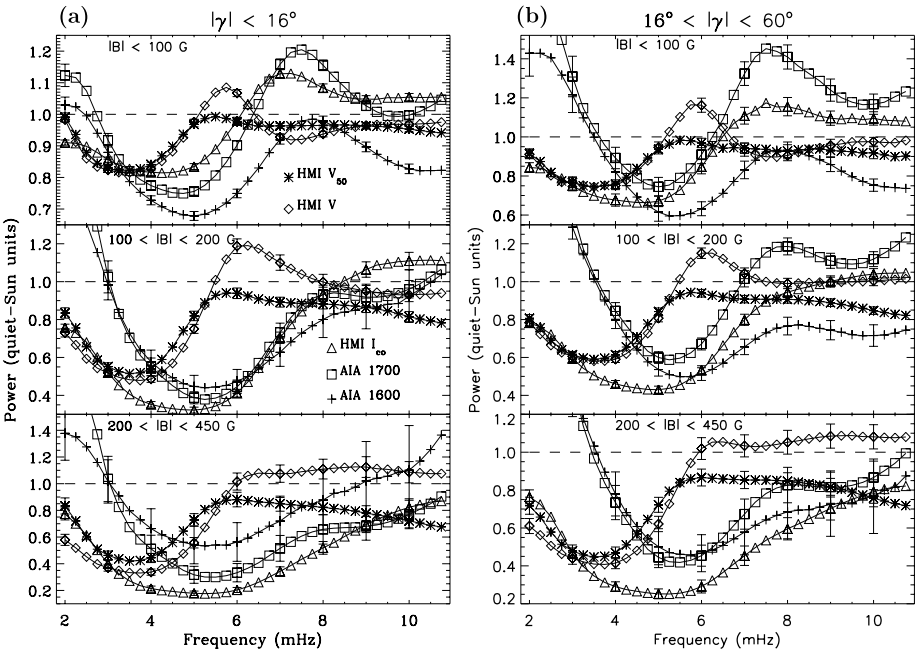


Figure 11 Frequency [ν] variation of Figures 6, 8, and 10 averaged over $\gamma < 16°$ (panel a) and $16° < \gamma < 60°$ (panel b). The different symbols correspond to different observables used in this study, and are marked in panels (a) and (b).

(7700 Å) observations (Moretti *et al.*, 2007). The 1700-Å emissions, $I_{\rm uv1}$, exhibit the most prominent and largest amplitude excess in the 6 – 9 mHz range.

The close proximity, in height of formation, of $I_{\rm co}$ and $I_{\rm uv1}$ is reflected clearly in the very similar $B-\nu$ and $\gamma-\nu$ variations seen in the right panels of Figures 7 and 8 and the left panels of Figures 9 and 10. In agreement with our earlier inference based on results from v, power over inclined fields ($20° < |\gamma| < 60°$) measured from $I_{\rm uv1}$ (Figure 10, left panel) corresponds to small-scale strong-flux concentrations outside of sunspots, and the spatial maps (Figure 4) make this enhanced high-ν power over such regions very obvious. Overall, it is clear that the spatial extent of power halos and their frequency extent increase as a function of height in the atmosphere. This is consistent with that expected from expanding field lines and hence upwardly rising locations of $\beta = 1$ layer around sunspots and other flux structures. The slow migration towards higher ν of the power halos in height in the atmosphere may relate to the height variation of the acoustic cut-off frequency and preferential conversions of waves of such frequency around the $\beta = 1$ layer.

3.3. Comparisons of Frequency Variation of Power over Height

A comparative picture of the variation of power *versus* ν, as well as observation height, is better seen by plotting the vertical cross sections of Figures 5 – 10 averaged over representative ranges of B and γ. This is done in Figure 11, for $|\gamma| < 16°$ (panel a) and $16° < |\gamma| < 60°$ (panel b). The error bars plotted are standard errors estimated assuming that each measurement over all pixels falling within the bin sizes used for B and γ captures an independent realization of the same wave process determined by the values taken by the above physical

quantities over the set of pixels and hence that the scatter within such bins is random. For clarity, we have plotted error bars only at every 1 mHz.

The more or less uniform values of power, with values close to one without any halo, above the acoustic cut-off of 5.3 mHz, in the measurements from v_{50} corresponding to the lowest heights (plotted as connected asterisks in Figure 11), especially over weak field regions ($B < 100$ G) is clear. The migration of peak power, for the observable v (plotted as connected diamonds), to higher ν as B increases is obvious. Our new finding of the existence of a spatially compact halo in maps from v, at ν above about 6.5 mHz with quite a broad peak, is seen in the bottom panels (for $200 < B < 450$ G) of Figure 11.

Overall, the largest amplitude power excess is observed from $I_{\rm uv1}$ over inclined-field regions ($16^\circ < \gamma < 60^\circ$) peaking at $\nu \approx 7.5$ mHz, and this is due to the fact that, at these heights, all magnetic structures, small and large, are surrounded by bright halos. It is also to be noted that power-halo amplitudes, as well as their ν variation, are very similar for $I_{\rm co}$ and $I_{\rm uv1}$ (squares and triangles in Figure 11, respectively), reinforcing our overall inference that variation of magnetic field (or $\beta = 1$ layer) over height in the atmosphere controls the wave physics behind the power excess.

3.4. Wavenumber Dependence of Power Spectra over Height

It is well known that the solar oscillation power spectra exhibit ridge structure well beyond the photospheric cut-off frequency of ≈ 5.3 mHz (Libbrecht, 1988), and that it is not due to the resonant p-modes as in the interior but are understood to be due to an atmospheric interference between waves directly from the sub-surface sources and those refracting back from the solar interior (Kumar *et al.*, 1990; Kumar and Lu, 1991). The high-frequency power halos are reported to introduce slight shifts in the locations of ridges in the power spectra (Schunker and Braun, 2011): at a given frequency, active regions shift the wavenumbers of high-frequency waves to slightly higher values as compared to quiet Sun; higher wavenumber for a given frequency means that the peak-power locations as a function of frequency (*i.e.* the ridges) shift downwards and hence reduced frequencies. Here, we check if changes occur for the same active region as a function of observation height (*i.e.* as a function of different observables used here).

We azimuthally average the three-dimensional (k_x, k_y, ν) power spectra at each ν to make $k_{\rm h} - \nu$ spectra. We show the results for one active region, NOAA 11092, in Figure 12. We produce the $k_{\rm h}$ variation of the spectra at two frequencies: 3.27 and 6 mHz. The spectra estimated from all intensities, $I_{\rm c}$, $I_{\rm co}$, $I_{\rm uv1}$, and $I_{\rm uv2}$ are grouped together (top panels), and those from the velocities v_{50} and v are shown in the bottom panels of Figure 12. Vertical-dashed lines across both set of panels mark locations of lowest radial order p-modes as determined by the peak locations of power in v (dotted lines in the bottom panels). It is clear that at $\nu = 3.27$ mHz, there is no change in the wavenumber dependence as a function of height (Figure 12a) and this is expected from the overall evanescent nature of such frequency waves. At $\nu = 6$ mHz (Figure 12b), the major change is a shift towards lower $k_{\rm h}$ of the locations of power peaks obtained from $I_{\rm c}$ than the rest of the observables. This change in power spectra corresponds to the well-observed positive shifts, *i.e.* increase in frequencies, related to the "correlated noise" mechanism that explains the reversal of asymmetry in p-mode power spectra from intensities in relation to that from velocities (Nigam *et al.*, 1998; Kumar and Basu, 1999). These increased frequencies observed in $I_{\rm c}$ correspond to the decrease in $k_{\rm h}$ seen in Figure 12b, and hence could not be due to the absence of power halos in $I_{\rm c}$. However, the other intensities $I_{\rm co}$, $I_{\rm uv1}$, and $I_{\rm uv2}$, which are from progressively increasing heights in the atmosphere, do not show these shifts and have power peaks at almost

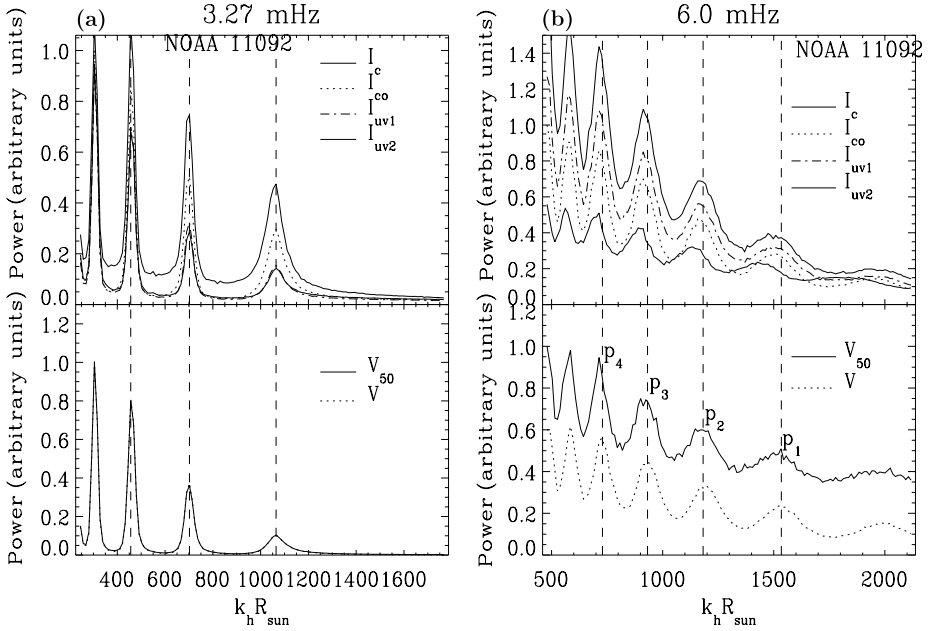


Figure 12 Modal power spectra *versus* $k_{\rm h}$ at two different values of ν for the different observables used in this study (different line styles correspond to different observables as marked within each panel): (a) for $\nu = 3.27$ mHz, and (b) for $\nu = 6$ mHz. Top-row panels correspond to spectra from intensities and the bottom panels to those from velocities. See text for further explanations.

the same positions as those from v. This again possibly relates to the decreasing correlated noise as distances from the acoustic sources increase for these intensities from higher layers. Hence, the above changes in power-peak locations are possibly not due to the power halos *per se*. Additionally, comparing v_{50} and v power spectra in Figure 12b (bottom panel), the increasing high-frequency power halo in v does not appear to leave any significant changes in the power spectra, although there are slight shifts noticeable in the locations of p_3- and p_4-ridges. However, such slight changes, with respect to v, are also seen in higher atmospheric intensities (top panel of Figure 12b). Thus, we can conclude that there are no clear associations of variations in high-frequency power excess *versus* height with any significant changes in the power-peak locations. The detected changes, in comparison with non-magnetic regions, as reported by Schunker and Braun (2011) are of larger magnitude than any changes caused by variations in power halo *versus* height.

A thorough analysis of differences in the spectral power-peak locations and asymmetries in power profiles, from among those obtained from intensities at different heights, and their relationship to the physical mechanisms controlled by correlated noise and acoustic source depths (Nigam *et al.*, 1998; Kumar and Basu, 1999), is beyond the scope of the work reported here, and we focus on this in a separate study, to be reported in the near future.

4. Discussion and Summary

The physics of interactions between acoustic waves and magnetic fields in the solar atmosphere leads to a rich variety of observable dynamical phenomena, the understanding of

which is crucial to mapping the thermal and magnetic structuring in height of the solar atmosphere. A large body of theoretical and observational studies of such physics exists and a significant fraction of which, as explained in Section 1, is relevant for the detailed analyses that we have made here of the high-frequency power halos around active regions. The spatial reorganization of otherwise uniform injection of acoustic waves of frequency above the photospheric cut-off of ≈ 5.3 mHz from the sub-surface and photospheric layers by the overlying and arching magnetic field of active regions and strong-field flux elements is clearly brought out in our analyses of the wave-power distribution estimated from observables spanning a height range of 0 – 430 km above the photosphere. In accordance with the central theme of theoretical investigations referred to in Section 1 (*e.g.* Rosenthal *et al.*, 2002; Bogdan *et al.*, 2003; Khomenko and Collados, 2009, see also the review by Khomenko, 2009), our results presented and discussed in detail in Section 3 clearly bring out the importance of interactions between fast-acoustic waves from the lower atmospheric high-β regions and the expanding and spreading magnetic-field canopy that separates the low-β region above. Although we have not modeled the higher atmospheric magnetic field from the observed vector field at photospheric heights and have not estimated and mapped the $\beta = 1$ layer, results obtained from an analysis of photospheric B- and γ-dependence of power maps, especially for the newly identified high-frequency secondary power halo peaking at about $\nu = 8$ mHz (Figures 2(b), 5, and 6), appear to agree with the theory and simulations performed by Khomenko and Collados (2009): refraction and subsequent reflection of fast magneto-acoustic waves around the $\beta = 1$ layer as agents of additional wave-energy deposition at these layers and hence the power excess. Such refracted fast magneto-acoustic waves, possibly depending on the angles of incidence of acoustic waves prior to conversion, may converge towards the stronger-field base region that fans out from the sunspots and other strong-field structures, and cause a focussing effect thereby increasing wave amplitudes and thus the compact halos seen in maps from v (Figure 2(b)). The reduction in power encircling the compact halos itself is again identified as a signature wave refraction under two possible scenarios:

i) the refraction and horizontal propagation region (*i.e.* the intermediate 100 – 200 G field $\beta = 1$ layer) being located at a height significantly above the v formation height (of 140 km), and/or
ii) vanishing velocity signals in the vertical direction due to horizontal propagation around the observation height (for v).

The spatially extended and weak halo surrounding this reduced-power region is suggested to result from the field-aligned slow magneto-acoustic wave due to the fast-to-slow-mode conversion. In regions where opposite-polarity canopy fields meet to produce neutral lines, it is conceivable that both the refracted fast waves and converted slow waves moving in opposite directions toward the neutral lines dump or focus wave energy causing greater power seen at these locations (Schunker and Braun, 2011).

Other major results of this study can be summarized as follows:

i) The visibility of enhanced power is a strong function of height in the atmosphere; absence of power excess in continuum intensity [I_c] is due to the corresponding height being significantly below the wave-mode conversions happening around the $\beta = 1$ layers, and hence not due to any incompressive wave; this is confirmed by the presence of strong halos in maps from line-core intensity I_{co}, which form at higher layers.
ii) The well-observed 6-mHz halo is the strongest in maps from Doppler velocities [v] forming at about 150 km above the photosphere, and it spreads out (spatially) and gets weaker with height as seen from AIA 1700- and 1600-Å emissions; this feature reflects

the spreading and weakening magnetic field (and hence increasing height for $\beta = 1$ locations) with height.

iii) Frequencies of peak power gradually shift to higher values, along with a spreading in frequency extent, as height increases in the atmosphere; in the upper photosphere, power halos (from v, $I_{\rm co}$, and $I_{\rm uv1}$) exhibit twin peaks, one centered around 6 mHz and the other around 8 mHz.

iv) On the whole, the largest power excess is seen over horizontal magnetic-field locations (as inferred from the photospheric field) for each observable (or height) and the largest among these are seen from the $I_{\rm uv1}$ at about 7.5 mHz.

v) There are no significant changes in peak positions (in wavenumber $k_{\rm h}$) of ridges in power spectra due to height-dependent changes in high-frequency power excess.

Overall, it is clear that the upper-photospheric and lower-chromospheric regions covered by the magnetic canopy spatially redistribute incoming high-frequency acoustic-wave energy from below into a mixture of slow and fast magneto-acoustic waves, through mode conversions around the $\beta = 1$ layer, so as to cause enhanced power around photospheric strong fields. The B and γ dependences of power halos brought out in Figures 5 – 11, possibly include more intricate signatures of above wave interactions, which we have not been able to discern from our current analyses. As mentioned earlier, magnetic-field extrapolations above the photospheric layers and models of atmospheric structure need to be combined to estimate the height variation of β, and sound and Alfvén speeds, which in turn would lead to a clear and unambiguous identification of physical mechanisms behind the power halos. We intend to follow up the present work with such attempts.

Acknowledgements S. Couvidat, K. Hayashi, and Xudong Sun are supported by NASA grants NNG05GH14G to the SDO/HMI project at Stanford University. The data used here are courtesy of NASA/SDO and the HMI and AIA science teams. Data intensive computations performed in this work made use of the High Performance Computational facility of the Indian Institute of Astrophysics, Bangalore. Discussions with C.R. Sangeetha (IIA, Bangalore) are acknowledged.

References

Bogdan, T.J., Carlsson, M., Hansteen, V., McMurry, A., Rosenthal, C.S., Johnson, M., Petty-Powell, S., Zita, E.J., Stein, R.F., McIntosh, S.W., Nordlund, A.: 2003, *Astrophys. J.* **599**, 626. ADS:2003ApJ...599..626B, doi:10.1086/378512.

Bogdan, T.J., Judge, P.G.: 2006, *Roy. Soc. Phil. Trans. A* **364**, 313. ADS:2006RSPTA.364..313B, doi:10.1098/rsta.2005.1701.

Borrero, J.H., Tomczyk, S., Kubo, M., Socas-Navarro, H., Schou, J., Couvidat, S., Bogart, R.: 2011, *Solar Phys.* **273**, 267. ADS:2011SoPh..273..267B, doi:10.1007/s11207-010-9515-6.

Braun, D.C., Lindsey, C., Fan, Y., Jefferies, S.M.: 1992, *Astrophys. J.* **392**, 739. ADS:1992ApJ...392..739B, doi:10.1086/171477.

Brown, T.M., Bogdan, T.J., Lites, B.W., Thomas, J.H.: 1992, *Astrophys. J. Lett.* **394**, L65. ADS:1992ApJ...394L..65B, doi:10.1086/186473.

Cally, P.S.: 2006, *Roy. Soc. Phil. Trans. A* **364**, 333. ADS:2006RSPTA.364..333C, doi:10.1098/rsta.2005.1702.

Finsterle, W., Jefferies, S.M., Cacciani, A., Rapex, P., McInctosh, S.W.: 2004, *Astrophys. J.* **613**, L185. ADS:2004ApJ...613L.185F, doi:10.1086/424996.

Fleck, B., Couvidat, S., Straus, T.: 2011, *Solar Phys.* **271**, 27. ADS:2011SoPh..271...27F, doi:10.1007/s11207-011-9783-9.

Fossum, A., Carlsson, M.: 2005, *Astrophys. J.* **625**, 556. ADS:2005ApJ...625..556F, doi:10.1086/429614.

Hanasoge, S.M.: 2008, *Astrophys. J.* **680**, 1457. ADS:2008ApJ...680.1457H, doi:10.1086/587934.

Hanasoge, S.M.: 2009, *Astron. Astrophys.* **503**, 595. ADS:2009A&A...503..595H, doi:10.1051/0004-6361/200912449.

Hill, F., Jain, K., Tripathy, S., Kholikov, S., Gonzalez Hernandez, I., Leibacher, J., Howe, R., Baudin, F., Carlsson, M., Chaplin, W., Tarbell, T.: 2011, *Bull. Am. Astron. Soc.* **43**, 21.11.

Hindman, B.W., Brown, T.M.: 1998, *Astrophys. J.* **504**, 1029. ADS:1998ApJ...504.1029H, doi:10.1086/306128.

Hoeksema, J.T., Liu, Y., Hayashi, K., et al.: 2013. Solar Phys., in preparation.

Howe, R., Komm, R., Gonzalez Hernand, I., Jain, K., Hill, F., Haber, D.A., Bogart, R.: 2010, AGU Fall Meeting. Abstract No. SH11A-1601.

Jacoutot, L., Kosovichev, A.G., Wray, A., Mansour, N.N.: 2008, *Astrophys. J.* **684**, L51. ADS:2008ApJ...684L..51J, doi:10.1086/592042.

Jain, R., Haber, D.: 2002, *Astron. Astrophys.* **387**, 1092. ADS:2002A&A...387.1092J, doi:10.1051/0004-6361:20020310.

Khomenko, E.: 2009, In: Dikpati, M., Arentoft, T., Gonzalez Hernandez, I., Lindsey, C., Hill, F. (eds.) *Solar-Stellar Dynamos as Revealed by Helio- and Asteroseismology*, *ASP Conference Series* **CS-416**, Astron. Soc. Pacific, San Francisco, 31. ADS:2009ASPC..416...31K.

Khomenko, E., Cally, P.S.: 2012, *Astrophys. J.* **746**, 68. ADS:2012ApJ...746...68K, doi:10.1088/0004-637X/746/1/68.

Khomenko, E., Collados, M.: 2006, *Astrophys. J.* **653**, 739. ADS:2006ApJ...653..739K, doi:10.1086/507760.

Khomenko, E., Collados, M.: 2008, *Astrophys. J.* **689**, 1379. ADS:2008ApJ...689.1379K, doi:10.1086/592681.

Khomenko, E., Kosovichev, A.G., Collados, M., Parchevsky, K., Olshevsky, V.: 2009, *Astrophys. J.* **694**, 411. ADS:2009ApJ...694..411K, doi:10.1088/0004-637X/694/1/411.

Khomenko, E., Collados, M.: 2009, *Astron. Astrophys.* **506**, L5. ADS:2009A&A...506L...5K, doi:10.1051/0004-6361/200913030.

Kumar, P., Duvall, T.L. Jr., Harvey, J.W., Jefferies, S.M., Pomerantz, M.A., Thompson, M.J.: 1990, In: Osaki, Y., Shibahashi, H. (eds.) *Proc. Oji. Internat. Seminar*, Springer, Berlin, 87.

Kumar, P., Basu, S.: 1999, *Astrophys. J.* **519**, 389. ADS:1999ApJ...519..389K, doi:10.1086/307368.

Kumar, P., Lu, E.: 1991, *Astrophys. J. Lett.* **375**, L35. ADS:1991ApJ...375L..35K, doi:10.1086/186082.

Kuridze, D., Zaqarashvili, T.V., Shergelashvili, B.M., Poedts, S.: 2008, *Ann. Geophys.* **26**, 2983. ADS:2008AnGeo..26.2983K, doi:10.5194/angeo-26-2983-2008.

Ladenkov, O.V., Hill, F., Egamberdiev, S.A., Chou, D.Y.: 2002, *Astron. Lett.* **28**, 411. ADS:2002AstL...28..411L, doi:10.1134/1.1484142.

Leka, K.D., Barnes, G., Crouch, A.D., Metcalf, T.R., Gary, G.A., Jing, J., Liu, Y.: 2009, *Solar Phys.* **260**, 83. ADS:2009SoPh..260...83L, doi:10.1007/s11207-009-9440-8.

Lemen, J.R., Title, A.M., Akin, D.J., Boerner, P.F., Chou, C., Drake, J.F., Duncan, D.W., *et al.*: 2012, *Solar Phys.* **275**, 17. ADS:2012SoPh..275...17L, doi:10.1007/s11207-011-9776-8.

Libbrecht, K.: 1988, *Astrophys. J.* **334**, 510. ADS:1988ApJ...334..510L. doi:10.1086/166855.

Lites, B.W., Rutten, R.J., Kalkofen, W.: 1993, *Astrophys. J.* **414**, 345. ADS:1993ApJ...414..345L, doi:10.1086/173081.

Metcalf, T.R.: 1994, *Solar Phys.* **155**, 235. ADS:1994SoPh..155..235M, doi:10.1007/BF00680593.

Metcalf, T.R., Leka, K.D., Barnes, G., Lites, B.W., Georgoulis, M.K., Pevtsov, A.A., Balasubramaniam, K.S., *et al.*: 2006, *Solar Phys.* **237**, 267. ADS:2006SoPh..237..267M, doi:10.1007/s11207-006-0170-x.

Moretti, P.F., Jefferies, S.M., Armstrong, J.D., McIntosh, S.W.: 2007, *Astron. Astrophys.* **471**, 961. ADS:2007A&A...471..961M, doi:10.1051/0004-6361:20077247.

Nagashima, K., Sekii, T., Kosovichev, A.G., *et al.*: 2007, *Publ. Astron. Soc. Japan* **59**, S631. ADS:2007PASJ...59S.631N.

Nigam, R., Kosovichev, A.G., Scherrer, P.H., Schou, J.: 1998, *Astrophys. J.* **495**, L115. ADS:1998ApJ...495L.115N, doi:10.1086/311219.

Norton, A.A., Graham, J.P., Ulrich, R.K., Schou, J., Tomczyk, S., Liu, Y., Lites, B.W., Lopez Ariste, A., Bush, R.I., Socas-Navarro, H., Scherrer, P.H.: 2006, *Solar Phys.* **239**, 69. ADS:2006SoPh..239...69N, doi:10.1007/s11207-006-0279-y.

Nutto, C., Steiner, O., Schaffenberger, W., Roth, M.: 2012, *Astron. Astrophys.* **538**, 79. ADS:2012A&A...538A..79N, doi:10.1051/0004-6361/201016083.

Rajaguru, S.P., Sankarasubramanian, K., Wachter, R., Scherrer, P.H.: 2007, *Astrophys. J. Lett.* **654**, L175. ADS:2007ApJ...654L.175R, doi:10.1086/511266.

Rajaguru, S.P., Wachter, R., Sankarasubramanian, K., Couvidat, S.: 2010, *Astrophys. J.* **721**, L86. ADS:2010ApJ...721L..86R, doi:10.1088/2041-8205/721/2/L86.

Rosenthal, C.S., Bogdan, T.J., Carlsson, M., Dorch, S.B.F., Hansteen, V., McIntosh, S.W., McMurry, A., Nordlund, A., Stein, R.F.: 2002, *Astrophys. J.* **564**, 508. ADS:2002ApJ...564..508R, doi:10.1086/324214.

Scherrer, P.H., Schou, J., Bush, R.I., Kosovichev, A.G., Bogart, R.S., Hoeksema, J.T., *et al.*: 2012, *Solar Phys.* **275**, 207. ADS:2012SoPh..275..207S, doi:10.1007/s11207-011-9834-2.

Schunker, H., Braun, D.C.: 2011, *Solar Phys.* **268**, 349. ADS:2011SoPh..268..349S, doi:10.1007/s11207-010-9550-3.
Schunker, H., Cally, P.S.: 2006, *Mon. Not. Roy. Astron. Soc.* **372**, 551. ADS:2006MNRAS.372..551S. doi:10.1111/j.1365-2966.2006.10855.x.
Shelyag, S., Zharkov, S., Fedun, V., Erdlyi, R., Thompson, M.J.: 2009, *Astron. Astrophys.* **501**, 735. ADS:2009A&A...501..735S, doi:10.1051/0004-6361/200911709.
Thomas, J.H., Stanchfield, D.C.H.: 2000, *Astrophys. J.* **537**, 1086. ADS:2000ApJ...537.1086T, doi:10.1086/309046.
Toner, C.G., LaBonte, B.J.: 1993, *Astrophys. J.* **415**, 847. ADS:1993ApJ...415..847T, doi:10.1086/173206.
Vigeesh, G., Hasan, S.S., Steiner, O.: 2009, *Astron. Astrophys.* **508**, 951. ADS:2009A&A...508..951V, doi:10.1051/0004-6361/200912450.

DOI 10.1007/978-1-4899-8005-2_8
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0075-9

Global-Oscillation Eigenfunction Measurements of Solar Meridional Flow

M. Woodard · J. Schou · A.C. Birch · T.P. Larson

Received: 1 March 2012 / Accepted: 12 July 2012 / Published online: 25 July 2012

Abstract We describe and apply a new helioseismic method for measuring solar subsurface axisymmetric meridional and zonal flow. The method is based on a theoretical model of the response of global-oscillation eigenfunctions to the flow velocity and uses cross spectra of the time-varying coefficients in the spherical-harmonic expansion of the photospheric Doppler-velocity field. Eigenfunction changes modify the leakage matrix, which describes the sensitivity of the spherical-harmonic coefficients to the global-oscillation modes. The form of the leakage matrix in turn affects the theoretically expected spherical-harmonic cross spectra. Estimates of internal meridional and zonal flow were obtained by fitting the theoretical flow-dependent cross spectra to spherical-harmonic cross spectra computed from approximately 500 days of full-disk Dopplergrams from the *Helioseismic and Magnetic Imager* (HMI) on the SDO spacecraft. The zonal-flow measurements, parameterized in the form of "a" coefficients, substantially agree with measurements obtained from conventional global-mode-frequency analysis. The meridional-flow estimates, in the form of depth-weighted averages of the flow velocity, are similar to estimates obtained from earlier analyses, for oscillation modes that penetrate the outermost one-third of the convection zone. For more deeply penetrating modes, the inferred flow velocity increases significantly with penetration depth, indicating the need for either a modification of the simple conveyor-belt picture of meridional flow or improvement in the cross-spectral model.

Keywords Solar interior · Mass flows · Helioseismology · Solar dynamo

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

M. Woodard (✉) · A.C. Birch
NorthWest Research Associates, Inc., 3380 Mitchell Lane, Boulder, CO 80301-5410, USA
e-mail: mfw@cora.nwra.com

J. Schou · T.P. Larson
W. W. Hansen Experimental Physics Laboratory, Stanford University, Stanford, CA 94305-4085, USA

A.C. Birch
Max-Planck-Institut für Sonnensystemforschung, 37191 Katlenburg-Lindau, Germany

1. Introduction

Meridional flows serve to transport fluid, energy, angular momentum, and magnetic flux within the interiors of stars. Such flows are seen in purely hydrodynamic simulations of solar convection (*e.g.* Miesch *et al.*, 2000). In the outer portion of the solar convection zone, meridional flow is observed as a motion of material from the Equator to the Poles (Duvall, 1979). Like the dominant zonal-flow pattern (differential rotation), meridional flow is also thought to play an important role in the magnetic dynamo, in spite of its relatively small, 10 to 20 $\mathrm{m\,s^{-1}}$, velocity amplitude. Conservation of mass demands that at some, yet undetermined, depth the poleward flow of mass must reverse and become an equatorward flow. This "conveyor belt" picture plays a key role in so-called flux-transport scenarios of the magnetic cycle (Sheeley, 2005). In the model of Dikpati and Gilman (2006), the conveyor belt extends to the rotational shear layer (tachocline) near the bottom of the convection zone, where the dynamo action occurs.

Meridional flow has been inferred from direct Doppler measurements of the solar photosphere (Duvall, 1979; Labonte and Howard, 1982), from correlation-tracking and related studies of supergranular-scale convection (Hathaway, 2011; Gizon, Duvall, and Schou, 2003a, 2003b; Schou, 2003), and from helioseismic analysis, beginning with the work of Giles *et al.* (1997). Meridional motions of photospheric magnetic features have also been studied, although the rates of motion of small, large, compact, and extended features differ widely (*e.g.* Sivaraman *et al.*, 2010). Studies of meridional motion have been carried out using a number of helioseismic methods. These include studies which use ring analysis (Schou and Bogart, 1998; González Hernández *et al.*, 1999; Haber *et al.*, 2002), time–distance analysis (Giles, Duvall, and Scherrer, 1998; Giles, 2000; Hughes and Thompson, 2003; Zhao and Kosovichev, 2004; Chou and Ladenkov, 2005; Duvall and Hanasoge, 2009), Fourier–Hankel analysis (Braun and Fan, 1998, 1999), and global cross-spectral analysis (Woodard, 2009). These methods have been applied to time series of Doppler images collected since the mid-1990s by the ground-based *Global Oscillation Network Group* (Harvey *et al.*, 1996) and by the *Michelson Doppler Imager* (MDI: Scherrer *et al.*, 1995) on the ESA-NASA *Solar and Heliospheric Observatory*.

Because p- and f-mode oscillations are much more sensitive to conditions in the outer layers of the Sun than in the deep interior, we have a much clearer picture of the flows in the outermost convection zone than in the tachocline. While most of the seismic studies have concentrated on the outermost convection zone, serious efforts have been made to probe deeper layers. In particular, Giles (2000) performed inversions of helioseismic travel-time data, derived from two years of MDI medium-ℓ Doppler images, and found the amplitude of meridional flow to decrease gradually down to the bottom of the convection zone. A counterflow emerged from the analysis, but only when a mass-conserving constraint was imposed. More recently, Duvall and Hanasoge (2009), using about six years of medium-ℓ data, measured travel times sensitive to meridional and zonal flow and compared them with theoretical expectations. The results of the Duvall and Hanasoge analysis for the deep convection zone are hard to reconcile with those of Giles, as they suggest a meridional-flow velocity increasing with depth. Thus seismic data have yet to clarify the issue of the return flow. Recent estimates of subsurface flow based on the correlation tracking of supergranular cells (Hathaway, 2011) suggest a meridional conveyor belt with a counterflow about 35 Mm below the photosphere, perhaps indicating a shallow dynamo. Studies based on the motion of magnetic elements (*e.g.* Sivaraman *et al.*, 2010) also suggest a shallow meridional flow. However, this picture is at odds with the results of other supergranulation studies, as well as the seismic studies.

In this article we explore a new approach to measuring subsurface meridional and zonal flow that considers the effect of the internal velocity field on the eigenfunctions of the Sun's normal oscillation modes. Changes in the eigenfunctions affect the leakage of oscillation signal into spherical-harmonic components of the observed velocity field. The theoretical effect of a flow on oscillation data can therefore be computed by an extension of the traditional leakage-matrix formalism developed to describe the non-uniform spatial sampling of helioseismic observations. The theoretical model is applied to cross spectra of spherical-harmonic time series of full-disk Dopplergrams from the *Helioseismic and Magnetic Imager* (HMI: Schou *et al.*, 2012) on the NASA *Solar Dynamics Observatory*.

2. Theoretical Basis for Large-Scale Helioseismic Flow Measurements

In this section we describe a theoretical framework that underlies a "leakage-matrix" approach to measuring large-scale flow with global-oscillation data. The data of interest are the spherical-harmonic and temporal-frequency amplitudes $[\varphi^{\ell m}(\omega)]$ of the line-of-sight component of the time-varying photospheric velocity field derived from long sequences of Doppler images. The indices ℓ and m denote spherical-harmonic degree and azimuthal order, and ω is angular frequency. Our framework directly extends that of Schou (1992, hereafter S92), which dealt mainly with the problem of estimating parameters – the eigenfrequencies, linewidths, and amplitudes – which characterize the spectral profiles of individual oscillation modes in these data. Global-oscillation data also contain information about mode eigenfunctions, which are influenced by flows and aspherical structure.

A number of approaches have been developed to model the effect of subsurface flow on solar oscillations. In the current approach, the influence of the flow-velocity field on the form of the solar eigenfunctions plays a leading role. In a Sun with flows, an eigenfunction is a sum of spherical-harmonic components, rather than a single spherical harmonic. The pattern of leakage into different spherical-harmonic components is sensitive to the form of the flow-velocity field.

The likelihood function used by S92 to estimate mode parameters depends on global-helioseismic data through the cross-spectral combinations $\varphi^{\ell' m'}(\omega)\varphi^{\ell m *}(\omega)$ and on solar-model parameters through the statistical expectations $E[\varphi^{\ell' m'}(\omega)\varphi^{\ell m *}(\omega)]$ of these combinations. We use "E" to denote the average over an ensemble of wave fields. The cross-spectral expectations are the components of the data covariance matrices discussed in S92. A key working hypothesis of the local helioseismology of the quiet Sun is that the wave field is a stationary Gaussian random process (Gizon and Birch, 2004). Cross-spectral data have also been used in helioseismic direct-modeling analysis (*e.g.* Woodard, 2009), although flow-dependent eigenfunctions have yet to play an explicit role in this work. The existence of specifically cross-spectral (*i.e.* $\ell', m' \neq \ell, m$) information is a direct consequence of leakage (see S92). Moreover, the complex-valued cross spectra provide information not present in power spectra, namely about the phase of the leakage. In the subsections that immediately follow we develop expressions for the flow dependence of the cross spectra, starting with a discussion of the eigenfunction dependence.

2.1. The Effect of Steady, Axisymmetric Flow on Oscillation Eigenfunctions

The problem of computing theoretical eigenfunctions of global adiabatic oscillations of a Sun with steady axisymmetric meridional and zonal flow has been addressed by a number of investigators (Woodard, 1989, 2000; Vorontsov, 2007, 2011; Gough and Hindman, 2010;

Schad, Timmer, and Roth, 2011). The flow velocity $[\boldsymbol{u}]$ of an axisymmetric flow can be written as a sum of poloidal (meridional) and toroidal (zonal) vector-spherical-harmonic functions $[\mathbf{P}_s(r,\theta,\phi)$ and $\mathbf{T}_s(r,\theta,\phi),\ s=0,1,2,\ldots]$ with

$$\mathbf{P}_s = \hat{\mathbf{r}} u_s(r) Y_s^0(\theta,\phi) + \hat{\boldsymbol{\theta}} v_s(r) \partial_\theta Y_s^0(\theta,\phi) \tag{1}$$

and

$$\mathbf{T}_s = -\hat{\boldsymbol{\phi}} w_s(r) \partial_\theta Y_s^0(\theta,\phi), \tag{2}$$

where Y_s^0 are the scalar spherical harmonics of degree s and azimuthal order 0. The coordinates (r,θ,ϕ) are spherical–polar coordinates, with corresponding unit vectors $(\hat{\mathbf{r}},\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\phi}})$. The functions $u_s(r)$, $v_s(r)$, and $w_s(r)$ are, respectively, the amplitudes at radius r of the r, θ, and ϕ components of the velocity multipole of degree s (Lavely and Ritzwoller, 1992). As in earlier treatments of the effect of these flows on waves, we adopt the constraint that the divergence of the mass flux vanishes throughout the solar interior. For steady axisymmetric flow, the mass-flux constraint (also referred to as an "anelastic" constraint), which amounts to a relation between $u_s(r)$ and $v_s(r)$, is equivalent to the mass continuity equation. In modeling Earth- or near-Earth-based helioseismic data, it is convenient to formulate the problem of wave motion in a uniformly rotating coordinate frame, whose axis coincides with the solar-rotation axis. In our analysis, oscillation and flow velocities, and therefore mode frequencies, will be referred to such a frame rotating at the Earth-orbital frequency of approximately 31.7 nHz.

The functions $w_1(r), w_3(r), w_5(r), \ldots$ correspond to large-scale zonal angular velocity patterns that are symmetrical about the solar Equator and produce observable frequency perturbations whose m-dependence is characterized by the coefficients $a_1, a_3, a_5, \ldots$, respectively (Ritzwoller and Lavely, 1991; Schou, Christensen-Dalsgaard, and Thompson, 1994). The even-s profiles $w_2, w_4, w_6, \ldots$, which describe zonal patterns which are antisymmetric about the Equator, produce no frequency shifts of first order in the angular velocity. However, all of the w_s profiles affect the eigenfunctions, thereby permitting measurements of the odd-parity component of the flow, as well as measurements of the dominant, symmetric component, which are independent of the measurements obtained from mode frequencies. Meridional flow has no first-order effect on eigenfrequencies and the estimated frequency perturbations are probably too small to measure (Roth and Stix, 2008; Chatterjee and Antia, 2009). Meridional flow does, however, affect the eigenfunctions.

Theoretical treatments assume that the mode-labeling scheme of a Sun with an internal velocity with components given by Equations (1) and (2) is independent of the internal velocity. Thus the eigenfunctions and eigenfrequencies can be labeled $(\boldsymbol{\xi}_i,\omega_i) = (\boldsymbol{\xi}_{n\ell m},\omega_{n\ell m})$ as the adiabatic modes of a static, spherical star. We recall that the displacement eigenfunctions $[\tilde{\boldsymbol{\xi}}_i]$ of a static star can be written in the form

$$\tilde{\boldsymbol{\xi}}_i = \tilde{\boldsymbol{\xi}}_{n\ell m}(r,\theta,\phi) = \left[U_{n\ell}(r)\hat{\boldsymbol{r}} + V_{n\ell}(r)\left(\hat{\boldsymbol{\theta}}\partial_\theta + \hat{\boldsymbol{\phi}}\frac{\partial_\phi}{\sin\theta}\right)\right] Y_\ell^m(\theta,\phi), \tag{3}$$

where the radial order $[n]$ characterizes the radial structure of the eigenfunction. The functions $U_{n\ell}(r)$ and $V_{n\ell}(r)$ can be chosen to be real-valued, and therefore the modes are standing wave patterns in r and θ and progressive in ϕ.

Flows and other departures from spherical symmetry affect the solar eigenfunctions in important ways. As the p- and f-mode oscillations of interest to this study are typically observed over a limited range of solar radius near the photosphere, the angular dependence of the eigenfunctions is of greater observational interest than the radial dependence. The axisymmetry of the flows that we are considering implies that the eigenfunctions have a

simple $\exp(\mathrm{i}m\phi)$ dependence on ϕ. Therefore the latitudinal structure of the eigenfunctions is of primary interest to us. (As the basis functions of our temporal Fourier expansions are $\exp(-\mathrm{i}\omega t)$, modes of positive m propagate in the direction of increasing ϕ, also taken as the direction of solar rotation.) To be more specific, the outstanding effect of Sun-like differential rotation alone is to displace the latitudinal nodes of the eigenfunctions, in an accordion-like fashion, such that the latitudinal displacement at a given latitude is approximately equal but opposite in sign for modes of the same n and ℓ, but opposite m. In contrast, Sun-like meridional flow adds a small latitudinal-traveling-wave component to an otherwise latitudinal standing wave field, which in first order depends only on the magnitude of m at fixed n and ℓ. The effect of meridional flow can alternatively be viewed as a distortion of the surfaces of constant azimuthal phase, which in the absence of flows are meridional planes.

As in previous theoretical treatments, we model the effect of flows on the eigenfunctions under the assumption that the eigenfunctions of the actual Sun can be expanded in terms of the eigenfunctions of a static reference model. Thus

$$\boldsymbol{\xi}_i = \sum_j c_i^j \, \tilde{\boldsymbol{\xi}}_j. \tag{4}$$

For this work, we use the expressions given in Woodard (1989, 2000) for the expansion coefficients $[c_i^j]$ pertaining to zonal and meridional flow, respectively. In these treatments, $c_i^j = c_{n\ell m}^{n'\ell' m'} = 0$, unless $(n', m') = (n, m)$. Furthermore, the $c_{n\ell m}^{n\ell' m}$ depend on subsurface flow through the advective matrix elements

$$M_\ell^{\ell'} = -\mathrm{i} \int_\odot \mathrm{d}m \; \tilde{\boldsymbol{\xi}}_{n\ell' m}^* \cdot (\boldsymbol{u} \cdot \nabla \tilde{\boldsymbol{\xi}}_{n\ell m}), \tag{5}$$

where $\mathrm{d}m$ denotes a mass element. In this article, the eigenfunctions are assumed to be normalized

$$\int_\odot \mathrm{d}m \; \tilde{\boldsymbol{\xi}}_{n\ell m}^* \cdot \tilde{\boldsymbol{\xi}}_{n\ell m} = 1. \tag{6}$$

The present study considers oscillation modes of degree up to 300. Since $|\ell' - \ell| \leq s \ll \ell$ for the typical ℓ value of interest, it is convenient to use asymptotic expressions for the advective matrix elements. Following Appendix A of Vorontsov (2011, hereafter V11), we find that

$$M_\ell^{\ell'} = \sum_s a_s^t \, \ell \, P_s^t(m/\ell), \quad t \equiv \ell' - \ell, \tag{7}$$

where the P_s^t are associated Legendre functions and the complex (n, ℓ)-dependent coefficients a_s^t extend the familiar a-coefficient expansion of the mode-frequency splittings (the case $t = 0$) to matrix elements of arbitrary $|t| \ll \ell$. The asymptotic relations (A6) and (A7) of V11 yield expressions for the non-vanishing generalized a-coefficients. For odd $s + t$

$$\mathrm{Re}\big(a_s^t\big) = (-1)^{\frac{s-t-1}{2}} \, \frac{(s-t)!!\,(s+t)!!}{(s+t)!} \left(\frac{2s+1}{4\pi}\right)^{\frac{1}{2}} \left\langle \frac{w_s(r)}{r} \right\rangle_{n\ell}, \tag{8}$$

while for even $s + t$

$$\mathrm{Im}\big(a_s^t\big) = t \, (-1)^{\frac{s-t}{2}} \, \frac{(s-t-1)!!\,(s+t-1)!!}{(s+t)!} \left(\frac{2s+1}{4\pi}\right)^{\frac{1}{2}} \left\langle \frac{v_s(r)}{r} \right\rangle_{n\ell}, \tag{9}$$

where Re and Im denote real and imaginary parts, $\langle q(r) \rangle_{n\ell} \equiv \int_0^{R_\odot} q(r) \rho_0(r) r^2 \, [U_{n\ell}^2 + l(l+1) V_{n\ell}^2] \, \mathrm{d}r$ and $\rho_0(r)$ denotes the mass density profile. Thus, to asymptotic order, the mode eigenfunctions can be parameterized by the a_s^t-coefficients.

The asymptotic sensitivity expressions provide an approximate theoretical basis for inverting the measurable a_s^t parameters for subsurface flow. We have yet to carry out inversions, but we note that the depth sensitivity of the a_s^t for both meridional and zonal flow is proportional to the radial distribution of mode kinetic energy. This circumstance strongly suggests that optimal inversions for the meridional-flow profiles $[v_s(r)]$ have a depth-resolution/noise trade-off similar to that of inversions of the conventional a-coefficients for the zonal profiles $w_s(r)$. The fact that the radial profile $[u_s(r)]$ of the meridional flow does not appear in the asymptotic sensitivity expressions suggests that the mass-flux constraint is essential in obtaining accurate estimates of the radial component of the meridional-flow velocity.

2.2. The Sensitivity of Spherical-Harmonic Data to Oscillation Mode Amplitudes

We make the usual approximation that the observed Doppler velocity is linear in the mode amplitudes, so that

$$\varphi^{\ell m} = \sum_i L_i^{\ell m} a^i, \tag{10}$$

where the amplitudes $[a^i]$ are defined by the expansion

$$\boldsymbol{v} = \sum_i a^i \boldsymbol{\xi}_i, \tag{11}$$

of the oscillatory velocity field $[\boldsymbol{v}]$ in modes of the actual Sun. In these equations, and in what follows, we work in the frequency domain unless otherwise stated. The structure of the leakage matrix $[L_i^{\ell m}]$ depends on both the form of the eigenfunctions and on observational and instrumental effects, the dominant effect owing to our current inability to see more than half the solar surface. We note that Equation (10), which relates the observed signal at a given temporal frequency to the mode amplitude at that frequency, holds strictly for time-independent leakage. The equation thus ignores the effect of the non-uniform motion of seismic instruments in the uniformly rotating frame (largely due to the Earth's orbital eccentricity and inclination to the solar equatorial plane). The equation also ignores the effect of gaps in the observed time series. To obtain Equation (11), we have ignored the difference between displacement and velocity eigenfunctions, which is non-zero in a rotating star. On the basis of Equation (13.26) of Unno *et al.* (1989), we estimate that the difference in question can be neglected for our problem.

It is convenient to write the leakage matrix as the product of a factor describing the sensitivity to the amplitudes of the reference modes and a factor expressing the flow-dependent sensitivity of the reference amplitudes to the amplitudes of the exact eigenfunctions. By analogy with the full leakage relation, the reference leakage relation is

$$\varphi^{\ell m} = \sum_i \tilde{L}_i^{\ell m} \tilde{a}^i, \tag{12}$$

where the reference amplitudes $[\tilde{a}^i]$ are defined by the expansion

$$\boldsymbol{v} = \sum_i \tilde{a}^i \tilde{\boldsymbol{\xi}}_i, \tag{13}$$

analogous to Equation (11), in reference modes. The equivalence of the two expansions for the velocity field implies (using Equation (4))

$$\tilde{a}^i = \sum_j c_j^i a^j, \tag{14}$$

which, together with Equation (12), yields the expression

$$L_i^{\ell m} = \sum_j \tilde{L}_j^{\ell m}\, c_i^j, \tag{15}$$

for the overall leakage matrix.

2.3. Flow Dependence of Cross Spectra

For maximum-likelihood modeling we need to know how the expected spherical-harmonic cross spectra depend on solar-model parameters, including the generalized a-coefficients. Starting from Equation (10) and assuming, as in S92, that the solar modes are excited independently, we obtain the general expression

$$E\left[\varphi^{\ell' m'} \varphi^{\ell m *}\right] = \sum_i L_i^{\ell' m'}\, L_i^{\ell m *}\, E\left[\left|a^i\right|^2\right] + r^{\ell' m', \ell m}, \tag{16}$$

where the term $r^{\ell' m', \ell m}$ has been included to allow for a non-seismic background contribution. With the help of Equation (15) this becomes

$$E\left[\varphi^{\ell' m'} \varphi^{\ell m *}\right] = \sum_{i,j,k} \tilde{L}_j^{\ell' m'}\, \tilde{L}_k^{\ell m *}\, E\left[\left|a^i\right|^2\right] c_i^j\, c_i^{k *} + r^{\ell' m', \ell m}. \tag{17}$$

The cross-spectral model depends non-linearly on the a_3- and a_5-coefficients, through the eigenfunction coefficients, and on other parameters of interest. In the following section, we describe the iterative fitting procedure used to obtain flow parameters from observed cross spectra. The procedure requires the computation of both theoretical cross spectra and their sensitivities to small changes in the fitted parameters. According to the previous equation, cross-spectral sensitivity to small changes in the flow velocity depends on eigenfunction sensitivity. (We ignore the flow dependence of background terms when computing sensitivity.) In the asymptotic approximation, the coefficients c_i^j, and therefore their sensitivities to flow-parameter changes, can be chosen to be real, while those of the relatively weak meridional flow, and their sensitivities, can be taken to be approximately imaginary. This circumstance greatly simplifies the task of distinguishing zonal and meridional flow in seismic data.

3. Analysis and Results

In the previous section we presented a framework for extending the maximum-likelihood method of S92 to measure axisymmetric subsurface flow. The procedure uses cross-spectral input in the "raw" form $\varphi^{\ell' m'} \varphi^{\ell m *}$ obtained from time series of duration sufficient to resolve the frequency profiles of the modes of interest. In this article we describe an analysis of cross-spectral data obtained by averaging many independent cross spectra. The present, preliminary analysis is somewhat simpler and less rigorous than that of S92. The central-limit theorem ensures that the statistics of averaged cross spectra tend toward a Gaussian distribution. For normally distributed data a likelihood analysis amounts to least-squares fitting. An optimal fitting procedure would have to be based on realistic estimates of the covariance matrix of the cross-spectral data (*e.g.* Anderson, Duvall, and Jefferies, 1990; Woodard, 2007). For simplicity, however, we assumed a unit data covariance matrix. The present analysis was based on approximately 500 days of HMI Doppler images.

3.1. Least-Squares Fitting Procedure

The cross-spectral model used for the analysis is summarized in Equation (17). The spectral profile $[E[|a^i|^2]]$ of the mode amplitude is modeled as a Lorentzian characterized by a frequency centroid, a linewidth, and an amplitude determined from HMI "pipeline" analysis. While the pipeline analysis allows for m-dependence, as well as n- and ℓ-dependence, in the mode-frequency centroids, the width and amplitude parameters are constrained to depend only on n and ℓ. The m-dependence of the mode frequencies is parameterized in the usual way in terms of a-coefficients.

The leakage model that we use for the HMI Doppler data is similar to that used by a number of other investigators (Schou and Brown, 1994; Korzennik, Rabello-Soares, and Schou, 2004). The model takes into account both the radial and horizontal components of the solar eigenfunctions. The ratio of the horizontal and vertical components of the eigenfunctions, at the level of the solar atmosphere where oscillations are observed, is computed from a standard asymptotic formula (*e.g.* Korzennik, Rabello-Soares, and Schou, 2004). The model also takes into account the spatial apodization that is applied to the Doppler images to suppress the signal near the solar limb. The observation point is arbitrarily located at the zero of longitude of the spherical–polar coordinate system of the rotating frame. With these conventions, the leakage-matrix elements are real and it follows from the discussion near the end of Section 2 that solar-like meridional flow contributes mainly to the imaginary part of the observed cross spectra while zonal flow contributes to the real part.

In this analysis we estimate the $s = 3$ and $s = 5$ components of zonal flow and the $s = 2$ component of meridional flow. Whereas Woodard (1989, 2000) computed eigenfunction expansion coefficients $[c_i^j]$ for the case of depth-independent $v_s(r)/r$ and $w_s(r)/r$ profiles, Vorontsov (V11) considered general profiles. Perusal of V11 reveals that the eigenfunctions of a given (n, ℓ) multiplet are asymptotically identical to the eigenfunctions of a Sun with depth-independent $v_s(r)/r$ and $w_s(r)/r$ equal to $\langle v_s(r)/r\rangle_{n\ell}$ and $\langle w_s(r)/r\rangle_{n\ell}$. The expressions of Woodard have been adapted for our present work by the appropriate substitution of r-independent profiles by r-averaged profiles in the expressions for eigenfunction expansion coefficients. By Equations (8) and (9), the eigenfunctions of these flows, of a particular multiplet, can be described by three independent parameters, for example $a_3 = a_3^0$, $a_5 = a_5^0$ for zonal flow, and a_2^2 for meridional flow. All other a_s^t-coefficients not proportional to these coefficients by Equations (8) and (9) are assumed to be zero in our analysis. Note that we have made no attempt to measure the radial component of the flow velocity, as the asymptotic expressions for the eigenfunctions do not depend on $u_s(r)$.

The current study was based on the first seven 72-day time series of spherical-harmonic coefficients of HMI Dopplergrams. Cross spectra of the form $\varphi^{\ell+\Delta\ell,m}\varphi^{\ell m*}$, for $\Delta\ell = 0, 2, 4$, were computed from each 72-day series and added to form composite cross spectra for all (ℓ, m) satisfying $-\ell \le m \le \ell$ for $20 \le \ell \le 300$. The background components $r^{(\ell+\Delta\ell,m),(\ell,m)}$ of the $\Delta\ell = 2$ and 4 cross spectra, like those of the power spectra ($\Delta\ell = 0$), are significant and we fit the m-dependence of the background at each $\Delta\ell$ to a polynomial function of m. To normalize the cross spectra, a single scale factor was also applied to all the theoretical cross spectra at each n and ℓ, before adding the background components. The eigenfunction and background parameters and the overall scale factor are the only parameters that were fitted in this analysis, other parameters, characterizing the modes and the instrumental leakage having already been determined as part of the pipeline analysis. In particular, since the mode-frequency parameters are not being fitted for, our fitting scheme yields information about the a-coefficients only through the eigenfunction-leakage dependence of the cross spectra. Unlike the fitting method of S92, the current fitting procedure gives equal weight to

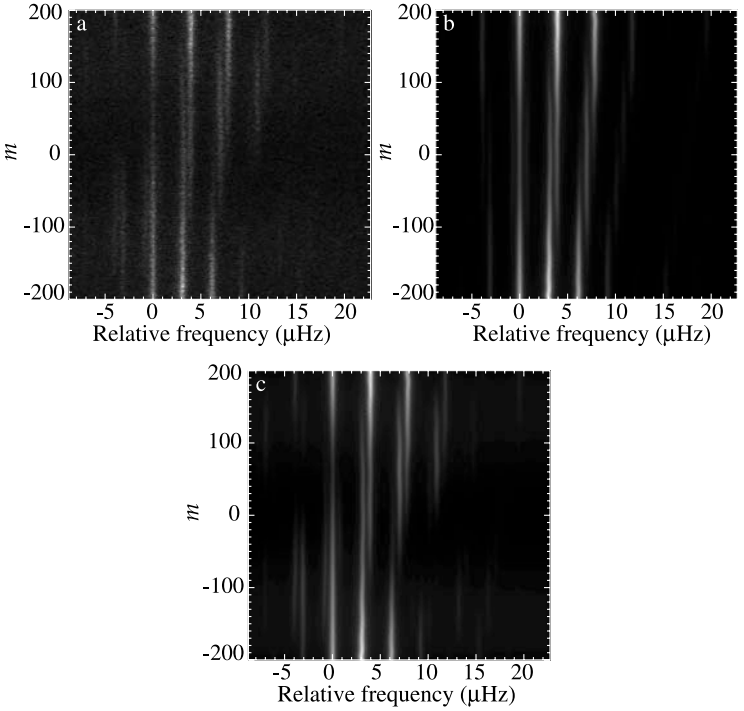


Figure 1 The real parts of cross spectra for $\Delta\ell = 0$ (power spectra) at $\ell = 200$ as a function of m and ν relative to the frequency of the f-mode multiplet of degree ℓ: (a) the cross spectrum computed from HMI data; (b) the best fit to (a) ignoring the effect of flows on the mode eigenfunctions; (c) the best fit to (a) when the effects of differential rotation and meridional flow on the eigenfunctions are included.

individual cross-spectral samples and is therefore suboptimal from the standpoint of signal-to-noise.

3.2. Meridional- and Zonal-Flow Estimates

To estimate the parameters for given (n, ℓ), we first removed the dominant effect of solar rotation in the cross spectra by sliding the observed and theoretical cross spectra in frequency at each m to compensate for the observed m-dependent frequency splitting, estimated from pipeline analysis. Examples of "de-rotated" cross spectra are shown in Figures 1 through 5. It should be noted that the panels of the greyscale plots have been individually scaled. "Relative frequency" is frequency minus the frequency $[\nu_{n\ell}]$ of the targeted multiplet. The lack of centering about $\nu_{n\ell}$ of the cross-spectral leakage pattern is a consequence of a horizontal component of oscillatory motion that is shifted in phase from the radial motion. In an f-mode with $m = \ell$, for example, the motion of a parcel of gas is essentially circular. The combination of circular motion and longitudinally dependent angle between the local vertical and the line-of-sight changes the apparent m and ℓ of the surface modal pattern. The fitting procedure used the de-rotated cross-spectral data for all $m = -\ell, \ell$ and $\nu = \omega/2\pi$ within the narrower of the intervals $[\nu_{n\ell} - 2.5\,\Delta_\ell\nu,\ \nu_{n\ell} + 6.5\,\Delta_\ell\nu]$ and $\nu_{n\ell} \pm 150$ μHz, where $\nu_{n\ell}$ is the de-rotated (multiplet) frequency of the modes and $\Delta_\ell\nu \equiv \nu_{n,\ell+1} - \nu_{n\ell}$.

As the cross-spectral data depend non-linearly on the zonal-flow velocity, the a_3- and a_5-parameters were estimated from the data in an iterative fashion. For each iteration, the

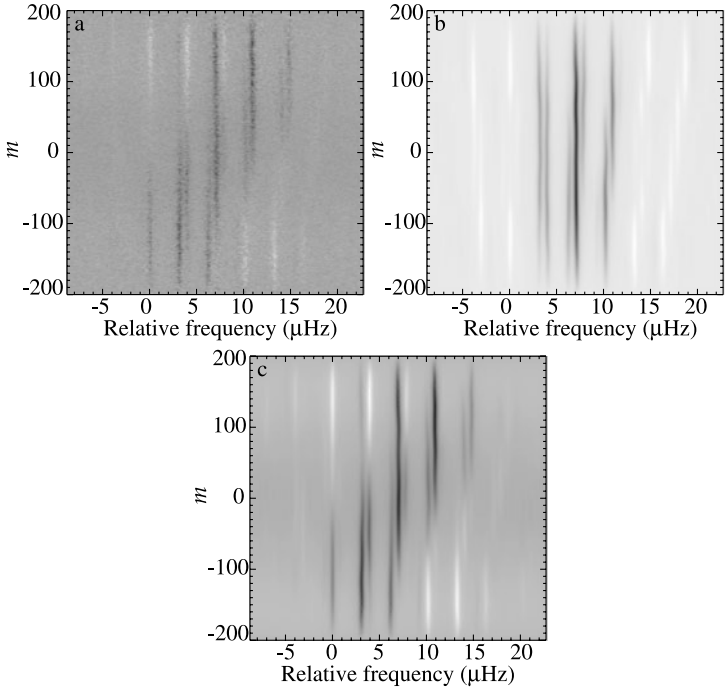


Figure 2 The real parts of cross spectra for $\Delta\ell = 2$ at $\ell = 200$ as a function of m and ν relative to the frequency of the f-mode multiplet of degree ℓ: (a) the cross spectrum computed from HMI data; (b) the best fit to (a) ignoring the effect of flows on the mode eigenfunctions; (c) the best fit to (a) when the effects of differential rotation and meridional flow on the eigenfunctions are included.

overall scale factor and background parameters were first determined from a linear fit to the observed power spectra ($\Delta\ell = 0$). The power-spectral background was modeled as a second-order polynomial function of m^2. Then the values of a_3 and a_5 were updated incrementally based on a simple linearized fitting to the real parts of the observed $\Delta\ell = 2$ and 4 cross spectra. The values of a_3 and a_5 for the first iteration were obtained from mode-frequency measurements and, as noted previously, meridional flow does not affect the real parts of cross spectra in our model. The fitting procedure for zonal flow required only a few iterations to converge. Having determined the overall scale factor and the zonal-flow parameters, we then obtained the meridional-flow parameter a_2^2 from linear fits to the imaginary parts of both the $\Delta\ell = 2$ and 4 cross spectra. Quadratic background models were again assumed. Figure 6 shows the multiplets for which flow-dependent parameters were obtained.

To compare the fitted and observed data more quantitatively, we averaged the de-rotated cross spectra over m for positive and negative m, as shown in Figures 7 through 11. These figures illustrate the degree to which the cross-spectral model is improved by the inclusion of flow-dependent eigenfunction leakages. While the signature of even the weaker, meridional, flow is quite pronounced at $\ell = 200$, it is much less pronounced at $\ell = 100$. At the lowest ℓs that we examined, even the signature of the stronger, zonal flow is hard to discern in the cross spectra of individual multiplets.

Figures 12 and 13 show the a_3- and a_5-coefficients, obtained from the real part of the combined $\Delta\ell = 2$ and 4 cross-spectral leakage fits, as a function of $\nu_{n\ell}/\ell$, which increases with the penetration depth of the modes. The same coefficients, obtained from pipeline

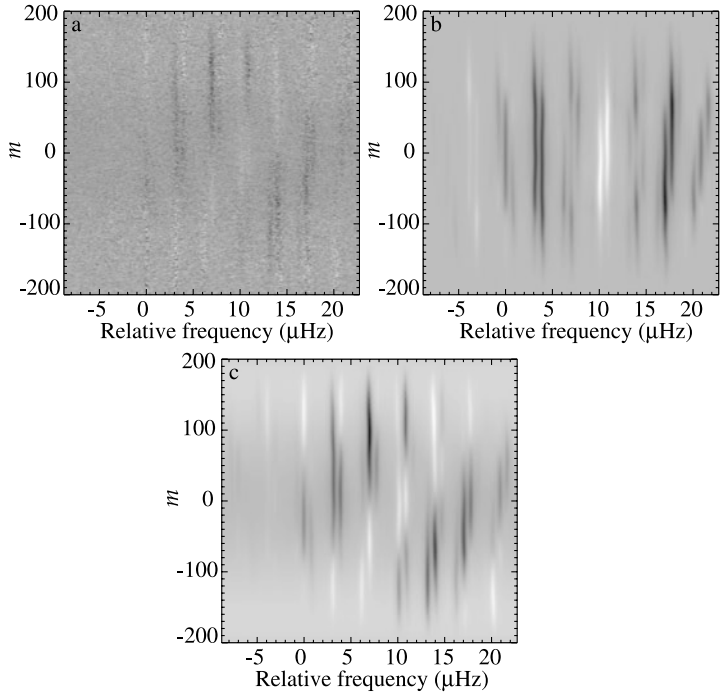


Figure 3 The real parts of cross spectra for $\Delta\ell = 4$ at $\ell = 200$ as a function of m and ν relative to the frequency of the f-mode multiplet of degree ℓ: (a) the cross spectrum computed from HMI data; (b) the best fit to (a) ignoring the effect of flows on the mode eigenfunctions; (c) the best fit to (a) when the effects of differential rotation and meridional flow on the eigenfunctions are included.

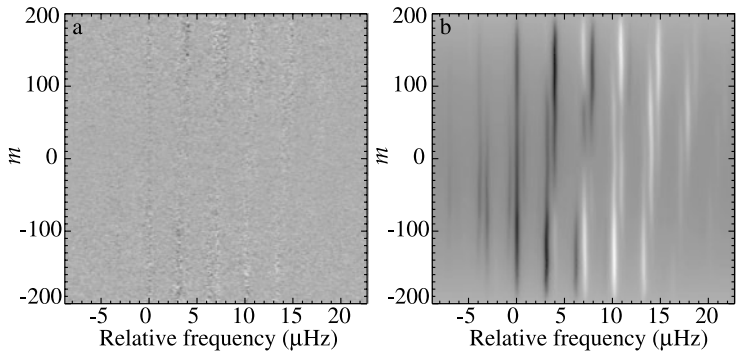


Figure 4 The imaginary parts of cross spectra for $\Delta\ell = 2$ at $\ell = 200$ as a function of m and ν relative to the frequency of the f-mode multiplet of degree ℓ: (a) the cross spectrum computed from HMI data; (b) the best fit to (a) when the effects of differential rotation and meridional flow on the eigenfunctions are included.

mode-frequency analysis, are shown in Figures 14 and 15. The rapid decline in the magnitudes of the a-coefficients at high $\nu_{n\ell}/\ell$ is the signature of the tachocline layer near the bottom of the convection zone. As in the mode-frequency measurement, the scatter in the $a_s(n\ell)$ measurements obtained from the leak fits increases with increasing $\nu_{n\ell}/\ell$, illustrat-

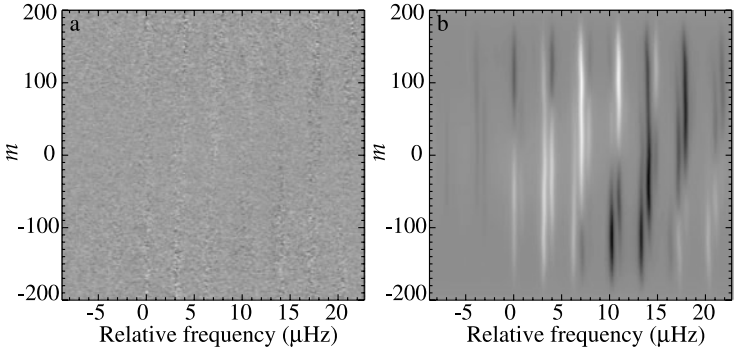


Figure 5 The imaginary parts of cross spectra for $\Delta\ell = 4$ at $\ell = 200$ as a function of m and ν relative to the frequency of the f-mode multiplet of degree ℓ: (a) the cross spectrum computed from HMI data; (b) the best fit to (a) when the effects of differential rotation and meridional flow on the eigenfunctions are included.

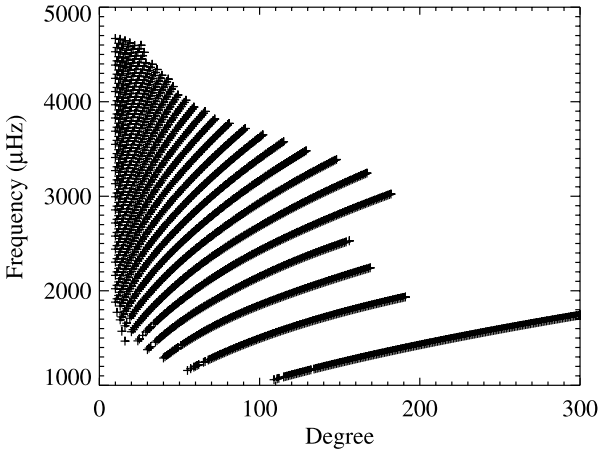


Figure 6 Locations in frequency–wavenumber space of the p- and f-mode multiplets used in the analysis of HMI cross-spectral data.

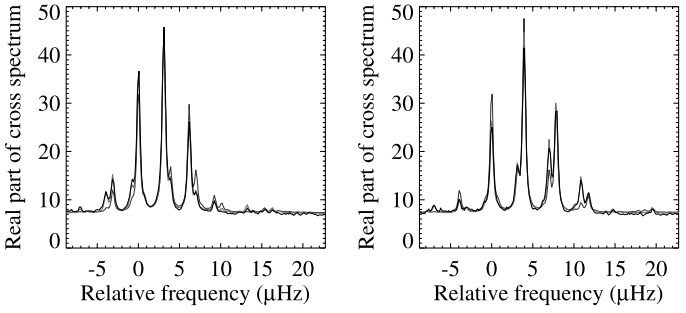


Figure 7 The real parts of the $\Delta\ell = 0$ cross spectra of Figure 1 averaged over negative m (left) and positive m (right). The black curves are from the observed cross spectra. The blue curves are from the fitted cross spectra for the case where the effect of flows on mode eigenfunctions has been ignored. The red curves are from the fitted cross spectra for the case where the effect of differential rotation and meridional flow has been included.

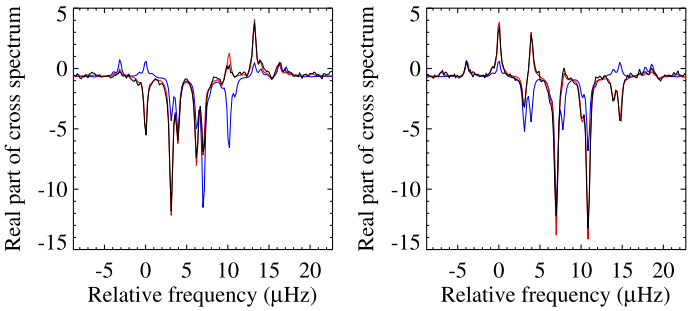


Figure 8 The real parts of the $\Delta\ell = 2$ cross spectra of Figure 2 averaged over negative m (left) and positive m (right). The black curves are from the observed cross spectra. The blue curves are from the fitted cross spectra for the case where the effect of flows on mode eigenfunctions has been ignored. The red curves are from the fitted cross spectra for the case where the effect of differential rotation and meridional flow has been included.

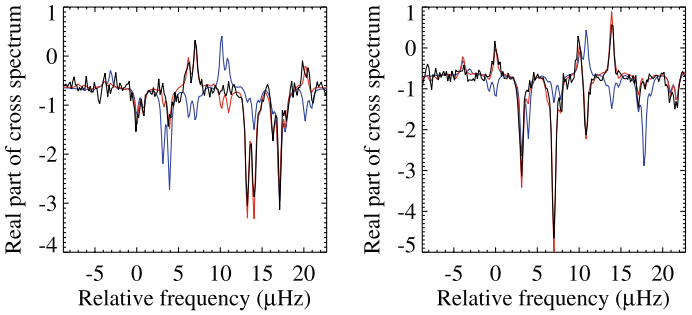


Figure 9 The real parts of the $\Delta\ell = 4$ cross spectra of Figure 3 averaged over negative m (left) and positive m (right). The black curves are from the observed cross spectra. The blue curves are from the fitted cross spectra for the case where the effect of flows on mode eigenfunctions has been ignored. The red curves are from the fitted cross spectra for the case where the effect of differential rotation and meridional flow has been included.

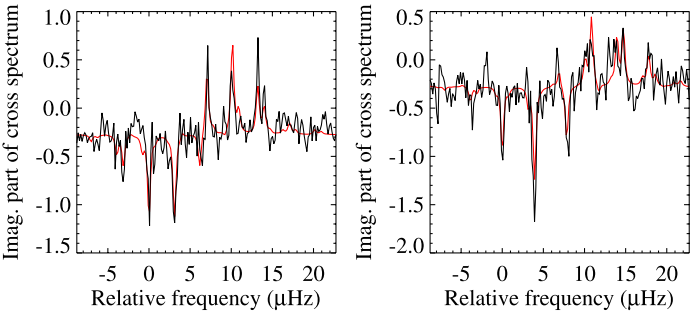


Figure 10 The imaginary parts of the $\Delta\ell = 2$ cross spectra of Figure 4 averaged over negative m (left) and positive m (right). The black curves are from the observed cross spectra. The red curves are from the fitted cross spectra for the case where the effect of differential rotation and meridional flow has been included.

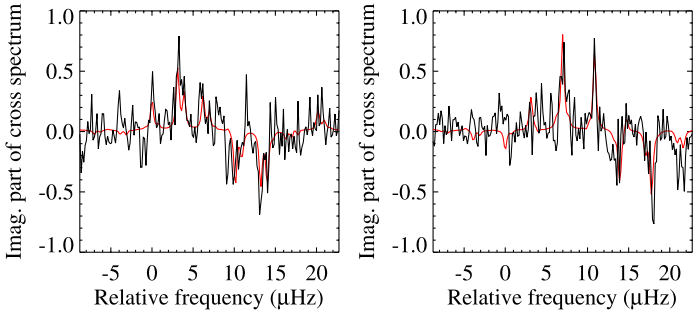


Figure 11 The imaginary parts of the $\Delta\ell = 4$ cross spectra of Figure 5 averaged over negative m (left) and positive m (right). The black curves are from the observed cross spectra. The red curves are from the fitted cross spectra for the case where the effect of differential rotation and meridional flow has been included.

Figure 12 The a_3-coefficient determined from cross-spectral analysis, as described in the text.

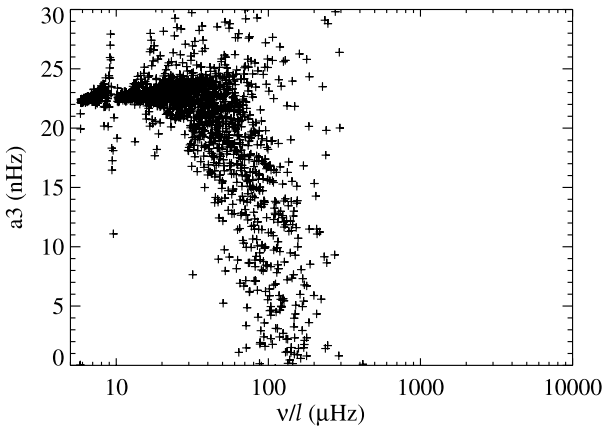
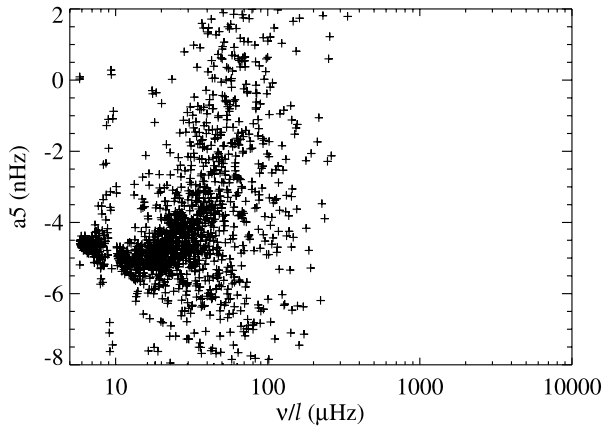


Figure 13 The a_5-coefficient determined from cross-spectral analysis.



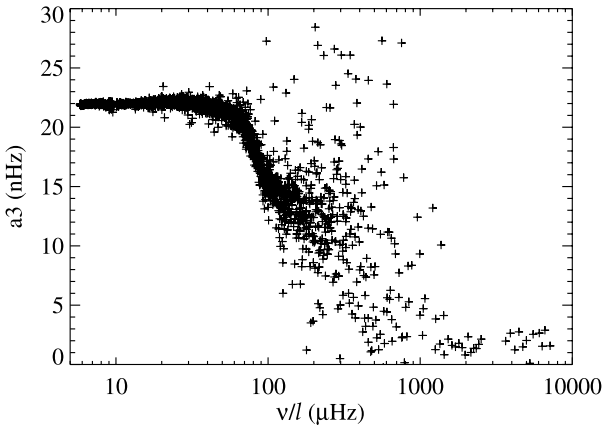


Figure 14 The a_3-coefficient determined from the m-dependence of the measured oscillation frequencies as a function of $\nu_{n\ell}/\ell$.

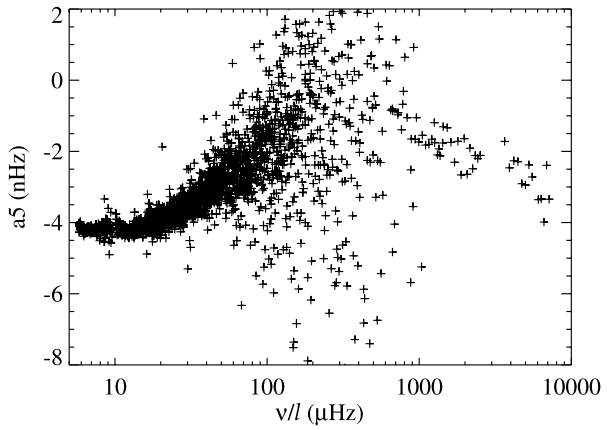


Figure 15 The a_5-coefficient determined from the m-dependence of the measured oscillation frequencies.

ing a generic difficulty in measuring flows deep in the convection zone. The meridional-flow rates obtained from the imaginary part of the $\Delta\ell = 2$ and 4 cross-spectral fits are shown in Figures 16 and 17. The flow measure plotted is $u_{\max} \equiv (45/16\pi)^{1/2}\, R_\odot \langle v_2(r)/r \rangle_{n\ell}$, which is related to the a_2^2-coefficient by Equation (9). The quantity $u_{\max}$ can be interpreted as the peak surface velocity of a hypothetical flow with $v_2(r)/r = \langle v_2(r)/r \rangle_{n\ell}$. Rather than a downturn at high $\nu_{n\ell}/\ell$, which would be expected for a simple conveyor-belt model of meridional flow in the convection zone, the magnitude of $u_{\max}$ shows a conspicuous rise towards the base of the convection zone, the origin of which is unclear (see Section 4 below).

4. Discussion and Conclusions

The preliminary analysis that we have carried out both confirms the basic soundness of the leakage-matrix approach to measuring axisymmetric large-scale flow and indicates where improvement might be needed. In view of the approximate nature of the cross-spectral model, we are encouraged by the level of agreement between the new estimates of zonal flow and the earlier measurements based on mode-frequency splittings. The meridional-flow measurements, on the other hand, are less satisfactory. For oscillations of $\nu/\ell < 30$ μHz,

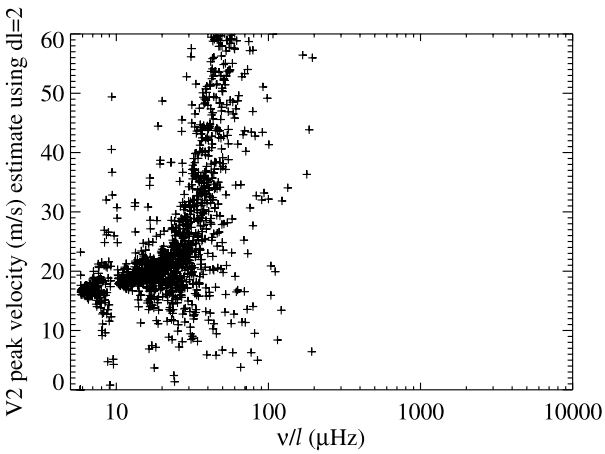


Figure 16 The peak surface velocity [$u_{\max}$] determined from fits to $\Delta\ell = 2$ cross spectra as a function of $\nu_{n\ell}/\ell$.

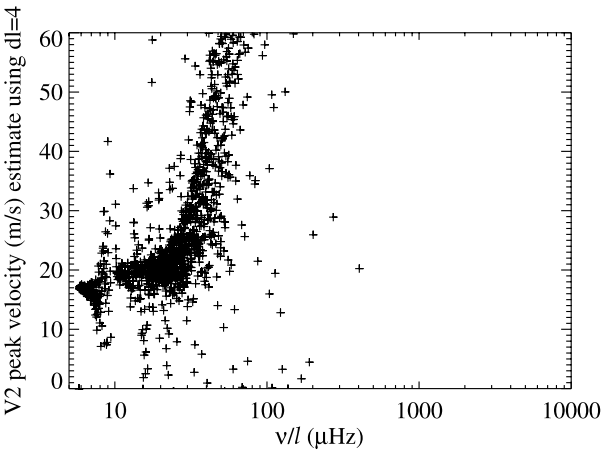


Figure 17 The peak surface velocity [$u_{\max}$] determined from fits to $\Delta\ell = 4$ cross spectra.

corresponding to a ray-turning radius of approximately $0.9R_\odot$, we obtain horizontal-flow velocities of 15 to 25 $\mathrm{m\,s^{-1}}$, in rough agreement with the bulk of seismic observations. For more deeply penetrating modes, the inferred flow velocity increases significantly with penetration depth, at odds with Giles' thesis, although in qualitative agreement with Duvall and Hanasoge (2009). As our deep-flow results are also theoretically unexpected, we view them cautiously.

As we have noted, our current fitting procedure is suboptimal, since the data are not properly weighted. It would be more appropriate to use a simple extension of the S92 method, which uses a weighting function based on the true likelihood function of the data. It would be desirable to model additional multipole components of the flow – odd-s components of meridional flow and even-s components of zonal flow – and fit all flow components, mode parameters, and other parameters together. It would also be desirable to extend the fitting to include cross-spectral data of odd $\Delta\ell \equiv \ell' - \ell$ and non-zero $\Delta m \equiv m' - m$. For future work we also envision using more exact, numerically computed, eigenfunctions.

We anticipate making further improvements to the leakage model. Duvall and Hanasoge (2009) point out that failure to model the center-to-limb variation in the light-travel time

between points in the photosphere and the observation point leads to an apparent material flow toward the center of the disk. If this effect were corrected for, the apparent increase in the meridional-flow velocity with depth would be even more dramatic than we see. By our own estimates, the light-travel-time effect is important even for waves propagating at a depth of 35 Mm. Thus there may be important flow-velocity artifacts at all depths that need to be understood and modeled. In our analysis the observed oscillation signal is used as a proxy for an ideal, Doppler-velocity signal at a fixed height in the solar atmosphere. We note that any effects that, like the light-travel-time effect, either retard or advance the observed oscillation signal relative to the ideal signal are potential sources of flow-velocity error. A radial variation in the phase of the mode eigenfunctions in combination with a center-to-limb variation in the height at which the oscillations are observed is one such effect. We note that non-adiabatic effects, such as radiative damping, on vertically trapped modes can produce radial phase variations. Thus it is important to correct for the light-travel-time effect and to understand and model the effects of atmospheric wave motion on the observations.

The analysis of Braun and Birch (2008) implies that, with 11 years of helioseismic data and a substantially better understanding of systematic errors, it will be possible to detect a meridional flow of a few $\mathrm{m\,s^{-1}}$, near the bottom of the convection zone. Even with less data and relatively modest improvements in estimating systematic error it should be possible to address important questions about shallow dynamos.

Acknowledgements We thank Doug Braun for useful discussions and Ashley Crouch for useful discussions and calculations. HMI data are courtesy of NASA/SDO and the HMI science team. This research was supported by NASA contracts NAS5-03114 and NNH09CF93C to NWRA/CoRA. HMI is supported by NASA contract NAS5-02139 to Stanford University.

References

Anderson, E.R., Duvall, T.L. Jr., Jefferies, S.M.: 1990, Modeling of solar oscillation power spectra. *Astrophys. J.* **364**, 699 – 705. doi:10.1086/169452.

Braun, D.C., Birch, A.C.: 2008, Prospects for the detection of the deep solar meridional circulation. *Astrophys. J. Lett.* **689**, L161 – L165. doi:10.1086/595884.

Braun, D.C., Fan, Y.: 1998, Helioseismic measurements of the subsurface meridional flow. *Astrophys. J. Lett.* **508**, L105 – L108. doi:10.1086/311727.

Braun, D.C., Fan, Y.: 1999, Erratum: helioseismic measurements of the subsurface meridional flow. *Astrophys. J. Lett.* **510**, L81. doi:10.1086/311804.

Chatterjee, P., Antia, H.M.: 2009, Solar flows and their effect on frequencies of acoustic modes. *Astrophys. J.* **707**, 208 – 217. doi:10.1088/0004-637X/707/1/208.

Chou, D.-Y., Ladenkov, O.: 2005, Evolution of solar subsurface meridional flows in the declining phase of Cycle 23. *Astrophys. J.* **630**, 1206 – 1212. doi:10.1086/432372.

Dikpati, M., Gilman, P.A.: 2006, Simulating and predicting solar cycles using a flux-transport dynamo. *Astrophys. J.* **649**, 498 – 514. doi:10.1086/506314.

Duvall, T.L. Jr.: 1979, Large-scale solar velocity fields. *Solar Phys.* **63**, 3 – 15. doi:10.1007/BF00155690.

Duvall, T.L. Jr., Hanasoge, S.M.: 2009, Measuring meridional circulation in the Sun. In: Dikpati, M., Arentoft, T., González Hernández, I., Lindsey, C., Hill, F. (eds.) *Solar-Stellar Dynamos as Revealed by Helio- and Asteroseismology: GONG 2008/SOHO 21* **CS-416**, Astron. Soc. Pac., San Francisco, 103.

Giles, P.M.: 2000, Time–distance measurements of large-scale flows in the solar convection zone. PhD thesis, Stanford University.

Giles, P.M., Duvall, T.L. Jr., Scherrer, P.H.: 1998, Time–distance measurements of subsurface rotation and meridional flow. In: Korzennik, S. (ed.) *Structure and Dynamics of the Interior of the Sun and Sun-like Stars* **SP-418**, ESA, Noordwijk, 775.

Giles, P.M., Duvall, T.L., Scherrer, P.H., Bogart, R.S.: 1997, A subsurface flow of material from the Sun's equator to its poles. *Nature* **390**, 52 – 54. doi:10.1038/36294.

Gizon, L., Birch, A.C.: 2004, Time–distance helioseismology: Noise estimation. *Astrophys. J.* **614**, 472 – 489. doi:10.1086/423367.

Gizon, L., Duvall, T.L., Schou, J.: 2003a, Wave-like properties of solar supergranulation. *Nature* **421**, 43 – 44.
Gizon, L., Duvall, T.L. Jr., Schou, J.: 2003b, Erratum: wave-like properties of solar supergranulation. *Nature* **421**, 764.
González Hernández, I., Patrón, J., Bogart, R.S., The SOI Ring Diagram Team: 1999, Meridional flows from ring diagram analysis. *Astrophys. J. Lett.* **510**, L153 – L156. doi:10.1086/311811.
Gough, D., Hindman, B.W.: 2010, Helioseismic detection of deep meridional flow. *Astrophys. J.* **714**, 960 – 970. doi:10.1088/0004-637X/714/1/960.
Haber, D.A., Hindman, B.W., Toomre, J., Bogart, R.S., Larsen, R.M., Hill, F.: 2002, Evolving submerged meridional circulation cells within the upper convection zone revealed by ring-diagram analysis. *Astrophys. J.* **570**, 855 – 864. doi:10.1086/339631.
Harvey, J.W., Hill, F., Hubbard, R.P., Kennedy, J.R., Leibacher, J.W., Pintar, J.A., Gilman, P.A., Noyes, R.W., Title, A.M., Toomre, J., Ulrich, R.K., Bhatnagar, A., Kennewell, J.A., Marquette, W., Patron, J., Saa, O., Yasukawa, E.: 1996, The Global Oscillation Network Group (GONG) project. *Science* **272**, 1284 – 1286. doi:10.1126/science.272.5266.1284.
Hathaway, D.H.: 2011, The Sun's shallow meridional circulation. arXiv:1103.1561.
Hughes, S.J., Thompson, M.J.: 2003, Time–distance helioseismology of subsurface flows. In: Sawaya-Lacoste, H. (ed.) *GONG 2002. Local and Global Helioseismology: The Present and Future* **SP-517**, ESA, Noordwijk, 307 – 310.
Korzennik, S.G., Rabello-Soares, M.C., Schou, J.: 2004, On the determination of Michelson Doppler imager high-degree mode frequencies. *Astrophys. J.* **602**, 481 – 516. doi:10.1086/381021.
Labonte, B.J., Howard, R.: 1982, Solar rotation measurements at Mount Wilson. III–Meridional flow and limbshift. *Solar Phys.* **80**, 361 – 372. doi:10.1007/BF00147982.
Lavely, E.M., Ritzwoller, M.H.: 1992, The effect of global-scale, steady-state convection and elastic-gravitational asphericities on helioseismic oscillations. *Phil. Trans. Roy. Soc. London A* **339**, 431 – 496.
Miesch, M.S., Elliott, J.R., Toomre, J., Clune, T.L., Glatzmaier, G.A., Gilman, P.A.: 2000, Three-dimensional spherical simulations of solar convection. I. Differential rotation and pattern evolution achieved with laminar and turbulent states. *Astrophys. J.* **532**, 593 – 615. doi:10.1086/308555.
Ritzwoller, M.H., Lavely, E.M.: 1991, A unified approach to the helioseismic forward and inverse problems of differential rotation. *Astrophys. J.* **369**, 557 – 566. doi:10.1086/169785.
Roth, M., Stix, M.: 2008, Meridional circulation and global solar oscillations. *Solar Phys.* **251**, 77 – 89. doi:10.1007/s11207-008-9232-6.
Schad, A., Timmer, J., Roth, M.: 2011, A unified approach to the helioseismic inversion problem of the solar meridional flow from global oscillations. *Astrophys. J.* **734**, 97. doi:10.1088/0004-637X/734/2/97.
Scherrer, P.H., Bogart, R.S., Bush, R.I., Hoeksema, J.T., Kosovichev, A.G., Schou, J., Rosenberg, W., Springer, L., Tarbell, T.D., Title, A., Wolfson, C.J., Zayer, I., MDI Engineering Team: 1995, The solar oscillations investigation – Michelson Doppler Imager. *Solar Phys.* **162**, 129 – 188. doi:10.1007/BF00733429.
Schou, J.: 1992, On the analysis of helioseismic data. PhD thesis, Aarhus University, Aarhus, Denmark [S92].
Schou, J.: 2003, Wavelike properties of solar supergranulation detected in Doppler shift data. *Astrophys. J. Lett.* **596**, L259 – L262. doi:10.1086/379529.
Schou, J., Bogart, R.S.: 1998, Flow and horizontal displacements from ring diagrams. *Astrophys. J. Lett.* **504**, L131. doi:10.1086/311575.
Schou, J., Brown, T.M.: 1994, Generation of artificial helioseismic time-series. *Astron. Astrophys. Suppl.* **107**, 541 – 550.
Schou, J., Christensen-Dalsgaard, J., Thompson, M.J.: 1994, On comparing helioseismic two-dimensional inversion methods. *Astrophys. J.* **433**, 389 – 416. doi:10.1086/174653.
Schou, J., Scherrer, P.H., Bush, R.I., Wachter, R., Couvidat, S., Rabello-Soares, M.C., Bogart, R.S., Hoeksema, J.T., Liu, Y., Duvall, T.L., Akin, D.J., Allard, B.A., Miles, J.W., Rairden, R., Shine, R.A., Tarbell, T.D., Title, A.M., Wolfson, C.J., Elmore, D.F., Norton, A.A., Tomczyk, S.: 2012, Design and ground calibration of the Helioseismic and Magnetic Imager (HMI) instrument on the Solar Dynamics Observatory (SDO). *Solar Phys.* **275**, 229 – 259. doi:10.1007/s11207-011-9842-2.
Sheeley, N.R. Jr.: 2005, Surface evolution of the Sun's magnetic field: A historical review of the flux-transport mechanism. *Living Rev. Solar Phys.* **2**, 5. http://www.livingreviews.org/lrsp-2005-5.
Sivaraman, K.R., Sivaraman, H., Gupta, S.S., Howard, R.F.: 2010, Return meridional flow in the convection zone from latitudinal motions of umbrae of sunspot groups. *Solar Phys.* **266**, 247 – 259. doi:10.1007/s11207-010-9620-6.
Unno, W., Osaki, Y., Ando, H., Saio, H., Shibahashi, H.: 1989, *Nonradial Oscillations of Stars*, Univ. of Tokyo Press, Tokyo.
Vorontsov, S.V.: 2007, Solar p modes of high degree l: Coupling by differential rotation. *Mon. Not. Roy. Astron. Soc.* **378**, 1499 – 1506. doi:10.1111/j.1365-2966.2007.11894.x.

Vorontsov, S.V.: 2011, Effects of differential rotation and meridional circulation in solar oscillations of high degree l. *Mon. Not. Roy. Astron. Soc.* **418**, 1146 – 1155. doi:10.1111/j.1365-2966.2011.19564.x [V11].

Woodard, M.F.: 1989, Distortion of high-degree solar p-mode eigenfunctions by latitudinal differential rotation. *Astrophys. J.* **347**, 1176 – 1182. doi:10.1086/168206.

Woodard, M.F.: 2000, Theoretical signature of solar meridional flow in global seismic data. *Solar Phys.* **197**, 11 – 20. doi:10.1023/A:1026508211960.

Woodard, M.F.: 2007, Probing supergranular flow in the solar interior. *Astrophys. J.* **668**, 1189 – 1195. doi:10.1086/521391.

Woodard, M.F.: 2009, Helioseismic measurement of large-scale solar flows. In: Dikpati, M., Arentoft, T., González Hernández, I., Lindsey, C., Hill, F. (eds.) *Solar-Stellar Dynamos as Revealed by Helio- and Asteroseismology: GONG 2008/SOHO 21* **CS-416**, Astron. Soc. Pac., San Francisco, 15.

Zhao, J., Kosovichev, A.G.: 2004, Torsional oscillation, meridional flows, and vorticity inferred in the upper convection zone of the sun by time–distance helioseismology. *Astrophys. J.* **603**, 776 – 784. doi:10.1086/381489.

DOI 10.1007/978-1-4899-8005-2_9
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0182-7

Analysis of the Helioseismic Power-Spectrum Diagram of a Sunspot

Junwei Zhao · Dean-Yi Chou

Received: 24 April 2012 / Accepted: 29 October 2012 / Published online: 16 November 2012

Abstract The continuous high spatial resolution Doppler observation of the Sun by the *Solar Dynamics Observatory/Helioseismic and Magnetic Imager* allows us to compute a helioseismic k–ω power-spectrum diagram using only oscillations inside a sunspot. Individual modal ridges can be clearly seen with reduced power in the k–ω diagram that is constructed from a 40-hour observation of a stable and round sunspot. Comparing this with the k–ω diagram obtained from a quiet-Sun region, one sees that inside the sunspot the f-mode ridge is more reduced in power than the p-mode ridges, especially at high wavenumbers. The p-mode ridges all shift toward lower wavenumber (or higher frequency) for a given frequency (or wavenumber), implying an increase of phase velocity beneath the sunspot. This is probably because the acoustic waves travel across the inclined magnetic field of the sunspot penumbra. Line-profile asymmetries exhibited in the p-mode ridges are more significant in the sunspot than in the quiet Sun. Convection inside the sunspot is also highly suppressed, and its characteristic spatial scale is substantially larger than the typical convection scale of the quiet Sun. These observational facts demand a better understanding of magnetoconvection and interactions of helioseismic waves with magnetic field.

Keywords Sun: helioseismology · Sun: oscillations · Sun: sunspots

1. Introduction

Analyzing helioseismic power-spectrum diagrams, also known as k–ω diagrams obtained from solar oscillation signals, is a useful way to derive solar atmospheric properties, in-

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

J. Zhao (✉)
W. W. Hansen Experimental Physics Laboratory, Stanford University, Stanford, CA 94305-4085, USA
e-mail: junwei@sun.stanford.edu

D.-Y. Chou
Physics Department, National Tsing Hua University, Hsinchu, Taiwan
e-mail: chou@phys.nthu.edu.tw

terior structures, and subsurface dynamics (*e.g.* Rhodes *et al.*, 1997; Schou *et al.*, 1998). While many global-scale properties of the Sun were inferred by analyzing k–ω diagrams obtained from the global-scale observations, some local-scale properties were derived by analyzing k–ω diagrams obtained from oscillations in local areas, known as ring-diagram analysis (Hill, 1988). Solar areas containing active regions have been studied using ring-diagram techniques to infer interior structures (Basu, Antia, and Bogart, 2004; Baldner, Bogart, and Basu, 2011) and subsurface flow fields (Haber *et al.*, 2004; Komm *et al.*, 2005) of active regions. However, the k–ω diagrams used in these analyses are mainly composed of oscillations outside of sunspots rather than those inside sunspots, hence they represent properties of large areas instead of properties of confined sunspot areas. Other local-helioseismology techniques, *e.g.* time-distance helioseismology (Duvall *et al.*, 1993b; Kosovichev, Duvall, and Scherrer, 2000; Gizon *et al.*, 2009), also need to apply filters over k–ω diagrams, but not to analyze the k–ω diagrams directly.

Despite the scientific interest and importance of constructing k–ω diagrams using only oscillations inside sunspots, this has not been successfully achieved due to either poor spatial resolution or limited temporal coverage. Penn and LaBonte (1993) attempted to compute k–ω diagrams of two sunspots using observations from the Mees Solar Observatory, but were unable to obtain distinct modal ridges. Using *Hinode* Ca H observations, Nagashima *et al.* (2007) studied acoustic power maps inside a sunspot, but did not analyze its k–ω diagram because of the short temporal coverage. Employing Hankel's decomposition, Braun, Duvall, and LaBonte (1987) analyzed the k–ω diagrams made from acoustic waves traveling into and out from sunspots separately, but these diagrams were not computed using oscillations observed directly inside sunspots. SOHO/MDI (Scherrer *et al.*, 1995) full-disk data did not have sufficient spatial resolution to construct k–ω diagrams using only signals inside sunspots, while their high-resolution data lacked sufficient duration. The *Helioseismic and Magnetic Imager* (Scherrer *et al.*, 2012; Schou *et al.*, 2012) onboard the *Solar Dynamics Observatory* (SDO/HMI) provides continuous observations of the Sun with a high spatial resolution and steady temporal cadence, which allows us to perform a k–ω diagram analysis for sunspots. In this article, we report some interesting properties found in the k–ω diagram of a sunspot. We introduce our data analysis procedure in Section 2 and present our results in Section 3. We discuss these results in Section 4 and give conclusions in Section 5.

2. Data Analysis

The SDO/HMI provides continuous observations of the full solar disk with a temporal cadence of 45 seconds and a spatial resolution of 1.0 arcsec. Dopplergrams observed by HMI, primarily used for helioseismology studies as recently demonstrated by Zhao *et al.* (2012), were used in this study. The sunspot located inside NOAA AR 11092 was selected for the k–ω power-spectrum analysis. This sunspot, as displayed in Figure 1, was located at the latitude of 12.4 °N, and remained stable and round in shape during the period of analysis. The magnetic field surrounding this sunspot was relatively simple. A period of 40 hours, covering 16:00 UT 2 August through 08:00 UT 4 August 2010, was selected for our analysis. This roughly corresponds to the sunspot's central meridian crossing, from 8.8° East of the central meridian to 8.8° West of it. This 40-hour period was then divided into five eight-hour segments, and each data segment was analyzed separately. The results obtained from each segment were then averaged to obtain the final results. Every eight-hour data segment was tracked with the sunspot's rotation speed as given by Howard (1990), and was remapped using Postel's projection with the remapping center at the sunspot center and with a sampling

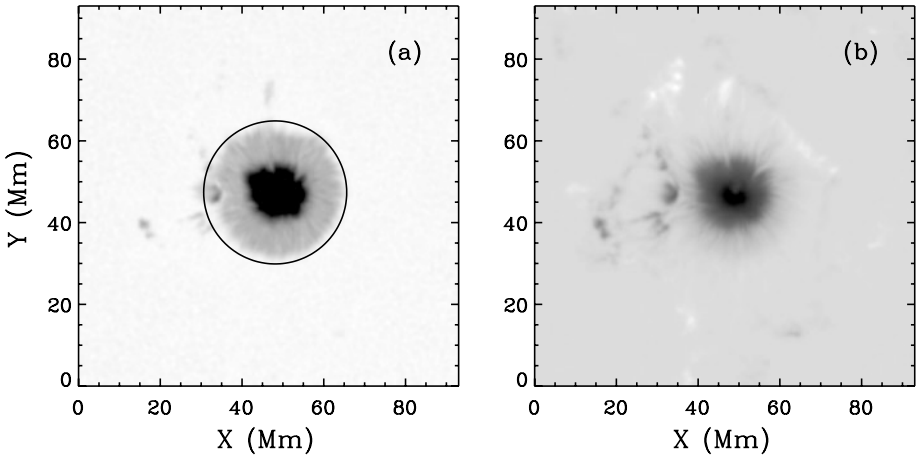


Figure 1 Continuum intensity image (a) and line-of-sight component of the magnetic field (b) for AR 11092. The magnetic-field strength is displayed with a range of −2000 to 500 Gauss. Oscillations observed inside the black circle in panel (a) are used to calculate the k–ω diagram for the sunspot.

resolution of 0.03° pixel^{-1}, slightly oversampling in most areas except near the solar disk center.

We used a Fourier transform to convert data from the space–time domain to the Fourier domain. The k–ω diagram was constructed in the three-dimensional Fourier domain. To avoid possible complications of the Fourier analysis because of the irregular shape of the sunspot boundary, we chose to use the oscillations inside a circular boundary as denoted by the black circle in Figure 1a. When performing Fourier transforms, we kept the entire box size unchanged as shown in Figure 1 and filled up the area outside of the circle with zeroes. This is a normal practice known as zero-padding in Fourier analysis. The area near the boundary of the inside and outside of the circle was tapered with a cosine bell to avoid any sharp transition that might yield spurious signals. For the k–ω diagram made following this procedure, the power at low-wavenumber areas becomes contaminated because of the small box size and the circular structure with data surrounded by zeroes. Therefore we did not take signals lower than $k = 0.3$ rad Mm^{-1} into account in our analysis.

To better understand the analysis results, it is useful to perform similar analyses over quiet-Sun regions for comparison. We selected a piece of quiet Sun that has the same disk location and temporal duration as the analyzed sunspot, covering the period of 16:00 UT 17 August through 08:00 UT 19 August 2010, *i.e.* fifteen days after our sunspot analysis period. The exact same analysis procedures were applied to these data segments, and we used the results as references to compare with the results from the sunspot.

3. Results

Figures 2a and b display the k–ω diagrams calculated from the sunspot region and the quiet-Sun region. Individual modal ridges can be clearly seen in both diagrams. As expected, the helioseismic power of each ridge is much weaker in the diagram obtained from the sunspot than that from the quiet-Sun. Compared with the p-mode (acoustic mode) power reductions for the sunspot, the power reduction in the f-mode (surface-gravity mode) ridge is more

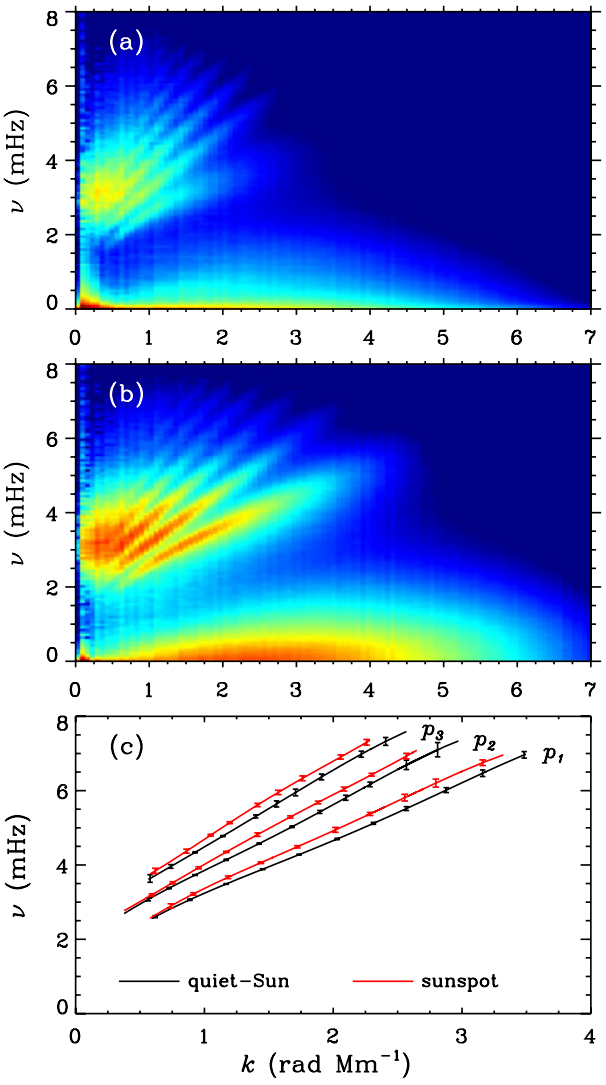


Figure 2 Panels (a) and (b) display the k–ω diagrams obtained from the sunspot and from the quiet-Sun region. Color scales in these two panels are the same. Panel (c) shows the ridge-peak locations, obtained by fitting the two k–ω diagrams, to better illustrate the locations of the modal ridges.

substantial, and the length of this ridge seems shorter due to the higher power reduction at higher wavenumbers. We fit the $p_1 - p_3$ modal ridges by peak-finding for the two k–ω diagrams to better illustrate how the location of each individual ridge obtained from the sunspot differs from that obtained from the quiet-Sun region. The fitting results are shown in Figure 2c. It is difficult to obtain a reliable fitting for the f-mode ridge of the sunspot. The three fitted modal ridges obtained from the sunspot all shift to the lower wavenumber side (alternatively speaking, higher frequency) relative to the ridges obtained from the quiet Sun for a given frequency (or wavenumber). The shift amount is not uniform along each individual ridge, and the largest shift is up to 0.3 mHz.

Based on the fitted modal ridges shown in Figure 2c, we calculated the phase velocity [$v_{\rm ph} = \omega/k$] along each ridge of the two diagrams and display the results in Figure 3a. The derived phase velocity in the sunspot region is often higher by up to 3 km s^{-1} than in the

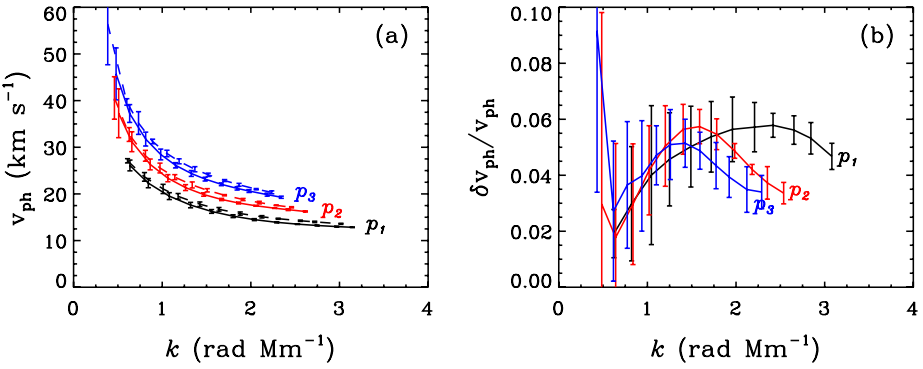


Figure 3 (a) Phase velocities calculated from fitted power ridges $p_1 - p_3$, with solid lines for the quiet-Sun region and dashed lines for the sunspot. (b) Perturbations of phase velocity obtained from the sunspot relative to that obtained from the quiet-Sun region.

quiet-Sun region. Figure 3b shows the phase-velocity perturbation for the sunspot relative to the quiet-Sun region $[\delta v_{\rm ph}/v_{\rm ph}]$ for each modal ridge. The calculated $\delta v_{\rm ph}/v_{\rm ph}$, although large in error bars, mostly fall in range of 0.02 to 0.06.

Comparisons of the ridge shape and location at selected frequencies are displayed in Figure 4. The curves are obtained by averaging a 0.2-mHz-wide band in the k–ω diagrams of Figures 2a and b. Figure 4 again shows that the helioseismic power of the f-mode ridge is more suppressed in the sunspot. All p-mode ridges show clear shifts toward lower wavenumbers; these are more significant at higher frequencies. The amount of ridge shift can be as large as half of the ridge-width for p_1, p_2, and p_3 ridges at the frequency of 5.0 and 6.0 mHz, as can be seen in Figures 4c and d. Figure 4 also shows that the modal ridge line-profile asymmetry is more significant in the sunspot power spectrum than in the power spectrum of the quiet-Sun region. It is well known that the modal ridges obtained from Doppler observations exhibit line-profile asymmetries (Duvall *et al.*, 1993a), which is believed to be caused by correlated noise (Nigam *et al.*, 1998; Georgobiani, Stein, and Nordlund, 2003). It is not immediately clear whether the more prominent line-profile asymmetries in sunspots can be explained by the more substantially correlated noise observed in these areas. On the other hand, recent studies (*e.g.* Chou *et al.*, 2009) found that the power absorption and local suppression of acoustic waves inside sunspots increase with wavenumber, and the more prominent line-profile asymmetry may possibly be caused by this effect.

Figure 5 shows comparisons of ridge shapes and locations at some selected wavenumbers. For each given wavenumber, the $p_1 - p_3$ ridges all display shifts toward higher frequency, and the amount of shift can also be up to half of the ridge width. The significant line-profile asymmetry is also clear in the p-mode ridges obtained from the sunspot region. The shift in the f-mode ridge seems to be toward lower frequency in panels (c) and (d) of Figure 5, but that may just be due to the enhancement of background noises near these frequencies.

It is also interesting to compare the convection power of the sunspot and the quiet-Sun region. The convection powers, which are primarily below 1.5 mHz, can be seen in the lower part of panels (a) and (b) of Figure 2. Following procedures employed by Chou *et al.* (1992) and Georgobiani *et al.* (2007), we calculated the convection power as a function of wavenumber by integrating the power below the straight line that separates the oscillation

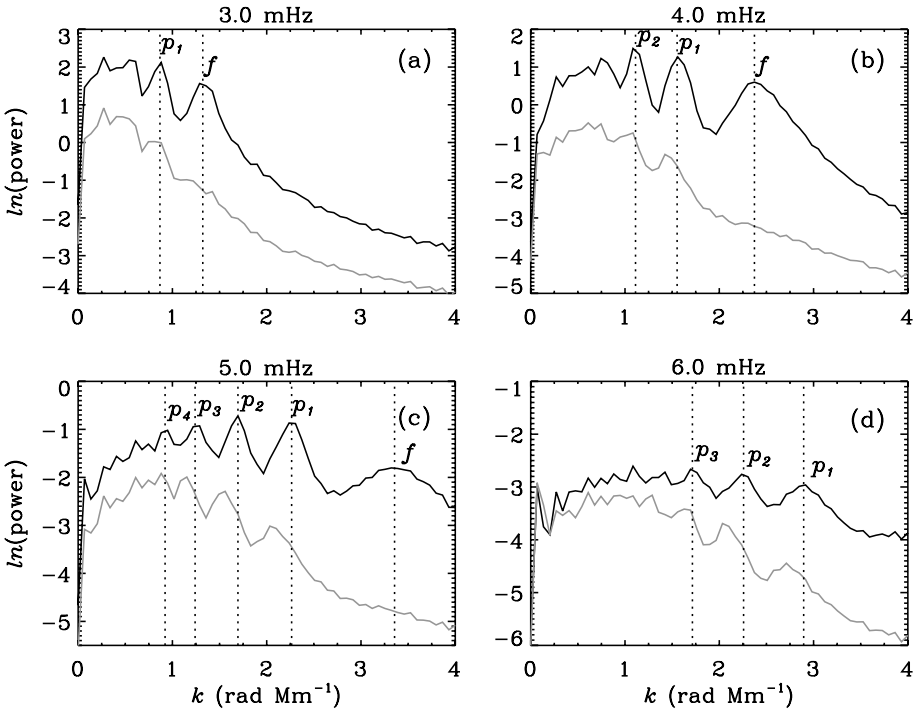


Figure 4 Comparisons of power spectra at different frequencies obtained from the sunspot (gray lines) and the quiet-Sun region (black lines). The vertical dashed lines indicate the locations of the ridge peaks for the quiet Sun.

power and the convection power, *i.e.,* $\nu = 1.03\,k$, where ν is in units of mHz and k in units of rad Mm^{-1}. Figure 6 shows the comparison of convection powers obtained from both k–ω diagrams. Not surprisingly, convection is highly suppressed inside the sunspot. The location of the convection peak is also shifted toward lower wavenumbers in the power spectrum obtained from the sunspot relative to that from the quiet Sun. For the quiet Sun, convection peaks at wavenumbers of approximately 2.7 rad Mm^{-1}, which corresponds to a characteristic spatial scale of 2.3 Mm, presumably representing the scale of solar granules. For the sunspot, convection peaks at approximately 1.9 rad Mm^{-1}, which corresponds to a characteristic spatial scale of 3.3 Mm, more than 40 % larger than the typical convection scale estimated for the quiet Sun. This is consistent with the result of time–distance analysis of a sunspot observed by *Hinode* Ca H line (Zhao, Kosovichev, and Sekii, 2010) in that the convection scale is larger inside sunspots than in the quiet Sun, but different in the convection cell sizes inside the sunspot, which was reported as 4 – 5 Mm in that study. The difference in sizes obtained from these two studies may be caused by the different spectral line formation heights. The reason for the convection cell size to appear larger inside the sunspot is not clear, but it is possible that magnetic field suppresses convection more strongly in smaller scales and leaves larger-scale convection more prominent. Moreover, we cannot completely rule out the possibility that the penumbral filament, which is typically larger than the granules, leaks into our analysis of the convection power inside the sunspot.

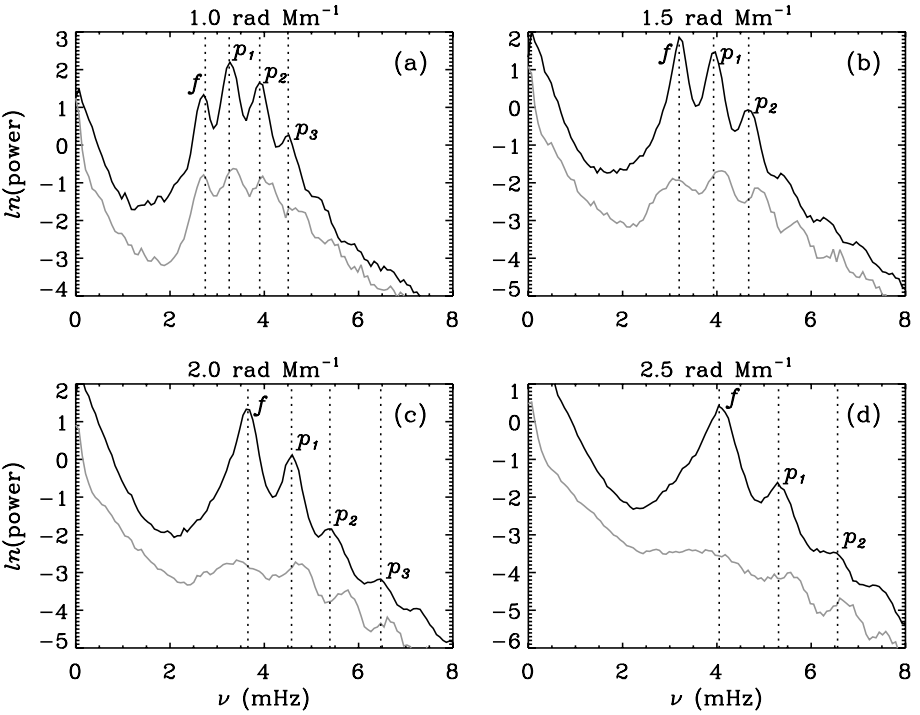
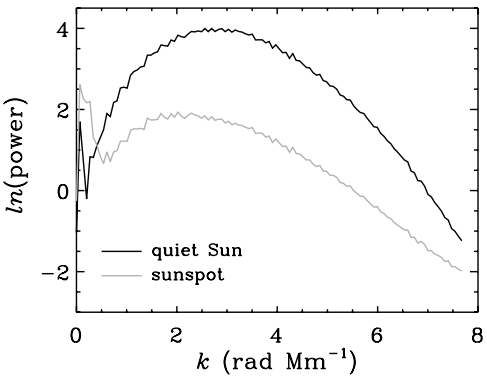


Figure 5 Comparisons of power spectra at different wavenumbers obtained from the sunspot (gray lines) and the quiet-Sun region (black lines). The vertical dashed lines indicate the locations of the ridge peaks for the quiet Sun.

Figure 6 Comparison of the convection power of the sunspot and the quiet-Sun region.



4. Discussion

4.1. Possible Measurement Uncertainties

We have constructed the k–ω power-spectrum diagram using only oscillations inside a sunspot observed by SDO/HMI. Distinct f- and p-mode ridges are clearly seen, which demonstrates that resonant oscillations exist inside the sunspot, primarily inside the sunspot

penumbra. It is not clear to what degree scattered light, *i.e.* the light scattered into the dark sunspot umbra from the brighter surrounding areas, plays a role in our computed k–ω diagram. It was estimated that in an HMI-observed sunspot umbra, the scattered light comprises approximately 20 % of the light intensity (T. Duvall, private communication, 2012). However, the power spectrum seen in Figure 2a is mainly derived from the sunspot penumbra, as is clear from the two experiments that we performed. One experiment was to mask out all oscillations in the umbra, and the other to enhance the oscillations in the umbra to the same amplitude as the oscillations in the penumbra. In both experiments, the k–ω diagram does not differ much from the one shown in Figure 2a except in the very low-k area, where the results are already deemed not useful due to the selection of a small area and zero-padding. Therefore, the effect of scattered light on the computed k–ω diagram of the sunspot is believed to be negligible in the diagram area that we are interested in. We also attempted to construct a k–ω diagram by using only oscillation signals inside the sunspot umbra, but were unable to detect distinct modal ridges, probably because of the small size of the umbra and its weak oscillation amplitude.

The uncertainties in determining Doppler velocities inside the sunspot may also have some effects on the k–ω diagram constructed in this study. Norton *et al.* (2006) discussed the sensitivity of the Fe I 6173 Å line to the magnetic field, and Couvidat *et al.* (2012) examined the uncertainties of deriving Doppler velocities, particularly inside active regions, using different methods. The uncertainties in Doppler velocities may cause some uncertainties in power measurements, but they are not expected to play a significant role in shifting the locations of the modal ridges found in this study.

4.2. Implications for Sunspot Subsurface Structure

It is interesting to understand why the p-mode ridges shift toward lower wavenumbers (or higher frequency). Schunker *et al.* (2008) explored the wave motions with an inclined magnetic field inside the sunspot penumbra, and Felipe, Khomenko, and Collados (2010) discussed the wave mode conversions in the magnetized plasma by using 3D MHD simulations. More recently, Kitiashvili *et al.* (2011) showed through MHD simulations that vertical magnetic field would introduce some acoustic-frequency shifts toward higher frequency, but an inclined magnetic field would shift the frequency more substantially. This is consistent with our observation of modal-ridge shifts toward higher frequency in the sunspot, where the sunspot penumbra, the region with a strongly inclined magnetic field, contributes most to our computed k–ω diagram. It is possible that the modal-ridge shift is caused by that part of the acoustic waves converts to fast magnetoacoustic waves when passing the magnetic field beneath the sunspot, and the phase velocity of fast magnetoacoustic wave is determined by

$$v_{\rm ph}^2 = \frac{1}{2}\left(c_{\rm s}^2 + v_{\rm A}^2\right) + \frac{1}{2}\sqrt{\left(c_{\rm s}^2 + v_{\rm A}^2\right)^2 - 4c_{\rm s}^2 v_{\rm A}^2 \cos^2\psi}, \tag{1}$$

where $c_{\rm s}$, $v_{\rm A}$, and ψ are the sound speed, the Alfvén speed, and the angle between the directions of the wave-vector and magnetic field. The observed power spectrum is an aggregate from different depths beneath the sunspot. Accordingly, it is possible in principle to invert the observed power spectrum for the interior magnetic-field distribution, although this inversion procedure is highly complicated.

The relative change in phase velocity $[\delta v_{\rm ph}/v_{\rm ph}]$ shown in Figure 3b may provide some information of the region below the sunspot. The low values of $\delta v_{\rm ph}/v_{\rm ph}$, which fall into a range of $0.02 - 0.06$, together with Equation (1) suggests that the averaged $\delta c_{\rm s}/c_{\rm s}$ and $v_{\rm A}^2/c_{\rm s}^2$

inside the sunspot is only moderate, where δc_s is the sound-speed change inside the sunspot relative to the quiet Sun. If $\delta c_s/c_s \ll 1$ and $v_A^2/c_s^2 \ll 1$, from Equation (1), we have

$$\frac{\delta v_{ph}}{v_{ph}} \approx \frac{\delta c_s}{c_s} + \frac{1}{2}\sin^2\psi\left(\frac{v_A^2}{c_s^2}\right). \tag{2}$$

The observed value of $\delta v_{ph}/v_{ph}$ gives a constraint on the value of the averaged $\delta c_s/c_s + \sin^2\psi(v_A^2/c_s^2)$, and that is $0.02 < \delta c_s/c_s + \sin^2\psi(v_A^2/c_s^2) < 0.06$. Here $\delta c_s/c_s + \sin^2\psi(v_A^2/c_s^2)$ is averaged over the three-dimensional sunspot region weighted with a kernel, which is proportional to the mode energy density, although the mode energy density inside the sunspot is unknown. Unlike $\sin^2\psi(v_A^2/c_s^2)$, which is always positive, $\delta c_s/c_s$ could be positive or negative below the sunspot, since $\delta c_s/c_s \sim \delta\gamma/2\gamma + \delta T/2T$ and $v_A^2/2c_s^2 = 1/\beta$, where γ is the adiabatic index and β the ratio of gas pressure to magnetic pressure. $|\delta c_s/c_s|$ is lower than $\sin^2\psi(v_A^2/c_s^2)$ near the surface if $\sin^2\psi$ is not too low. Thus, most of the contribution to $\delta v_{ph}/v_{ph}$ could come from $\sin^2\psi(v_A^2/c_s^2)$. Another possible contribution to $\delta v_{ph}/v_{ph}$, which is not shown in Equation (1), is the change in the mode cavity, although it could be weaker than that in Equation (2). The upper turning point inside the sunspot is deeper than in the quiet Sun, thus its contribution to $\delta v_{ph}/v_{ph}$ is positive.

An interesting phenomenon shown in Figure 3b is that $\delta v_{ph}/v_{ph}$ has a general trend for each n: it first increases and then decreases with k. This could be explained by Equation (2). It is generally believed that v_A^2/c_s^2 is significant only near the surface. Since the modes of higher k penetrate less deeply into the solar interior, the averaged v_A^2/c_s^2 increases with k. On the other hand, the modes of higher k propagate less vertically near the surface, and the magnetic field inside the sunspot penumbra has a strong horizontal component near the surface. This causes a decrease of ψ with k, or the average $\sin^2\psi$ increases with k. Therefore, the average v_A^2/c_s^2 increases with k while the average $\sin^2\psi$ decreases with k. This yields a maximum of $\sin^2\psi(v_A^2/c_s^2)$ or $\delta v_{ph}/v_{ph}$ at some k.

4.3. Relevance to Previous Studies

The recent power spectral analysis by Schunker and Braun (2011) showed that the modal ridges obtained from the sunspot vicinity, where acoustic halos are detectable, shift to higher wavenumber relative to the quiet Sun. This shift is in a direction opposite to the one that we find inside the sunspot area in this study. It seems that the modal ridges shift toward lower wavenumbers in areas where the acoustic power is reduced and shift toward higher wavenumbers in areas where the acoustic power is enhanced. Moreover, the phase velocity becomes faster in the sunspot penumbra, possibly implying faster magnetoacoustic waves, and the phase velocity becomes slower in the sunspot vicinity, possibly implying slower magnetoacoustic waves in the area. It is interesting to note that in that same study by Schunker and Braun (2011), the modal-ridge lines showed a weaker line-asymmetry than in the quiet-Sun region, again in contrast to what we found inside a sunspot in this study. This is also a curious phenomenon, but it is not clear whether these opposite line-asymmetry trends are related to the different atmospheric heights of the sunspot penumbra and sunspot vicinity, or to the faster and slower magnetoacoustic waves discussed above.

The modal-ridge shift observed inside the sunspot may complicate the interpretation of the measured travel-time shifts or phase shifts in sunspot areas where a phase-velocity filter or a modal-ridge filter were applied. Because these filters choose parameters based on quiet-Sun areas, the power ridges of sunspots may not be filtered in the same way as the power ridges of quiet-Sun areas because the modal ridges of sunspots are systematically located

on the lower-wavenumber (or higher-frequency) side. However, this effect may not be significant, because Zhao, Kosovichev, and Sekii (2010) showed that the measurements with and without phase-speed filters gave results that qualitatively agreed well. This observed modal-ridge shift in sunspot regions may not have an effect on the subsurface properties in active regions derived from the ring-diagram analysis, such as given by Baldner, Bogart, and Basu (2011), because the acoustic power of the sunspot contributed little to the power spectrum used for the ring-diagram analysis, which usually analyzes a much larger area than a sunspot.

5. Conclusion

Using 40 hours of continuous SDO/HMI Doppler observations of a stable sunspot, we constructed a k–ω power-spectrum diagram that shows distinct modal ridges with suppressed helioseismic power. More notably, compared with the k–ω diagram obtained from a quiet-Sun region, the f-mode ridge in the sunspot diagram becomes more suppressed in power than the p-mode ridges, especially at high wavenumbers. All p-mode ridges shift toward the lower-wavenumber (or higher-frequency) for a given frequency (or wavenumber). The phase velocity computed from our fitted p-mode ridges shows an increase of $2-6$ % inside the sunspot region. The inclined magnetic field in the sunspot penumbra may be responsible for the p-mode ridge shifts and the increase of the phase velocity. The line-profile asymmetries in the p-mode ridges for the sunspot region are more prominent than the asymmetries observed in the quiet Sun, which may be relevant for the higher power absorption and suppression ratio in higher wavenumbers inside the sunspot. Convection inside the sunspot is also highly suppressed, but exhibits a characteristic spatial scale approximately 40 % larger than the typical granulation scale in the quiet Sun. Realistic MHD simulations of magneto-convection with helioseismic waves included (*e.g.* Rempel *et al.*, 2009) are very useful for understanding all these observed facts inside the sunspot region. We expect that an analysis similar to this one, but using numerical simulation data, will help in better understanding the sunspot subsurface structure.

Acknowledgements SDO is a NASA mission, and HMI project is supported by NASA contract NAS5-02139. We thank Mark Cheung for useful discussions on convection scales inside sunspots. We also thank the anonymous referee for useful comments that improved the quality of this article.

References

Baldner, C.S., Bogart, R.S., Basu, S.: 2011, *J. Phys. Conf. Ser.* **271**, 012006.
Basu, S., Antia, H.M., Bogart, R.S.: 2004, *Astrophys. J.* **610**, 1157.
Braun, D.C., Duvall, T.L. Jr., LaBonte, B.J.: 1987, *Astrophys. J.* **319**, L27.
Chou, D.-Y., Chen, C.-S., Ou, K.-T., Wang, C.-C.: 1992, *Astrophys. J.* **396**, 333.
Chou, D.-Y., Yang, M.-H., Zhao, H., Liang, Z.-C., Sun, M.-T.: 2009, *Astrophys. J.* **706**, 909.
Couvidat, S., Rajaguru, S.P., Wachter, R., Sankarasubramanian, K., Schou, J., Scherrer, P.H.: 2012, *Solar Phys.* **278**, 217. ADS:2012SoPh..278..217C, doi:10.1007/s11207-011-9927-y.
Duvall, T.L. Jr., Jefferies, S.M., Harvey, J.W., Osaki, Y., Pomerantz, M.A.: 1993a, *Astrophys. J.* **410**, 829.
Duvall, T.L. Jr., Jefferies, S.M., Harvey, J.W., Pomerantz, M.A.: 1993b, *Nature* **362**, 430.
Felipe, T., Khomenko, E., Collados, M.: 2010, *Astrophys. J.* **719**, 357.
Georgobiani, D., Stein, R.F., Nordlund, Å.: 2003, *Astrophys. J.* **596**, 698.
Georgobiani, D., Zhao, J., Kosovichev, A.G., Benson, D., Stein, R.F., Nordlund, Å.: 2007, *Astrophys. J.* **657**, 1157.
Gizon, L., Schunker, H., Baldner, C.S., Basu, S., Birch, A.C., Bogart, R.S., Braun, D.C., Cameron, R., Duvall, T.L., Hanasoge, S.M., *et al.*: 2009, *Space Sci. Rev.* **144**, 249.

Haber, D.A., Hindman, B.W., Toomre, J., Thompson, M.J.: 2004, *Solar Phys.* **220**, 371. ADS: 2004SoPh..220..371H, doi:10.1023/B:SOLA.0000031405.52911.08.
Hill, F.: 1988, *Astrophys. J.* **333**, 996.
Howard, R.F.: 1990, *Solar Phys.* **126**, 299. ADS:1990SoPh..126..299H, doi:10.1007/BF00153052.
Kitiashvili, I.N., Kosovichev, A.G., Mansour, N.N., Wray, A.A.: 2011, *Solar Phys.* **268**, 283. ADS: 2011SoPh..268..283K, doi:10.1007/s11207-010-9679-0.
Komm, R., Howe, R., Hill, F., González Hernández, I., Toner, C., Corbard, T.: 2005, *Astrophys. J.* **631**, 636.
Kosovichev, A.G., Duvall, T.L. Jr., Scherrer, P.H.: 2000, *Solar Phys.* **192**, 159. ADS:2000SoPh..192..159K, doi:10.1023/A:1005251208431.
Nagashima, K., Sekii, T., Kosovichev, A.G., Shibahashi, H., Tsuneta, S., Ichimoto, K., Katsukawa, Y., Lites, B., Nagata, S., Shimizu, T., *et al.*: 2007, *Publ. Astron. Soc. Japan* **59**, S631.
Nigam, R., Kosovichev, A.G., Scherrer, P.H., Schou, J.: 1998, *Astrophys. J. Lett.* **495**, L115.
Norton, A.A., Graham, J.P., Ulrich, R.K., Schou, J., Tomczyk, S., Liu, Y., Lites, B.W., López Ariste, A., Bush, R.I., Socas-Navarro, H., Scherrer, P.H.: 2006, *Solar Phys.* **239**, 69. ADS:2006SoPh..239...69N, doi:10.1007/s11207-006-0279-y.
Penn, M.J., LaBonte, B.J.: 1993, *Astrophys. J.* **415**, 383.
Rempel, M., Schüssler, M., Cameron, R.H., Knölker, M.: 2009, *Science* **325**, 171.
Rhodes, E.J. Jr., Kosovichev, A.G., Schou, J., Scherrer, P.H., Reiter, J.: 1997, *Solar Phys.* **175**, 287. ADS:1997SoPh..175..287R, doi:10.1023/A:1004963425123.
Scherrer, P.H., Bogart, R.S., Bush, R.I., Hoeksema, J.T., Kosovichev, A.G., Schou, J., Rosenberg, W., Springer, L., Tarbell, T.D., Title, A., *et al.*: 1995, *Solar Phys.* **162**, 129. ADS:1995SoPh..162..129S, doi:10.1007/BF00733429.
Scherrer, P.H., Schou, J., Bush, R.I., Kosovichev, A.G., Bogart, R.S., Hoeksema, J.T., Liu, Y., Duvall, T.L. Jr., Zhao, J., Title, A.M., *et al.*: 2012, *Solar Phys.* **275**, 207. ADS:2012SoPh..275..207S, doi:10.1007/s11207-011-9834-2.
Schou, J., Antia, H.M., Basu, S., Bogart, R.S., Bush, R.I., Chitre, S.M., Christensen-Dalsgaard, J., di Mauro, M.P., Dziembowski, W.A., Eff-Darwich, A., *et al.*: 1998, *Astrophys. J.* **505**, 390.
Schou, J., Scherrer, P.H., Bush, R.I., Wachter, R., Couvidat, S., Rabello-Soares, M.C., Bogart, R.S., Hoeksema, J.T., Liu, Y., Duvall, T.L. Jr., *et al.*: 2012, *Solar Phys.* **275**, 229. ADS:2012SoPh..275..229S, doi:10.1007/s11207-011-9842-2.
Schunker, H., Braun, D.C.: 2011, *Solar Phys.* **268**, 349. ADS:2011SoPh..268..349S, doi:10.1007/s11207-010-9550-3.
Schunker, H., Braun, D.C., Lindsey, C., Cally, P.S.: 2008, *Solar Phys.* **251**, 341. ADS:2008SoPh..251..341S, doi:10.1007/s11207-008-9142-7.
Zhao, J., Kosovichev, A.G., Sekii, T.: 2010, *Astrophys. J.* **708**, 304.
Zhao, J., Couvidat, S., Bogart, R.S., Parchevsky, K.V., Birch, A.C., Duvall, T.L. Jr., Beck, J.G., Kosovichev, A.G., Scherrer, P.H.: 2012, *Solar Phys.* **275**, 375. ADS:2012SoPh..275..375Z, doi:10.1007/s11207-011-9757-y.

DOI 10.1007/978-1-4899-8005-2_10
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-013-0245-4

Revisiting the Solar Oblateness: Is Relevant Astrophysics Possible?

J.P. Rozelot · Z. Fazel

Received: 21 February 2012 / Accepted: 28 January 2013 / Published online: 26 February 2013

Abstract The measurement of solar oblateness has a rich history extending well back into the past. Until recently, its estimate has been actively disputed, as has its temporal dependence. Recent accurate observations of the solar shape gave cause for doubt, and so far only balloon flights or satellite experiments, such as those onboard SDO, seem to achieve the required sensitivity to measure the expected small deviations from sphericity. A shrinking or an expanding shape is ultimately linked to solar activity (likely not homologously with its change), as gravitational or magnetic fields, which are existing mechanisms for storing energy during a solar cycle, lead to distinct perturbations in the equilibrium solar-structure and changes in the diameter. It follows that a sensitive determination of the solar radius fluctuations might give information about the origin of the solar cycle. In periods of higher activity, the outer photospheric shape seems to become aspheric under the influence of higher-order multipole moments of the Sun, resulting both from the centrifugal force and the core rotation. An accurate determination of the shape of the Sun is thus one of the ways that we have now for peering into its interior, learning empirically about flows and motions there that would otherwise only be guessed at from theoretical considerations, developing more precise inferences, and ultimately building possible alternative gravitational theories.

Keywords Sun: diameter · Sun: rotation · Sun: activity · Celestial mechanics: astrophysical quantities · General relativy: alternative theories

Invited Talk

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

J.P. Rozelot (✉)
OCA-Lagrange, CNRS UMR 7293, Nice University, Bld de l'Observatoire, BP 4229, 06304 Nice Cedex 4, France
e-mail: Jean-Pierre.Rozelot@oca.eu

Z. Fazel
Tabriz University, Faculty of Phys., Tabriz, Iran
e-mail: z-fazel@tabrizu.ac.ir

1. Introduction

The rotation of the Sun induces a deformation of its shape (*i.e.* a deviation from a spherical figure), and the lowest-order harmonic can be characterized as an oblateness signal. Higher-order terms lead to latitudinal variations, in shape and in thermal profiles (Lefebvre and Rozelot, 2004; Rast and Ortiz, 2008), mainly because of the conditions that prevail in the medium, such as meridional circulation, torsional oscillations, or the radial gradient of rotation. In this sub-surface layer, the temperature, density, or pressure vary by many orders of magnitude: the medium changes from optically thick to optically thin and from mostly ionized to mostly neutral; the convective flow is compressible and is permeated by magnetic activity. All these phenomena make the solar surface difficult to model, especially the solar limb (Piau *et al.*, 2011).

The photospheric oblateness is, in principle, sensitive to the interior rotation rate (see for instance Goldreich and Schubert, 1968; Dicke, 1970; Ulrich and Hawkins, 1981; Paternó, Sofia, and Di Mauro, 1996; Lydon and Sofia, 1996; Armstrong and Kuhn, 1999). Thus, accurate observations may be useful because they provide constraints on the rotation of the core; efforts are made to achieve this by many means. In contrast to these core measurements, the solar rotation from a few thousand kilometers below the surface to a depth of about half a solar radius has been measured helioseismically (mainly by the *Global Oscillation Network Group* (GONG: Harvey *et al.*, 1996) and the *Michelson Doppler Imager* (MDI: Scherrer *et al.*, 1995) onboard the *Solar Heliospheric Observatory* (SOHO); today this is performed by means of the *Solar Dynamics Observatory* (SDO) and likely by means of *Picard* when space data will be re-analyzed). These measurements and analyses permitted substantial progress in our knowledge of the rotation rate of the surface at different latitudes and in the solar interior, together with their temporal dependence (Antia, Chitre, and Gough, 2008). There is evidence of distinct patterns of bands of faster and slower than average rotation, which at low latitudes move toward the Equator, and at high latitudes move toward the poles. This pattern is similar to the torsional oscillations observed at the solar surface by Howard and LaBonte (1980), and it has been shown that they penetrate inside the convection zone. Below that, it has been found that the core may rotate faster than the surface (Turck-Chièze and Couvidat, 2011), the origin of which could be the fossil remnant of the rotation of the young Sun (Dicke, 1970; Duez, 2009). These complex rotational processes from the core to the surface, which are variable with time, encourage many studies to make an accurate estimate of the solar shape, and the impact on the solar gravitational moments.

If we consider the observed outer solar figure, different measurements and analyses techniques, sometimes quasi-identical instruments and similar analyses methods yield incompatible results. As pointed out by Sofia *et al.* (2005), "it is presumptuous to infer that the cause of this controversy is that for the majority of the techniques, the results are at the borderline of the sensitivity of the technique". Limb observations have been made by a number of ground-based instruments, for which most of the results are inconsistent, as we see from discussing Figure 1. Above the Earth's atmosphere, observations have been made by the *Solar Disk Sextant* (SDS) balloon-borne experiment (Sofia, Heaps, and Twigg, 1994), the *Reuven Ramaty High-Energy Solar Spectroscopic Imager* (RHESSI: Fivian *et al.*, 2008), by the MDI instrument onboard SOHO (Emilio *et al.*, 2007), and by the HMI instrument onboard SDO, launched 12 February 2010. A dedicated mission (*Picard*) has been conceived to accurately measure the solar diameter in any heliographic latitude and from this basis, check the temporal dependence of the oblateness. Launch was in June 2011. It is clear that new progress will be made with SDO, which is capable of achieving precise solar-radius

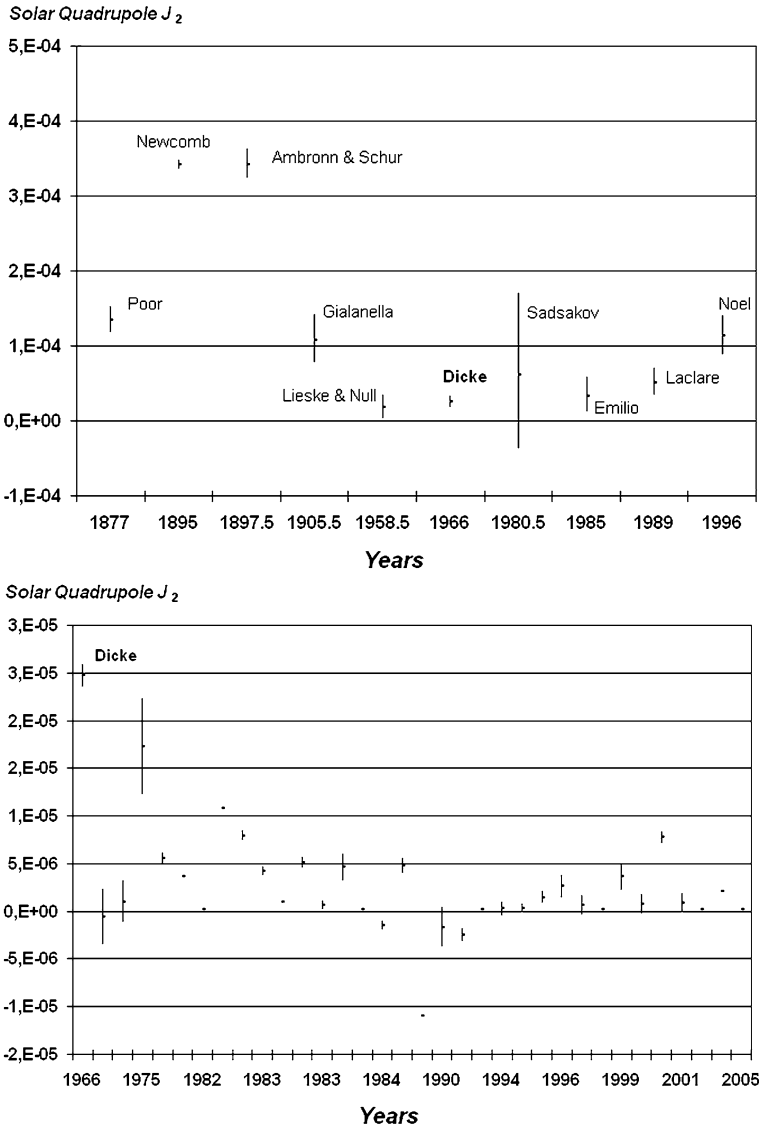


Figure 1 History of the solar gravitational quadrupole moment [J_2] deduced from measurements of the solar oblateness, compiled from the table given by Rozelot and Damiani (2011); points are tabulated in the same chronological order, corresponding to the mean date of the observation(s) (see footnote 3). The upper panel shows estimates about one order of magnitude higher than those listed in the bottom panel. The "historical" measurement of Dicke and Goldenberg (1967) has been plotted both in the upper and lower panels, to link the J_2 scales. Considering only estimates depicted in the lower panel, a dynamical flattening may be inferred, but it seems rather unlikely.

observations with an accuracy sufficient to directly test solar models (Kuhn, 2011). Moreover, these new astrometric results will likely support the view that changes deep in the solar interior are required to account for the observed irradiance variability (through the W parameter: Emilio *et al.*, 2000; Gough, 2001; Lefebvre and Rozelot, 2004). Lastly, it must be

mentioned that one of the major difficulties of measuring the solar limb profile is to visualize and separate the magnetic contamination. Analysis at the limb must not confuse the Sun's oblateness with brightness contributions, and here again, particular attention must be given to the analysis of the data. As pointed out by Kuhn *et al.* (1998), it is important to carry out a detailed analysis of the photospheric intensity distribution, including the brightness at the extreme limb, and the results will certainly have broad importance in solar (and stellar) physics. The possible brightness variations include not only the classical limb-darkening function, but also several other interesting possibilities associated with interior or surface structures (Fivian *et al.*, 2009).

From these above-mentioned features it is clear that the solar-limb structure is complex, and solar oblateness cannot simply be explained by the gravity darkening (or ellipticity dimming) due to centrifugal effects from the known surface rotation alone.

On the other hand, the limb variations that have already been implied by theories or that are currently required by new ones extend from General Relativity (Dicke, 1964) to those that explain the solar-core rotation (Dicke, 1981; Howe, 2009; Sturrock, 2009; Turck-Chièze and Couvidat, 2011). Lastly, surface physics sets the boundary conditions for the solar–stellar structure and evolution. It turns out that the solar oblateness can be considered as a fundamental parameter in solar physics, but its accurate determination, as well as its temporal variability, if any, remains somewhat elusive despite years of research activity in this field. We briefly comment on this point in the next section, taking into account the latest results.

2. A Quick Tour of the Oblateness Issue

From an historical point of view, it seems that Harzer (1891) was the first to introduce the concept of ellipticity into the equations of rotation of the Sun. Defining the flatness ε (sometimes noted as the visual oblateness) by the fractional Equator-to-Pole solar radius variation $[\Delta R = R_{\rm eq} - R_{\rm pol}]$ divided by the solar radius $R_\odot$, *i.e.*, $\varepsilon = \Delta R / R_\odot$, Harzer found $\varepsilon^2 \approx 1/25\,000$, a value certainly inexact, but already of good approximation.

Newcomb (1895) wondered if an oblateness of $\varepsilon \approx 5 \times 10^{-3}$ (or $\Delta R = 500$ mas) could not explain the anomalies of Mercury's observed motion. Even though this oblateness is erroneous, at least the idea deserves to be mentioned because the problem of the solar mass quadrupole moment is again present in the modern theories of gravitation. We recall here that Ambronn and Schur (1905) were able to measure an oblateness of $\varepsilon \approx 2.08 \times 10^{-6}$ by means of their heliometer[1] between 1892 and 1902, while Auwers (1891) concluded from the analysis of Mercury and Venus transits observed between 1874 and 1882, $\varepsilon \approx 3.96 \times 10^{-5}$ (or $\Delta R = 0.038'' \pm 0.023''$). This value was adopted in the ephemerides until 2000. Based on results obtained by means of the SDS experiment on 30 September 1992, Allen (2000) gives $\Delta R = 0.0086''$ (the SDS result is $\varepsilon = (8.63 \pm 0.88) \times 10^{-6}$).[2]

Figure 1[3] shows a history of the first solar gravitational moment deduced from observations up to now and compiled from the table given in Rozelot and Damiani (2011), computed

[1] As computed from their own measurements (see corrections made by Poor, 1905a, 1905b), but they quote in their published material ($\varepsilon = 1/1920$) *i.e.* $\approx 5.2 \times 10^{-4}$.

[2] Allen indicates this reference, but the Sofia result is $\Delta R = 0.0082''$.

[3] The following authors are taken into account in the bottom panel (references in Damiani *et al.*, 2011): Landgraf (1992), Hill and Stebbins (1975), Schatten (1975), Hill, Bos, and Goode (1982), Gough (1982), Pitjeva (2005), Kislik (1983), Dicke, Kuhn, and Libbrecht (1986), Bearsley (1987), Campbell *et al.* (1983),

via $J_2 = (2/3)(\Delta R - \Delta R_{\rm surf})/R_\odot$. The surface rotation rate, about 2 km s^{-1} at the equator, predicts $\Delta R_{\rm surf}$ to be ≈ 7.8 mas (Fivian *et al.*, 2008), or 7.77 mas, following Dicke (1970, Equation (80)), who took into account the solar latitude dependence. It is likely that this value could be slightly overestimated, as recently shown by Kuhn *et al.* (2012).[4] Figure 1 upper panel shows estimates that are about one order of magnitude higher than those listed in Figure 1 lower panel, which render these estimates rather improbable. The "historical" measurement of Dicke and Goldenberg (1967) has been plotted up and down to link the J_2 scales. Considering only estimates depicted in this last figure, a dynamic flattening may be inferred, contradicting helioseismic results. Owing to the difficulty of the measurements, the different criteria used for analyzing the data, mainly through the limb-darkening function and the possible magnetic contamination at the limb that can lead to an equatorial excess in size, it is highly likely that the oblateness may be constant in time.

The question whether the solar oblateness is variable with time deserves to be posed. Hill and Stebbins (1975) suggested a varying oblateness with an average equator-to-pole radius difference of about 7 mas. Then, Dicke, Kuhn, and Libbrecht (1987) suggested an amplitude variability of the oblateness in phase with the solar activity. Considering the observations over one solar cycle (1996 – 2008), the heliometer data obtained by Rösch and Rozelot (1996) at the Pic du Midi observatory still exhibit a very weak relationship (in phase) with the level of activity on the Sun's surface (Damiani *et al.*, 2011). Using the MDI instrument onboard the SOHO spacecraft and after correcting measurements for scattered light, Emilio *et al.* (2007) reported a solar shape almost purely oblate near solar maximum, and showing a significant hexadecapole shape near minimum. A phasing with solar activity was reported, which accounted for a difference between the equatorial and polar radii of 8.7 ± 2.8 mas in 1997 and 18.9 ± 1.9 mas in 2001. A mechanism was suggested to explain these changes: in times of high activity only the first solar gravitational moment has a significant effect, but in times of low activity, the second is predominant; this results in a decrease of the oblateness. The combination of the two terms leads to a solar complex shape (a bulge at the Equator and a depression at mid-latitudes that indicates the J_4 influence), suggesting that the outer solar atmosphere expands nonhomologously during the cycle (Emilio *et al.*, 2007; Rozelot, Damiani, and Pireaux, 2009). In contrast, Egidi *et al.* (2006) and Djafer *et al.* (2008), revisiting balloon-flight data recorded on 30 September 1992, 26 September 1994, 1 October 1995, and 10 October 1996, found an anti-correlation of the solar oblateness with activity. When the activity is zero, modern oblateness measurements *versus* solar activity lead to an oblateness of $\Delta R = 8.21$ mas ($\pm \approx 1$ mas), in good accordance with the theory (Damiani *et al.*, 2011, Figure 5). However, a comparison of all of these (observed) data is difficult because they are not free from a dependence on the limb-darkening function, which is a critical observational quantity, as previously mentioned. One possible conclusion to reconcile all data is to assume a constant oblateness with time, as helioseismic data analyzed by Antia and Basu (2004) show. It is expected that SDO data will enable us to reach a conclusion in one way or another. Preliminary results show, to a certain extent, a very small oblateness variability with time, which does not exceed a few mas of amplitude (on the order of around 0.1 mas for the quadrupole term, and of about 2 mas for the hexadecapolar term: Kuhn *et al.*, 2012).

Hill, Rabey, and Rosenwald (1986), Afanesava (1990), Duvall *et al.* (1984), Dicke, Kuhn, and Libbrecht (1987), Bursa (1986), Maier *et al.* (1992), Egidi *et al.* (2006), Paternó (1996), Rösch (2001), Emilio *et al.* (2007), Pjipers (1998), Reis Neto (2003), Damiani *et al.* (2011), Rösch (2003), Antia *et al.* (2008), Fivian *et al.* (2008) (not corrected), Fivian *et al.* (2009) (corrected).

[4] In terms of angles subtended at Earth, they report from SDO observations that the Sun's equatorial radius exceeds its polar radius by 7.2 ± 0.5 mas: the upper value overlaps that of Dicke.

3. Some Theoretical Statements

Computing the solar oblateness is not an easy task. Chandrasekhar (1933) was the pioneer in determining the oblateness of stars taking into account a uniform rotation rate combined with a density that increased from the surface to the core as a power law. Maeder (1999) generalized von Zeipel's theorem to account for differential rotation in the case of a "shellular" rotation law and found that the differential rotation increases the oblateness. However, in the solar case, the radial gradient of rotation $[\partial\omega/\partial r]$ varies with the heliographic latitude, which is negative from the Equator up to about 50° and is positive thereafter; hence the solar oblateness is not as pronounced as expected, since generally $\partial\omega/\partial r < 0$ (however, the exact shape critically depends on the rotation law in the external layers). Dicke (1970) found that the contribution to the oblateness induced by surface rotation including the effect of differential rotation is $\varepsilon = 0.81 \times 10^{-5}$ (or $\Delta R = 7.77$ mas), a value "nearly equal to that of a surface rotating at the angular velocity seen at around 45° latitude" (in fact 46°). This mean theoretical solar ellipticity may be distorted by local stresses mainly through magnetic fields, underlining the need to perform astrometric measurements at the limb.

4. Relevance of J_n to Solar Astrophysics and Other Fields

4.1. Solar Astrophysical Quantities

A number of questions are still pending about solar global properties. Among them is the question about the influence of solar core dynamics on the values of the global solar spin $[\vec{J_\odot}]$, on the multipolar moments $[J_n]$, and on the solar shape coefficients $[c_n]$. How does the temporal dependence of J_n, c_n, and $\vec{J_\odot}$ on the solar cycle affect the motion of the planets?

All these quantities are at the crossroad of solar physics, astrometry, and celestial mechanics. From the point of view of solar physics, their values reflect the physics of solar models: non-rigid rotation, solar latitudinal rotation, solar-core properties, solar-cycle variations, and stellar evolution. Knowing their temporal variations is essential for constraining solar-cycle modeling or solar-evolution theories.

A quite comprehensive approach was first made by Komm *et al.* (2003), who derived a temporal variation of the coefficient of P_4 (in the Legendre expansion of the rotation law) from helioseismic data, which is indicative of torsional oscillation patterns extending deep into the convection zone, while the angular momentum shows a 1.3-year period that hints at a long-term trend likely related to the solar activity cycle. Also through analyzing helioseismic data, Antia and Basu (2004) found a significant temporal variation in the angular momentum and in the gravitational multipole moments. For these authors, the quadrupole moment $[J_2]$ exhibits no noticeable temporal variation, while J_4 up to J_{12} do show a stronger variability. To interpret the discrepancy with the Kuhn, Libbrecht, and Dicke (1988) results, the authors argued that "most of the distortion from sphericity is the direct response to the centrifugal force on the rotating surface layers, and not from the asphericity of the gravitational field". This argument contradicts Dicke's statement: "a solar ellipticity requires a rotating distorted gravitational potential function, the most likely being a fast rotating solar core. There are only two feasible sources of internal stress to distort the solar core, a magnetic field and a Reynolds stress (fluid motion). The latter does not readily yield a rotating distortion, but a magnetic field demands it".

4.2. Relativistic Astrometry

Space–time is shaped by the presence of solar system bodies. The space–time curvature induced by the Sun leads to light deflection or corresponding time-delays in the propagation of signals. Precise astrometry in the solar neighborhood will thus require precise knowledge of the solar quadrupole moment and spin. Indeed, in addition to the solar mass monopole contribution (≈ 1.75 arcsec light deflection at grazing incidence), there is a quadrupolar one ($\approx 0.4 - 0.3$ μarcsec at grazing incidence) and rotational solar contribution ($\approx \pm 0.7$ μarcsec at grazing incidence) (Pireaux, 2002). Unfortunately, the contribution of the J_2-term to light deflection drops dramatically as the angle of incidence (*i.e.* the closest approach to the Sun) increases (non-grazing incidence). Hence the need for determining an accurate J_2.

4.3. Relativistic Celestial Mechanics

The solar quadrupole moment also plays a role in celestial mechanics. The relativistic precession of the perihelion of planets is a known phenomenon (Pireaux and Rozelot, 2003), ≈ 43 arcsec per century in the case of Mercury ($42.9794''$: Pireaux and Rozelot, 2003). As planetary ephemerides are today fitted to observational data sets (tracking data of spacecraft, VLBI angular positions, or flyby normal points), a precise knowledge of J_2 and its variation are necessary to improve the fitting process. The statistics of the obtained postfit residuals may help to better constrain the ephemerides. As an example, the knowledge of Mercury's orbit has made significant progress using this method, reaching a few meters to be compared with 800 m obtained through direct radar ranging (Cavanaugh *et al.*, 2007).

Furthermore, the gravitational field associated with the mass quadrupole of the Sun creates a Newtonian perturbation on the orbit of the planets that has an effect on planetary spins and on the ecliptic plane. Up to now, the terms that contain the contribution to J_2 implicitly assume that the Sun's equatorial plane and the orbital plane of the planet in question coincide, which is an approximation. Here also, more progress must be made even if some work has been done in this direction (Iorio, 2011; Xu *et al.*, 2011). Finally, through solar system spin–orbit coupling, J_2 and $\vec{J}_\odot$ will indirectly influence the orbital parameters of solar system bodies. For example, the Moon–Earth spin–orbit coupling propagates the influence of the solar quadrupole moment to the Moon. This has allowed us to set a dynamic upper limit to the solar quadrupole moment of $J_2 \leq 3 \times 10^{-6}$ (Rozelot and Rösch, 1997; Rozelot and Bois, 1998; Bois and Girard, 1999) through observed lunar librations, but a more improved approach remains to be made.

4.4. Possible Alternative Theories of Gravitation?

Currently, there is still a strong correlation between the post-Newtonian parameter $[\beta]$ and J_2 in planetary ephemerides. Hence, one cannot simultaneously fit for those two parameters, but a better knowledge of J_2 would help long-term solar system modeling.

Indeed, together with γ, which encodes the amount of curvature of space–time per unit rest-mass, the post-Newtonian parameter $[\beta]$ contributes to the relativistic precession of the planets. The latter parameter encodes the amount of nonlinearity in the superposition law of gravitation (with $\beta \equiv 1$ in General Relativity (GR)). Using a reasonable value for J_2 and the current best constraints on post-Newtonian parameters γ and β, GR is still a possible solution, but there is room for alternative theories as well (Pireaux and Rozelot, 2003). At present, the best values available for β and γ are $(\gamma - 1) \times 10^5 = 2.1 \pm 2.3$ (*Cassini* mission, Bertotti, Iess, and Tortora, 2003), $(\beta - 1) \times 10^4 = 1.2 \pm 1.1$ (Lunar Laser Ranging,

LLR: Williams, Turyshev, and Boggs, 2009, Equation (27)) and $\eta \times 10^4 = (4\beta - \gamma - 3) = 4.4 \pm 4.5$, through LLR also (Williams, Turyshev, and Boggs, 2009, Equation (26)).

By accurately measuring the perihelion advances of several planets, it will be possible in the near future to decorrelate all these quantities. Such a test is an objective of several space missions, such as *Beppi–Columbo*, *Gaia,* and *Lator*. The primary objective of this last mission is the measurement of the key post-Newtonian Eddington parameter γ with an accuracy of one part in 10^9 (a factor 30 000 beyond the current best result obtained by means of *Cassini* (Bertotti, Iess, and Tortora, 2003). Direct measurement of the solar quadrupole moment $[J_2]$ to an accuracy of one part in 200 of its size of around 10^{-7} is expected (Turyshev *et al.*, 2009). Currently, a first analysis of the *Messenger* flybys (Iorio, 2011) shows that the Lense–Thirring precession in the case of Mercury may have been canceled to a certain extent by the competing precession caused by a small mismodeling in the quadrupole mass moment $[J_2]$ of the Sun, which could be estimated at 1.8×10^{-7}, which is more compatible with the RHESSI results ($J_2 = (1.46 \pm 1.0) \times 10^{-7}$: Fivian *et al.*, 2008).

5. Conclusion

We emphasized the need to accurately capture real solar latitudinal intensity limb variations over a long period of time – a still challenging task (Schwarzschild, 2012). We highlighted that the limb's tiny departure from perfect circularity, *i.e.* the asphericity, is sensitive to the Sun's otherwise invisible interior conditions, allowing us to possibly learn empirically about flows and motions there that could otherwise only be guessed at from theoretical considerations. We then recalled that knowing the precise solar gravitational moments $[J_n]$ and their temporal dependence can be used to precisely determine solar parameters such as the angular momentum and its variation on the long term, which may help to locate the seat of the dynamo (Komm *et al.*, 2003). From all observed oblateness estimates, we also showed that the oblateness variation, in phase with the solar cycle, remains very faint (a few mas in amplitude), with a possible exchange in magnitude between the two first solar moments ($n = 2$ and 4). Furthermore, J_2 at first and surely J_4 are at the crossroads of relativistic astrometry and relativistic celestial mechanics. Indeed, i) one of the major effects of a dynamical flattening of the Sun due to its oblique rotation with respect to the ecliptic plane is a secular variation in the orbital elements of the planet, which is not negligible, and must not be neglected in computing the ephemerides of celestial bodies. ii) The gravitational moments J_2 and J_4 must be used, via their decorrelation with PPN parameters, as tests of alternative theories of gravitation.

Lastly, detailed analysis of the photospheric intensity distribution, including the brightness at the extreme limb, is important for determining the solar shape and its underlying parameters, shape coefficients, and oblateness, for which we showed that the most likely value of Δr is 7.77 mas at low activity (the best observed values give 8.21 ± 1 mas, which is within the theoretical estimate). The accurate knowledge of all above-mentioned quantities is certainly of broad importance in solar and stellar physics, mainly for solar-cycle modeling and time evolution. The possible brightness variations include not only the classical limb-darkening function, but also several other capabilities that enable us to deal with the surface as well as the interior structure (Isaak, 1999). The analysis must completely eliminate possible contributions to the solar photospheric intensity from an abundance of unresolved magnetic elements of very low magnetic-flux density. First results obtained in this way from SDO seem promising.

Together with space missions dedicated to testing the curvature of the solar system's gravity field with an accuracy better than 1 part in 10^9, determining the gravitational moments will allow us to estimate the solar-core velocity rate with an unequaled precision.

Acknowledgements This work was supported by the French CNRS and Tabriz University. JPR is indebted to A. Kosovichev for inviting him to participate in the SDO-3 meeting held at Stanford University. The authors thank all referees, who provided helpful suggestions to clarify specific issues.

References

Allen, C.W.: 2000, *Astrophysical Quantities*, 4th edn., Springer, Berlin.
Ambronn, L., Schur, A.C.W.: 1905, In: *Astron. Mitt. der Königlichen Sternwarte zu Göttingen*, Part 7 **III**, No. 3, 126.
Antia, H.M., Basu, S.: 2004, In: Dansey, D. (ed.) *Proc. SOHO 14/GONG 2004 Workshop* **SP-559**, ESA, Noordwijk, 301.
Antia, H.M., Chitre, S.M., Gough, D.O.: 2008, *Astron. Astrophys.* **477**, 657.
Armstrong, J., Kuhn, J.R.: 1999, *Astrophys. J.* **525**, 533.
Auwers, A.: 1891, *Astron. Nachr.* **128**, 361.
Bertotti, B., Iess, L., Tortora, P.: 2003, *Nature* **425**, 374.
Bois, E., Girard, J.F.: 1999, *Celest. Mech.* **73**, 329.
Cavanaugh, J.F., Smith, J.C., Sun, X., Bartels, A.E., Ramos-Izquierdo, L., Krebs, D.J., McGarry, J.F., Trunzo, R.M., Novo-Gradac, A., Britt, J.L., Karsh, J., Katz, R.B., Lukemire, A.T., Szymkiewicz, R., Berry, D.L., Swinski, J.P., Neumann, G.A., Zuber, M.T., Smith, D.E.: 2007, *Space Sci. Rev.* **131**, 451.
Chandrasekhar, S.: 1933, *Mon. Not. Roy. Astron. Soc.* **93**, 390.
Damiani, C., Rozelot, J.P., Lefebvre, S., Kilcik, A., Kosovichev, A.K.: 2011, *J. Atmos. Solar-Terr. Phys.* **73**, 241. doi:10.1016/j.jastp.2010.02.021.
Dicke, R.H.: 1964, *Nature* **202**, 432.
Dicke, R.H.: 1970, *Astrophys. J.* **159**, 1.
Dicke, R.H.: 1981, *Solar Phys.* **78**, 3. ADS:1982SoPh...78....3D. doi:10.1007/BF00151138.
Dicke, R.H., Goldenberg, H.M.: 1967, *Phys. Rev. Lett.* **18**, 313.
Dicke, R.H., Kuhn, J., Libbrecht, K.G.: 1987, *Astrophys. J.* **318**, 451.
Djafer, D., Thuillier, G., Sofia, S., Egidi, A.: 2008, *Solar Phys.* **247**, 225. ADS:2008SoPh..247..225D. doi:10.1007/s11207-007-9079-2.
Duez, V.: 2009, Ph.D. thesis, Ecole Polytechnique, Paris.
Egidi, A., Caccin, B., Sofia, S., Heaps, W., Hoegy, W., Twigg, L.: 2006, *Solar Phys.* **235**, 407. ADS:2006SoPh..235..407E. doi:10.1007/s11207-006-0073-x.
Emilio, M., Kuhn, J., Bush, R.I., Scherrer, P.: 2000, *Astrophys. J.* **543**, 1007.
Emilio, M., Bush, R.I., Kuhn, J., Scherrer, P.: 2007, *Astrophys. J. Lett.* **660**, L161.
Fivian, M.D., Hudson, H.S., Lin, R.P., Zahid, H.J.: 2008, *Science* **322**, 560.
Fivian, M.D., Hudson, H.S., Lin, R.P., Zahid, H.J.: 2009, *Science* **324**, 1143.
Goldreich, P., Schubert, G.: 1968, *Astrophys. J.* **154**, 1005.
Gough, D.: 2001, *Nature* **410**, 313.
Harvey, J.W., Hill, F., Hubbard, R.P., Kennedy, J.R., Leibacher, J.W., Pintar, J.A., Gilman, P.A., Noyes, R.W., Title, A.M., Toomre, J., Ulrich, R.K., Bhatnagar, A., Kennewell, J.A., Marquette, W., Patron, J., Saa, O., Yasukawa, E.: 1996, *Science* **274**, 1284.
Harzer, P.: 1891, *Astron. Nachr.* **127**, 17.
Hill, H.A., Stebbins, R.T.: 1975, *Astrophys. J.* **200**, 471.
Howard, R., LaBonte, B.J.: 1980, *Astrophys. J. Lett.* **239**, L33.
Howe, R.: 2009, *Living Rev. Solar Phys.* **6**, 1. solarphysics.livingreviews.org/Articles/lrsp-2009-1/.
Iorio, L.: 2011, arXiv:1109.0266v3 (7 October 2013).
Isaak, G.: 1999, In: Wilson, A. (ed.) *Proc 9th Meeting on Solar Physics: Magnetic Fields and Solar Processes* **SP-448**, ESA, Noordwijk, 1.
Komm, R., Howe, R., Durney, B.R., Hill, F.: 2003, *Astrophys. J.* **586**, 650.
Kuhn, J.: 2011, In: *Solar Dynamics and Magnetism from the Interior to the Atmosphere, SDO 3 Meeting*, Stanford Univ.
Kuhn, J., Libbrecht, K.G., Dicke, R.H.: 1988, *Science* **242**, 908.
Kuhn, J.R., Bush, R.I., Scheik, X., Scherrer, P.: 1998, *Nature* **392**, 155. doi:10.1038/32361.
Kuhn, J.R., Bush, R., Emilio, M., Scholl, I.F.: 2012, *Science* **337**, 1638.

Lefebvre, S., Rozelot, J.P.: 2004, *Astron. Astrophys.* **419**, 1133.
Lydon, T.J., Sofia, S.: 1996, *Phys. Rev. Lett.* **76**, 177.
Maeder, A.: 1999, *Astron. Astrophys.* **347**, 185.
Newcomb, S.: 1895, *Fundamental Constants of Astronomy*, US Gov. Print Office, Washington, 111.
Paternó, L., Sofia, S., Di Mauro, M.P.: 1996, *Astron. Astrophys.* **314**, 940.
Piau, L., Stein, R.F., Melo, S., Turck-Chièze, S., Thuillier, G., Hauchecorne, A.: 2011, In: Alecian, G., Belkacem, K., Samadi, R., Valls-Gabaud, D. (eds.) *Proc. French Soc. Astron. Astrophys.*, EDP Sciences, Les Ulis, 407.
Pireaux, S.: 2002, Ph.D. Thesis, Université Catholique de Louvain.
Pireaux, S., Rozelot, J.P.: 2003, *Astrophys. Space Sci.* **284**, 1159.
Poor, C.L.: 1905a, *Astrophys. J.* **22**, 103.
Poor, C.L.: 1905b, *Astrophys. J.* **22**, 305.
Rast, M.P., Ortiz, A.: 2008, *Astrophys. J.* **673**, 1209.
Rösch, J., Rozelot, J.P.: 1996, *C. R. Acad. Sci. Paris, Sér. II* **322**, 637.
Rozelot, J.P., Bois, E.: 1998, In: Balasubramaniam, K.S., Harvey, J., Rabin, D.S. (eds.) *Synoptic Solar Physics – 18th NSO/Sacramento Peak Summer Workshop* **CS-140**, Astron. Soc. Pac., San Francisco, 75.
Rozelot, J.P., Damiani, C.: 2011, *Eur. Phys. J. H* **36**, 407.
Rozelot, J.P., Damiani, C., Pireaux, S.: 2009, *Astrophys. J.* **703**, 1791.
Rozelot, J.P., Rösch, J.: 1997, *Solar Phys.* **172**, 11. ADS:1997SoPh..172...11R. doi:10.1023/A:1004919823102.
Scherrer, P.H., Bogart, R.S., Bush, R.I., Hoeksema, J.T., Kosovichev, A.G., Schou, J., Rosenberg, W., Springer, L., Tarbell, T.D., Title, A., Wolfson, C.J., Zayer, I. MDI Engineering Team: 1995, *Solar Phys.* **162**, 129. ADS:1995SoPh..162..129S. doi:10.1007/BF00733429.
Schwarzschild, B.M.: 2012, *Phys. Today* **65**, 14.
Sofia, S., Heaps, W., Twigg, L.W.: 1994, *Astrophys. J.* **427**, 1048.
Sofia, S., Basu, S., Demarque, P., Li, L., Thuillier, G.: 2005, *Astrophys. J. Lett.* **632**, L147.
Sturrock, P.A.: 2009, *Solar Phys.* **254**, 227. ADS:2009SoPh..254..227S. doi:10.1007/s11207-008-9291-8.
Turck-Chièze, S., Couvidat, S.: 2011, arXiv:1009.0852v2, 24 June.
Turyshev, S.G., Shao, M., Nordtvedt, K.L., Dittus, H., Laemmerzahl, C., Theil, S., Salomon, C., Reynaud, S., Damour, T., Johann, U., Bouyer, P., Touboul, P., Foulon, B., Bertolami, O., Pàramos, J.: 2009, *Exp. Astron.* **27**, 27.
Ulrich, R.K., Hawkins, G.W.: 1981, *Astrophys. J.* **246**, 985. Erratum *Astrophys. J.* **249**, 831.
Williams, J.G., Turyshev, S.G., Boggs, D.H.: 2009, *Int. J. Mod. Phys. D* **18**, 1129.
Xu, Y., Yang, Y., Zhang, Q., Xu, G.: 2011, *Mon. Not. Roy. Astron. Soc.* **415**, 3335.

DOI 10.1007/978-1-4899-8005-2_11
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0186-3

Variability of Solar Five-Minute Oscillations in the Corona as Observed by the *Extreme Ultraviolet Spectrophotometer* (ESP) on the *Solar Dynamics Observatory/Extreme Ultraviolet Variability Experiment* (SDO/EVE)

L. Didkovsky · A. Kosovichev · D. Judge · S. Wieman · T. Woods

Received: 11 March 2012 / Accepted: 30 October 2012 / Published online: 23 November 2012

Abstract Solar five-minute oscillations have been detected in the power spectra of two six-day time intervals from soft X-ray measurements of the Sun observed as a star using the *Extreme Ultraviolet Spectrophotometer* (ESP) onboard the *Solar Dynamics Observatory* (SDO)/*Extreme Ultraviolet Variability Experiment* (EVE). The frequencies of the largest amplitude peaks were found to match the known low-degree ($\ell = 0-3$) modes of global acoustic oscillations within 3.7 μHz and can be explained by a leakage of the global modes into the corona. Due to the strong variability of the solar atmosphere between the photosphere and the corona, the frequencies and amplitudes of the coronal oscillations are likely to vary with time. We investigated the variations in the power spectra for individual days and their association with changes of solar activity, *e.g.* with the mean level of the EUV irradiance, and its short-term variations caused by evolving active regions. Our analysis of samples of one-day oscillation power spectra for a 49-day period of low and intermediate solar activity showed little correlation with the mean EUV irradiance and the short-term variability of the irradiance. We suggest that some other changes in the solar atmosphere,

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

L. Didkovsky (✉) · D. Judge · S. Wieman
Space Sciences Center, University of Southern California, Los Angeles, USA
e-mail: leonid@usc.edu

D. Judge
e-mail: judge@usc.edu

S. Wieman
e-mail: wieman@usc.edu

A. Kosovichev
Hansen Experimental Physics Laboratory, Stanford University, Stanford, USA
e-mail: akosovichev@solar.stanford.edu

T. Woods
Laboratory for Atmospheric and Space Physics, University of Colorado at Boulder, Boulder, USA
e-mail: tom.woods@lasp.colorado.edu

e.g., magnetic fields and/or inter-network configuration may affect the mode leakage to the corona.

Keywords Solar p-modes · Helioseismology · Solar extreme ultraviolet irradiance · Spectrophotometer

1. Introduction

The frequencies of individual resonant acoustic modes (Claverie *et al.*, 1979) that are excited by turbulent convection are stable within about 0.4 μHz on the time scale of the solar cycle because they correspond to intrinsic phase relations of resonant waves in the solar interior. Five-minute oscillations with frequencies centered at about 3.3 mHz are trapped below the solar photosphere. However, several observations in photospheric and chromospheric lines, in the UV pass-bands, and in the coronal Fe XVI line (33.5 nm) demonstrate some leakage of these oscillations into the upper layers of the solar atmosphere (*e.g.* Judge, Tarbell, and Wilhelm, 2001; O'Shea, Muglach, and Fleck, 2002; McIntosh, Fleck, and Judge, 2003; Muglach, 2003, and De Pontieu, Erdelyi, and James, 2004). This leakage may be explained by the increased amplitude of the oscillations caused by the rapid density decrease (*e.g.* Gough, 1993), and/or by interaction with the network magnetic elements, which can channel the photospheric acoustic power to higher atmospheric layers at frequencies below the cutoff (Vecchio *et al.*, 2007). For a discussion of these observations see Didkovsky *et al.* (2011).

Erdelyi *et al.* (2007) investigated the acoustic response to a single point-source driver. Malins and Erdelyi (2007) used a numerical simulation to show that widely horizontally coherent velocity signals from *p*-modes may cause cavity modes in the chromosphere and surface waves in the transition region, and that fine structures are generated extending from a dynamic transition region into the lower corona, even in the absence of a magnetic field.

A detection of the response of the corona to the observed photospheric low-degree ($\ell = 0-3$) *p*-modes was reported by Didkovsky *et al.* (2011). The authors studied the oscillation power spectra of two six-day-long time series using the soft X-ray band-pass from the *Extreme Ultraviolet Spectrophotometer* (ESP) (Didkovsky *et al.*, 2012) onboard the *Solar Dynamics Observatory* (SDO)/*Extreme Ultraviolet Variability Experiment* (EVE: Woods *et al.*, 2012). The largest amplitude peaks in the five-minute spectral region were compared with the low-degree photospheric *p*-modes observed in Doppler velocity from the *Birmingham Solar Oscillation Network* (BiSON: Chaplin *et al.*, 1998) and in visible-light intensity (red channel) from the SOHO/VIRGO instrument (Andersen, 1991; Fröhlich *et al.*, 1997). This comparison showed that the frequencies of the coronal oscillations may deviate from the frequencies determined from the photospheric observations. This can be explained with the significant influence of the non-uniform distribution of the irradiance sources and variability of the upper atmosphere (Didkovsky *et al.*, 2011). The mean standard deviation of the coronal frequencies from the photospheric *p*-mode frequencies was ≈ 3.7 μHz in the frequency range of 2.4 mHz to 3.6 mHz, which is about two times larger than the uncertainty of the peaks in the power spectrum determined from the six-day time-series. This deviation was also confirmed by comparing the power spectra for two consecutive six-day time series with a spectrum for the combined 12-day period. The power spectrum for a single six-day time series showed more significant peaks with better correspondence to the photospheric *p*-mode spectrum than the combined 12-day spectrum. However, it was not clear whether the observed coronal oscillations were transmitted but distorted photospheric *p*-modes, or

if these oscillations were excited in the corona by localized impulsive perturbation related to solar activity processes (*e.g.* modeling of Goode, Gough, and Kosovichev, 1992; Andreev and Kosovichev, 1995, 1998; Bryson, Kosovichev, and Levy, 2005). In this follow-up work we investigate whether the oscillations were excited in the corona by impulsive sources of solar activity (*e.g.* by flares) by studying the observational data during higher solar activity. If they are caused by solar activity, this study could reveal a correlation between the appearance of the coronal five-minute oscillations and this activity. In contrast to that, if the observed coronal oscillations (Didkovsky *et al.*, 2011) were related to the transmission of photospheric *p*-modes through the upper atmosphere, the observations made during the higher solar activity may reveal that solar-irradiance variability and significant increases of the soft X-ray irradiance during solar flares add some solar "noise" to the data time series, and low-amplitude oscillation peaks in the power spectra may be masked by these noise peaks.

In this work we study the variability of coronal five-minute oscillations by analyzing 49 one-day power spectra for various solar activity observing conditions that range from the lowest solar activity observed during the SDO mission in the middle of May 2010 to intermediate solar-activity levels in 2011. As in Didkovsky *et al.* (2011), we use soft X-ray observations without spatial resolution in the zeroth-order channel of SDO/EVE/ESP.

2. SDO/EVE/ESP Channels

EVE (Woods *et al.*, 2012) is one of three instrument suites on SDO. It provides solar EUV-irradiance measurements that are unprecedented in terms of spectral resolution, temporal cadence, accuracy, and precision. Furthermore, the EVE program will incorporate physics-based models of solar EUV irradiance to advance the understanding of solar dynamics based on short- and long-term activity of solar magnetic features. ESP (Didkovsky *et al.*, 2012) is one of five channels in the EVE suite. It is an advanced version of the SOHO/CELIAS/SEM (Hovestadt *et al.*, 1995; Judge *et al.*, 1998) and is designed to measure solar EUV irradiance in four first-order bands of the diffraction grating centered around 19 nm, 25 nm, 30 nm, and 36 nm, as well as in a soft X-ray band from 0.1 to 7.0 nm (the energy range is 0.18 to 12.4 keV) in the zeroth-order of the grating. Each band's detector system converts the photo-current into a count rate (frequency). The count rates are integrated over 0.25-second increments and transmitted to the EVE Science and Operations Center for data processing. An algorithm for converting the measured count rates into solar irradiance and the ESP calibration parameters is described by Didkovsky *et al.* (2007, 2012).

3. Observations

Our analysis of the ESP measurements was based on datasets for a long series of observations covering a much wider range of solar activity compared to the two six-day time intervals analyzed by Didkovsky *et al.* (2011). Due to the high sensitivity of the ESP zeroth-order soft X-ray signal to solar activity and because of significant contamination of the power spectra in the five-minute band by the impulsive increases in irradiance during solar flares, we used data for the periods of low-to-intermediate solar activity, without strong solar flares. Based on these conditions, five data intervals were chosen (Table 1, Figure 1). Figure 1 shows these (thick horizontal bars) on the background of variations of soft X-ray solar irradiance related to the increased phase of solar activity cycle. Intervals one to

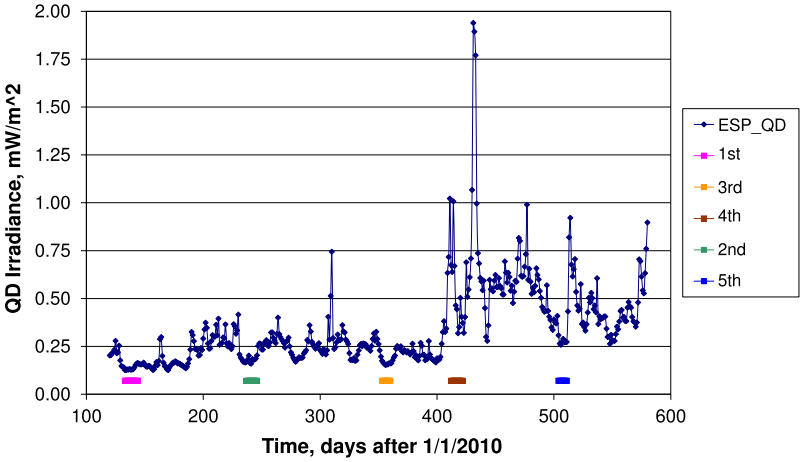


Figure 1 ESP zeroth-order (0.1 to 7.0 nm) quad-diode (QD) soft X-ray irradiance measured for the first 461 days of the SDO mission, 245 days in 2010 and 216 days in 2011. The time intervals chosen are shown as wide horizontal bars below the irradiance curve. The mean number of analyzed days for each time interval was about ten, with the total number of days equal to 49. The irradiance curve represents daily mean values while the actual cadence is 0.25 seconds.

three were chosen to represent the lowest periods of solar irradiance observed by SDO in 2010, which are 0.137, 0.181, and 0.160 mW m^{-2}. The fourth time interval was chosen between two periods of relatively high solar activity (Figure 1) with a mean solar irradiance of 0.531 mW m^{-2}. The fifth time interval represents a return to relatively low solar activity with a mean irradiance of 0.276 mW m^{-2}. Thus, these five time intervals cover a wide range of solar conditions for the periods of decreased solar activity. To establish a point of reference with our previous analysis, the first time interval matches the two six-day intervals analyzed previously (Didkovsky *et al.*, 2011). Three of the time intervals (the second, third, and fifth) correspond to the lower-irradiance "spots" on the irradiance curve (Figure 1) with decreased solar activity, while the fourth time interval includes some C- and M-class solar flares to include some stronger disturbances of the solar atmosphere. Table 1 shows some details of the five-interval database. The lowest daily mean irradiance since the start of the SDO mission was 0.126 mW m^{-2} on 2010 DOY 133, the lowest daily standard deviation (STD) was 3.31×10^{-3} mW m^{-2} on 2010 DOY 135, the highest flare-related mean irradiance value was 0.674 mW m^{-2} on 2011 DOY 047, and the highest flare-related STD of 0.505 mW m^{-2} was detected on 2011 DOY 055 (see Table 2).

4. Data Reduction

The data reduction was based on the ESP zeroth-order time series, whose original (level 0D) effective count rate (counts s^{-1}) was corrected for energetic-particle events and temperature changes of dark counts (Didkovsky *et al.*, 2012). The data were interpolated to eliminate short gaps (about two minutes total) that occur when the filters in the ESP filter wheel and observing modes change during routine daily calibration. The first step was to calculate a power spectrum for each of the 49 days. Then, a power-law curve for each spectrum was determined in a manner similar to that described by Didkovsky *et al.* (2011) for the best fit

Table 1 Some details of the datasets. Columns (left to right) are time intervals, range of days within each time interval, days with GOES class B or larger solar flares, and solar flare class.

Time interval	Range of days	Days with solar flares $\geq B1$	GOES solar flare class
1	13 – 24 May 2010	13 May 2010	B2.8
		22 May 2010	B1.3
		23 May 2010	B1.4, B1.3, B1.1
		24 May 2010	B1.1
2	25 Aug – 3 Sep 2010	27 Aug 2010	B2.2, B1.6, B2.3, B4.7, B2.2, B1.9, B1.3, B1.1, B2.2, B1.9, B1.3, B1.1
		28 Aug 2010	B1.0
		29 Aug 2010	B1.7, B1.4
		30 Aug 2010	B1.4, B1.8
		3 Sep 2010	B2.8, B1.2
3	19 – 26 Dec 2010	19 Dec 2010	B1.1
4	16 – 26 Feb 2011	16 Feb 2011	C2.0, M1.0, C2.2, C5.9, C2.2, M1.1, C9.9, C3.2, C1.0, M1.6, C7.7, C1.3, C1.1, C4.2, C2.8
		17 Feb 2011	C6.1, C2.3, C1.5, C1.0
5	19 – 26 May 2011	19 May 2011	
		20 May 2011	B2.0, B3.5, B3.0, B2.2
		21 May 2011	B1.6, B1.1, B2.7, B1.5
		22 May 2011	B1.3, B1.2
		23 May 2011	B1.9
		24 May 2011	B2.0
		25 May 2011	B1.9, B4.1, C1.1, C1.4, B3.0
		26 May 2011	B1.3, B4.4, B1.9

of the spectrum in the range of frequencies from 2.0 mHz to 10.0 mHz, which includes our range of interest between 2.4 mHz and 4.0 mHz,

$$I1_i(f) = A_i \times f^{-n_i}, \tag{1}$$

where $I1$ is an array for the power-law spectral density, i is the day number, A_i is a constant, f is the frequency, and n_i is the power-law index. The third step was to calculate a running mean $[RM]$ curve that represents the local power change (local meaning in the five-minute band, for instance) in the power spectrum. Because we used one-day power spectra with a relatively low (11.6 μHz) frequency resolution and consequently a low confidence in the frequencies of individual peaks, the running-mean window of integration was chosen to be 13 seconds to reduce the influence of individual peaks in the power spectrum, but preserve the whole power change within the range of the power-law curve,

$$I2_i(f) = RM_i, \tag{2}$$

where $I2$ is an array that represents the running-mean function. This is the standard IDL procedure, *i.e.* MEDIAN. The final step was to calculate the mean sum $[S_i]$ of the ratios

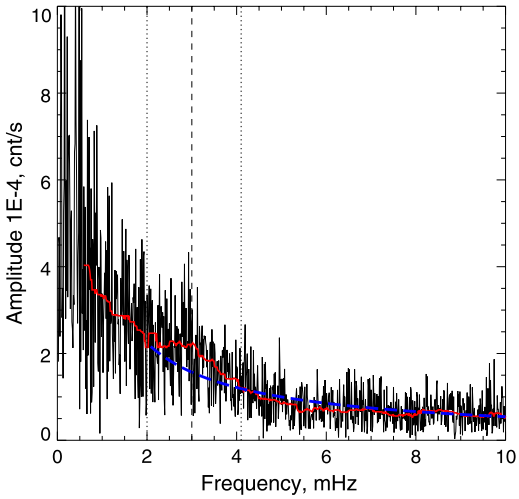


Figure 2 ESP zeroth-order (0.1 to 7.0 nm) power spectrum for 13 May 2010 (2010 DOY 133). The dashed (blue) line is a power-law curve ($I1 = 0.18 f^{-0.88}$), the red line ($I2$) is a running-mean curve. The two dotted vertical lines represent the edges of the power increase and the dashed vertical line shows the frequency of the increase maximum (see Table 4). The S-ratio for this example is 1.23.

between the $I2_i$ and $I1_i$ over the frequency range of five-minute oscillations ($f_1 = 2.4$ mHz to $f_2 = 4.0$ mHz),

$$S_i = \frac{\sum_{f_1}^{f_2} \frac{I2_f}{I1}}{m}, \tag{3}$$

where m is the number of frequency bins within the frequency range. If the power spectrum showed an increase in the five-minute band, S_i was greater than unity. As an example of this data-reduction algorithm, Figure 2 shows a power spectrum for 13 May 2010 with $I1$ as a dashed (blue) line and $I2$ as a red line. In addition to the power increase in the five-minute region, some peaks with frequencies above the cut-off frequency of 5.5 mHz demonstrate increased amplitudes without the clear indication of increased power represented by the red line (Figure 2). For some other days the power increase in the frequency region from 6 to 10 mHz is clearly observed. This increase may be similar to that found by Gurman *et al.* (1982) in the frequency range from 5.8 to 7.8 mHz using SMM observations. An example of these increases in the five-minute region and above the cut-off frequency is shown in Figure 3.

5. Results

The results of the five-minute S-ratio calculations for the one-day power spectra are shown in Table 2 along with the changes of spectral irradiance, fluctuations of this irradiance (STD), and the maximum amplitude of the filter curve $I2$ (Equation (2)) in the five-minute range. Columns (one to five) are days of observations in the (YYYY DOY) format, the daily mean irradiance [10^{-1} mW m^{-2}], the daily standard deviation [10^{-3} mW m^{-2}], the ratio [S], and the filter $I2$ maximum amplitude in the five-minute range [10^{-4} counts s^{-1}]. Each of the five time intervals is separated by an empty row. The S-ratio column in Table 2 is marked as N/A if the change of the spectral density in the power spectrum shows strong low-frequency variations that could be caused by strong changes of solar irradiance related to a solar flare, for instance. The gaps in the $I2$ column are either for days in which the power spectra do not show any increase in the five-minute region (the S-ratio is < 1) or for days where the

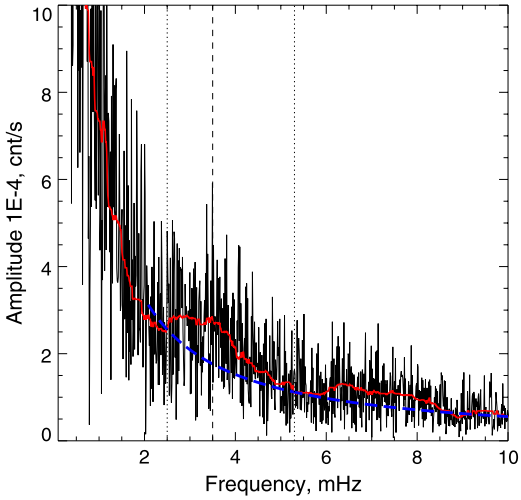


Figure 3 ESP zeroth-order (0.1 to 7.0 nm) power spectrum for the 23 May 2010 (DOY 142). The dashed line is a power-law curve ($I1 = 7.010^{-4} f^{-1.1}$), the red line ($I2$) is a running-mean curve. The two dotted vertical lines represent the edges of the power increase and the dashed vertical line shows the frequency corresponding to the increase maximum (see Table 4). The S-ratio for this example is 1.37. Power increases above the cut-off frequency, at about 6 – 8 mHz, are more clearly visible than in Figure 2. To facilitate the comparison with Figure 2, the vertical scale is the same as in that figure.

Table 2 Comparison of the S-ratios (Equation (3)) in the five-minute band with the daily mean soft X-ray irradiance, the standard deviation of this irradiance (STD), and the filter $I2$ maximum amplitude (Equation (2)).

Day [YYYY DOY]	Daily mean, irradiance	Standard deviation	S	$I2$
2010 133	1.26	7.54	1.23	2.3
2010 134	1.26	3.83	1.07	0.9
2010 135	1.29	3.31	1.14	1.4
2010 136	1.28	3.40	0.87	
2010 137	1.32	3.99	1.12	1.1
2010 138	1.28	3.47	1.06	1.1
2010 139	1.29	3.67	1.05	1.5
2010 140	1.30	3.90	1.07	0.9
2010 141	1.39	7.31	1.00	1.5
2010 142	1.52	9.58	1.37	2.9
2010 143	1.60	18.4	1.03	2.5
2010 144	1.63	11.7	1.01	2.5
2010 237	1.69	19.50	0.99	
2010 238	1.81	18.40	0.98	
2010 239	2.01	27.2	N/A	
2010 240	1.63	8.09	1.11	2.4
2010 241	1.59	6.68	1.05	2.4
2010 242	1.83	11.5	1.07	1.9
2010 243	1.80	6.96	0.98	
2010 244	1.81	5.43	1.03	1.7
2010 245	1.86	8.87	1.08	2.4
2010 246	2.03	13.6	1.02	3.1
2010 353	1.74	6.14	0.98	
2010 354	1.64	6.24	0.95	
2010 355	1.58	3.98	0.96	

Table 2 (*Continued.*)

Day [YYYY DOY]	Daily mean, irradiance	Standard deviation	S	$I2$
2010 356	1.52	3.41	0.78	
2010 357	1.55	4.47	0.95	
2010 358	1.56	4.91	0.88	
2010 359	1.62	5.09	N/A	
2010 360	1.61	4.68	0.90	
2011 047	6.74	288.0	1.24	225
2011 048	6.38	147.0	0.95	
2011 049	1.01	504.0	N/A	
2011 050	6.70	178.0	1.16	108
2011 051	4.64	75.9	0.97	
2011 052	4.44	112.0	N/A	
2011 053	3.18	25.1	0.96	
2011 054	3.50	42.2	0.93	
2011 055	5.03	505.0	N/A	
2011 056	4.05	36.9	N/A	
2011 057	3.72	60.6	1.08	28.4
2011 139	3.08	30.8	N/A	
2011 140	2.65	9.3	0.95	
2011 141	2.58	6.37	N/A	
2011 142	2.72	9.66	N/A	
2011 143	2.89	10.6	N/A	
2011 144	2.72	12.8	1.02	3.3
2011 145	2.73	32.1	N/A	
2011 146	2.73	20.1	N/A	

S-ratio in the S-ratio column is marked as N/A because of strong contamination from solar flares. As Table 2 shows, the largest number of these contaminated spectra corresponds to the fourth and fifth time intervals with significantly higher solar activity (see the STD in Table 2) than during the first three time intervals.

5.1. How Solar Activity Affects the Oscillations in the Corona

To analyze how the oscillations in the five-minute band represented by the S-ratios (Table 2) and by the amplitude of the filter curve $I2$ (Equation (2)) are related to the changes of the observing conditions, two parameters of these conditions, the daily mean irradiance and standard deviation of the irradiance, were compared with the changes of the S-ratios and the maximum amplitude of filter $I2$. Table 2 shows that S-ratios higher than unity are detected for the first, second, and fourth time series. Figures 4 – 6 show these time series. Column $I2$ in Table 2 for the amplitude of the filtered curve and the open circles in Figures 4, 5, 6 do not show any negative amplitude compared to the power-law curve (Equation (1)) for the days for which the S-ratio is lower than unity. The filtered curve $I2$ (Equation (2)) for these days is below the spectral density $I1$ (Equation (1)) and represents only noise. The S-ratios (Table 2) are plotted according to the linear scale shown at the right-hand side of the plots.

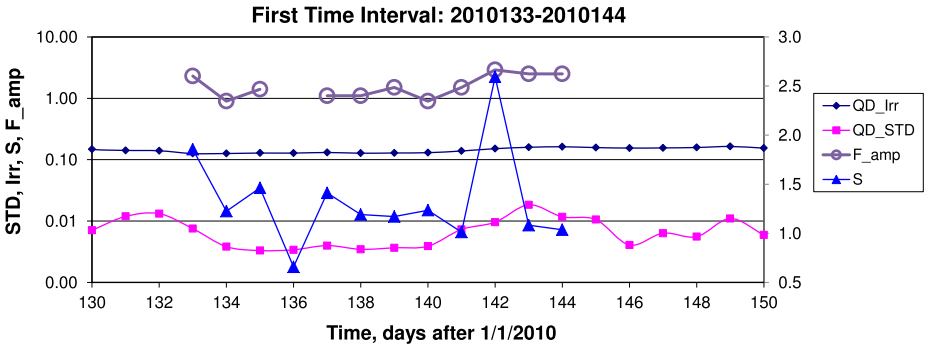


Figure 4 Time series for the first time interval used to analyze the correlation between the five-minute oscillation ratio (triangles), the daily mean irradiance (diamonds), and the STD of the irradiance (squares). Some of the positive correlations, *e.g.* between days 133 – 134 and 141 – 142, are changing to anti-correlations, for instance 140 – 141 and 142 – 143. The filter $I2$ (Equation (2)) maximum amplitudes (open circles) within the five-minute range are shown for the increases of the ratio $S > 1.0$ (right scale). The gap in the maximum amplitude of filter $I2$ (open circles) for day 136 in Figure 4 and other gaps in Figures 5 and 6 are related to the ratios with $S < 1.0$ (Table 2, S-ratio) for which there is no power increase. These spectra show the noise in the five-minute region.

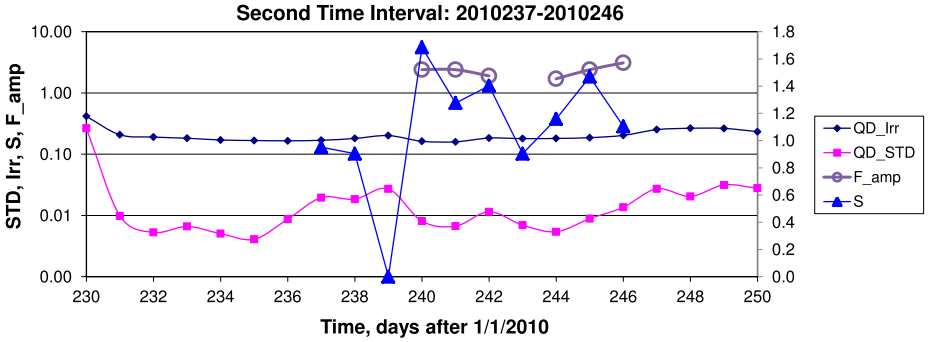


Figure 5 Time series for the second ten-day time interval (2010 DOY 237 – 246) used to analyze the correlation between the five-minute oscillation ratio (triangles), the daily mean irradiance (diamonds), and the STD of the irradiance (squares). Some positive correlations, *e.g.* between days 237 – 238, 240 – 241, and 242 – 243, are changing to anti-correlations, for instance 238 – 239, 243 – 244, and 245 – 246. The filter $I2$ (Equation (2)) maximum amplitudes (open circles) within the five-minute range are shown for the increases of the ratio $S > 1.0$ (right scale). The gaps in the maximum amplitude of filter $I2$ (open circles) for days 237 – 239, and 243 in Figure 5 are related to the ratios with $S < 1.0$ (Table 2, S-ratio), for which there are no power increases. These spectra show the noise in the five-minute region.

Table 3 summarizes data from Table 2 for a more detailed comparison of the observing conditions.

The technique used for this analysis is based on a comparison of spectral amplitudes in the five-minute range of the power spectra. This technique is very sensitive to the contamination of the spectra by flare-related increases of the irradiance. These increases affect both the daily mean solar irradiance and the standard deviation [STD] of this irradiance. If the local five-minute increase in the power spectrum is a result of this contamination (flare-related solar noise), one should expect a positive correlation between the S-ratio (or maximum amplitude of filter $I2$) and the flare-related increases of the irradiance and its STD. Thus, the

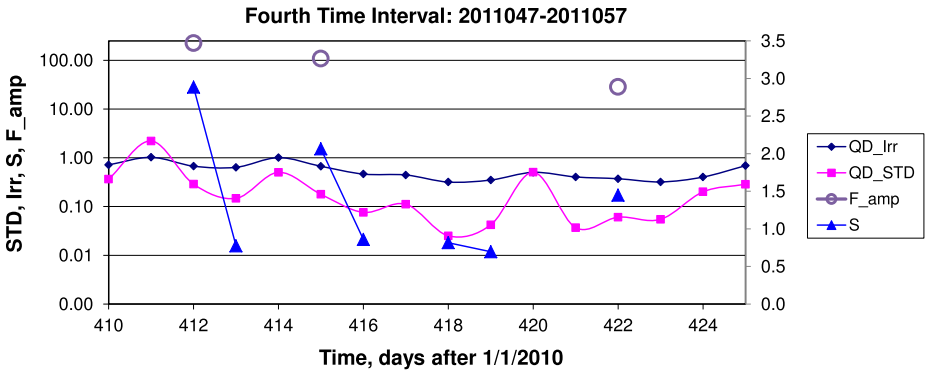


Figure 6 Time series for the fourth 11-day time interval (2011 DOY 47 – 57) shown as days 412 – 422 used to analyze the correlation between the five-minute oscillation ratio (triangles), the daily mean irradiance (diamonds), and the STD of the irradiance (squares). Some positive correlations, *e.g.* between days 412 – 413 and 415 – 416, are changing to anti-correlations, for instance 418 – 419. The filter *I*2 (Equation (2)) maximum amplitudes (open circles) within the five-minute range are shown for the increases of the ratio $S > 1.0$ (right scale). The gaps in the maximum amplitude of filter *I*2 (open circles) for days 413 – 414 and 416 – 421 in Figure 6 are related to the ratios with $S < 1.0$ (Table 2, *S*-ratio) for which there is no power increase. These spectra show noise in the five-minute region.

Table 3 A comparison of the *S*-ratios in the five-minute band and their variations with the mean irradiance for each of the five time intervals.

Days [YYYY DOY]	Mean irradiance 10^{-1} [mW m^{-2}]	Mean STD 10^{-3} [mW m^{-2}]	Mean *S*-ratio
2010 133 – 144	1.37 ± 1.35	6.68 ± 4.69	1.09 ± 0.12
2010 237 – 246	1.81 ± 1.45	12.6 ± 7.07	1.03 ± 0.05
2010 353 – 360	1.60 ± 0.07	4.86 ± 0.97	0.92 ± 0.07
2011 047 – 057	5.31 ± 2.02	180 ± 178	1.04 ± 0.12
2011 139 – 146	2.76 ± 0.15	16.5 ± 10.1	0.99 ± 0.05

technical goal of this analysis was to extract and compare this information from the daily spectra for different levels of solar activity.

5.1.1. Maximum Amplitude of Filter I2

Figures 4 – 6 show the maximum amplitude of filter *I*2 as open circles (see also last column in Table 2). For solar activity minimum conditions, the oscillation peaks in the power spectra are not masked by noise, and we assume that the maxima in the amplitude of filter *I2* for the first and second time intervals show the amplitudes of the five-minute oscillations in the corona related to the photospheric *p*-modes. For these conditions the mean amplitude is similar for the first and second time intervals: 1.69 and 2.16×10^{-4} counts s^{-1}. The fourth time interval shows significantly higher amplitudes (Figure 6 and Table 2) that indicate the contamination of the spectra by much higher solar-flare activity. For low solar-activity periods, the maximum amplitude of filter *I*2 is just another representation of the *S*-ratio, see Figure 4.

5.1.2. Daily Mean Irradiance

Table 2 and Figures 4 – 6 show that the daily mean soft X-ray irradiance is a significant source of the *S*-ratio change. Table 3 shows that the highest mean *S*-ratio of the oscillations

in the five-minute band was detected for the first time interval with the lowest daily mean irradiance. The cross-correlation between the changes of the irradiance and the S-ratios for the first time interval is low at 0.1. For the second time interval it becomes negative, -0.27, which indicates that an increase of irradiance (1.81 compared to 1.37) leads to a decrease of S-ratios for the observed five-minute oscillations in the corona. For the fourth time interval, the correlation is positive and high at 0.64. We interpret this as a contamination of the power spectra by solar-flare events. The amplitudes of the filter-curve maxima in Figures 4 – 6 (open circles) and the last column in Table 2 demonstrate such a flare-related increase. Because the spectral contamination is the result of the solar-flare low-frequency power transfer to the other frequency regions of the spectrum, including the five-minute region, we assume that it leads to the strong positive correlation detected for the fourth time interval.

5.1.3. Standard Deviation

The soft X-ray signal that ESP detects in the zeroth-order channel is a very sensitive probe of solar variability. Assuming that the observed five-minute oscillations in the corona represent a response of the corona to the photospheric acoustic modes and that the "transmission" of the solar atmosphere is a function of various disturbances and inhomogeneities between the photosphere and the corona, we can treat the standard deviation [STD] as an indicator of this solar "noise" in the five-minute oscillation signal. However, our results indicate that STD is not a unique parameter for estimating this transmission. For example, the cross-correlation between the S-ratio and STD for the first and second time intervals, 0.1 and -0.44, is either low or negative. For the third time interval for which STD was the lowest (Table 3), the S-ratios were all < 1.0 (Table 2). This may indicate that as suggested by several authors, *e.g.* Judge, Tarbell, and Wilhelm (2001), O'Shea, Muglach, and Fleck (2002), McIntosh, Fleck, and Judge (2003), Muglach (2003), and Vecchio *et al.* (2007), the connectivity between the photosphere and the corona depends on the configuration of the magnetic fields, which may also be a function of the STD. If the lowest STD during the third time interval is related to the decreased magnetic-field strength and to an ineffective configuration of the network, this may explain the absence of the power increases in the five-minute range of oscillations. The relatively high negative correlation (-0.44) for the second time interval with an STD about two times higher compared to the first time interval allows us to consider that S-ratios for the second time interval as well as for the first time interval were not the result of the spectral contamination in the power spectra. This evidence is consistent with another independent confirmation of the response of the corona to photospheric p-modes, *e.g.* based on the spectral analysis of the two six-day time series (Didkovsky *et al.*, 2011).

5.1.4. A Shift of the Maximum Frequency as a Function of Activity

We analyzed the oscillation spectra with the most significant power increases ($S \geq 1.1$, see Table 2, S) in the five-minute range to investigate the correlation between the frequency of the maximum of this increase and solar activity (see Table 2, standard deviation). In addition to this correlation, the ranges of the increases were also analyzed. Table 4 and Figure 5 show the results of this analysis.

Figure 7 shows a positive correlation between the increases of the STD and the shift of the maximum frequency: $R_1 = 0.62$. The same correlation ($R_2 = 0.66$) is found between the changes of STD and the mean frequency of the increase. Certainly, the significance of these correlations is low. Using the t-distribution, we find that $t = 1.4$, which is lower

Table 4 A comparison of the STD, the frequency maximum, and the ranges of power increases for parameter $S \geq 1.1$.

Days [YYYY DOY]	STD of the irradiance [10^{-6} W m^{-2}]	Max frequency [mHz]	Range of frequency [mHz]	Mean frequency [mHz]
2010 133	7.54	3.0	2.0 – 4.1	3.05
2010 135	3.31	2.4	2.1 – 4.1	3.1
2010 137	3.99	3.0	2.6 – 3.5	3.05
2010 142	9.58	3.5	2.5 – 5.3	3.9
2010 240	8.09	2.7	2.1 – 4.4	3.25
2011 047	288	3.6	2.4 – 5.4	3.9

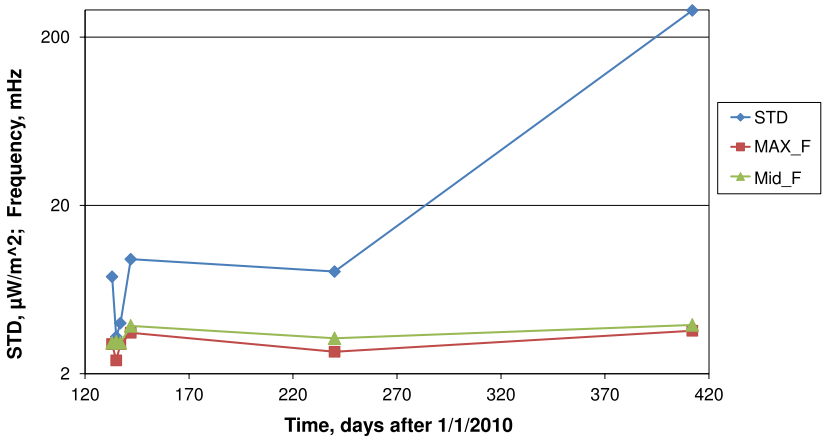


Figure 7 Comparison of the solar activity changes [STD: diamonds] with the changes of the frequency maximum increase [squares], and the mean frequency of the increase range [triangles]. Data for the power spectra are shown with increases of $S \geq 1.1$, see Table 2, S-ratio.

than the t-distribution number of 1.638 for three degrees of freedom; the significance is 0.1 (one tail). Two days with the highest STD show a significant shift of the right edge of the frequency range toward the cut-off frequency, see 2010 DOY 142 (5.3 mHz) and 2011 DOY 47 (5.3 mHz) in Table 4. This shift is consistent with the model proposed by Bryson, Kosovichev, and Levy (2005).

6. Concluding Remarks

Global solar oscillations in the five-minute range were initially detected in the corona using soft X-ray irradiance measurements from SDO/EVE/ESP (Didkovsky *et al.*, 2012). The variability of the five-minute S-ratios (Equation (3)) in the power spectra that characterize the strength of oscillations relative to noise was analyzed for five time intervals with low- and intermediate solar activity during a 379-day period (2010 DOY 133 – 2011 DOY 146) for 49 days.

The results of this analysis show that the best conditions for observing five-minute solar oscillations in the corona are at low solar activity. In the first time interval, 11 of the 12 days show power increases with an estimated S-ratio ≥ 1.0.

Our analysis shows that power increases in the spectra of coronal oscillations, interpreted as the response of the corona to photospheric p-modes due to their channeling (Vecchio *et al.*, 2007) or leakage (Erdelyi *et al.*, 2007; Malins and Erdelyi, 2007), are not related to an increased daily mean solar irradiance. This is clear from the S-ratio comparison between the first time interval and the other four time intervals, see Table 3. The higher daily mean irradiances observed for the second through fifth time intervals did not lead to a higher S-ratio.

The power increases in the oscillation spectra are not caused by the increases in the mean STD (third column in Table 3) either. This conclusion is based on the analysis of cross-correlations between the S-ratio and the mean STD. The correlation is low for all time intervals except the fourth. This result confirms that detected increases of the S-ratios in the five-minute range are not created by the spectral contamination in the power spectra from solar flares and are not instrumental or data-reduction artifacts, but represent the leakage of photospheric p-modes to the corona.

We interpret the significant positive correlation for the fourth time interval as an artifact and a reflection of much higher solar activity with an STD about 27 times higher than for the first time interval. The amplitude maxima of filter $I2$ for the fourth time interval are significantly larger than the amplitudes for the first and second time intervals and may be a demonstration of the spectral contamination in the five-minute region by the low-frequency flare signals. This is also clear from helioseismology results, which have shown shifts in the frequencies of the global modes related to the solar cycle, but not in the amplitudes.

Another conclusion based on these results is that large-scale solar oscillations detected in the corona are related to the leakage of the photospheric acoustic oscillations rather than to the excitation of these coronal oscillations by solar energetic events.

The high sensitivity of the five-minute oscillations in the corona to the changes of solar irradiance may be used as a diagnostic tool to characterize the connectivity in soft X-ray irradiance dynamics.

Acknowledgements This work was partially supported by the University of Colorado award 153-5979. Data courtesy of NASA/SDO and the EVE science team.

References

Andersen, B.N.: 1991, *Adv. Space Res.* **11**, 93.

Andreev, A.S., Kosovichev, A.G.: 1995, In: Hoeksema, J.T., Domingo, V., Fleck, B., Battrick, B. (eds.) *SOHO 4 – Helioseismology* **SP-376**, ESA, 471.

Andreev, A.S., Kosovichev, A.G.: 1998, In: Wilson, A. (ed.) *SOHO 6/GONG – Structure and Dynamics of the Interior of the Sun and Sun-like Stars* **SP-418**, ESA, Noordwijk, 87.

Bryson, S., Kosovichev, A., Levy, D.: 2005, *Physica D* **201**, 1.

Chaplin, W.J., Elsworth, Y., Isaak, G.R., Lines, R., McLeod, C.P., Miller, B.A., New, R.: 1998, *Mon. Not. Roy. Astron. Soc.* **300**, 1077.

Claverie, A., Isaak, G.R., McLeod, C.P., Van Der Raay, H.B., Roca Cortes, T.: 1979, *Nature* **282**, 591.

De Pontieu, B., Erdelyi, R., James, S.P.: 2004, *Nature* **430**, 536.

Didkovsky, L.V., Judge, D.L., Wieman, S., Harmon, M., Woods, T., Jones, A., Chamberlin, P., Woodraska, D., Eparvier, F., McMullin, D., Furst, M., Vest, R.: 2007, In: Fineschi, S., Viereck, A.R. (eds.) *Solar Physics and Space Weather Instrumentation II*, *SPIE* **6689**, 6689OP1.

Didkovsky, L., Judge, D., Kosovichev, A., Wieman, S., Woods, T.: 2011, *Astrophys. J. Lett.* **738**, L7.

Didkovsky, L., Judge, D., Wieman, S., Woods, T., Jones, A.: 2012, *Solar Phys.* **275**, 179.

Erdelyi, R., Malins, C., Toth, G., De Pontieu, B.: 2007, *Astron. Astrophys.* **467**, 1299.
Fröhlich, C., Andersen, B.N., Appourchaux, T., Berthomieu, G., Crommelynck, D.A., Domingo, V., Fichot, A., Finsterle, W., Gomez, M.F., Gough, D., Jemenez, A., Leifsen, T., Lombaerts, M., Pap, J.M., Provost, J., Cortes, T.R., Romero, J., Roth, H., Sekii, T., Telljohann, U., Toutain, T., Wehrli, C.: 1997, *Solar Phys.* **170**, 1. ADS:1997SoPh..170....1F, doi:10.1023/A:1004969622753.
Goode, P.R., Gough, D., Kosovichev, A.G.: 1992, *Astrophys. J.* **387**, 707.
Gough, D.O.: 1993, In: Zahn, J.-P., Zinn-Justin, J. (eds.) *Astrophysical Fluid Dynamics – Les Houches Session XLVII, 1987*, Elsevier, Amsterdam, 399.
Gurman, J.B., Leibacher, J.W., Shine, R.A., Woodgate, B.E., Henze, W.: 1982, *Astrophys. J.* **253**, 939.
Hovestadt, D., Hilchenbach, M., Burgi, A., Klecker, B., Laeverenz, P., Scholer, M., Grunwaldt, H., Axford, W.I., Livi, S., Marsch, E., Wilken, B., Winterhoff, H.P., Ipavich, F.M., Bedini, P., Coplan, M.A., Galvin, A.B., Gloeckler, G., Bochsler, P., Balsiger, H., Fisher, J., Geiss, J., Kallenbach, R., Wurz, P., Reiche, K.-U., Gliem, F., Judge, D.L., Ogawa, H.S., Hsieh, K.C., Mobius, E., Lee, M.A., Managadze, G.G., Verigin, M.I., Neugebauer, M.: 1995, *Solar Phys.* **162**, 441. ADS:1995SoPh..162..441H, doi:10.1007/BF00733436.
Judge, D.L., McMullin D, R., Ogawa, H.S., Hovestadt, D., Klecker, B., Hilchenbach, M., Mobius, E., Canfield, L.R., Vest, R.E., Watts, R., Tarrio, C., Kuhne, M., Wurz, P.: 1998, *Solar Phys.* **177**, 161. ADS:1998SoPh..177..161J, doi:10.1023/A:1004929011427.
Judge, P.G., Tarbell, T.D., Wilhelm, K.: 2001, *Astrophys. J.* **554**, 424.
Malins, C., Erdelyi, R.: 2007, *Solar Phys.* **246**, 41. ADS:2007SoPh..246...41M, doi:10.1007/s11207-007-9073-8.
McIntosh, S.W., Fleck, B., Judge, P.G.: 2003, *Astron. Astrophys.* **405**, 769.
Muglach, K.: 2003, *Astron. Astrophys.* **401**, 685.
O'Shea, E., Muglach, K., Fleck, B.: 2002, *Astron. Astrophys.* **387**, 642.
Vecchio, A., Gauzzi, G., Reardon, K.P., Janssen, K., Rimmele, T.: 2007, *Astron. Astrophys.* **461**, L1.
Woods, T.N., Eparvier, F.G., Hock, R., Jones, A., Woodraska, D., Judge, D., Didkosky, L., Lean, J., Mariska, J., Warren, H., McMullin, D., Chamberlin, P., Berthiaume, G., Bailey, S., Fuller-Rowell, T., Soika, J., Tobiska, W.K., Viereck, R.: 2012, *Solar Phys.* **275**, 115. ADS:2012SoPh..275..115W, doi:10.1007/s11207-009-9487-6.

DOI 10.1007/978-1-4899-8005-2_12
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0198-z

Solar Dynamo and Toroidal Field Instabilities

Alfio Bonanno

Received: 1 March 2012 / Accepted: 9 November 2012 / Published online: 7 December 2012

Abstract The possibility of non-axisymmetric (kink) instabilities of a toroidal field seated in the tachocline is much discussed in the literature. In this work, the basic properties of kink and quasi-interchange instabilities, produced by mixed toroidal and poloidal configuration, will be briefly reviewed. In particular, it will be shown that the unstable modes are strongly localized near the Equator and not near the Poles as often claimed in the literature. Based on the results of recent numerical simulations, it is argued that a non-zero helicity can already be produced at a non-linear level. A mean-field solar dynamo is then constructed with a *positive* α-effect in the overshoot layer localized near the Equator, and a meridional circulation with deep return flow. Finally, the possibility that the solar cycle is driven by an $\alpha\Omega$ dynamo generated by the negative subsurface shear in the supergranulation layer will also be discussed.

Keywords Solar dynamo · Magnetic fields

1. Introduction

An essential ingredient of any dynamo model is the α-effect, where α is the transport coefficient of the closure relation for the turbulent electromagnetic force. Owing to its pseudo-scalar nature this term represents a likely possibility to produce a dynamo action with an axially symmetric field.

From the theoretical point of view it is expected that both the kinetic helicity $[\langle \mathbf{v} \cdot \nabla \times \mathbf{v} \rangle]$ and the current helicity $[\langle \mathbf{b} \cdot \nabla \times \mathbf{b} \rangle]$ produced by the velocity $\mathbf{v}$ and magnetic field $\mathbf{b}$ fluctuations contribute to this term (Gruzinov and Diamond, 1994), although the relative importance of these terms is still a subject of debate (Silant'ev, 2000). On physical grounds

Invited talk.

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

A. Bonanno (✉)
INAF, Osservatorio Astrofisico di Catania, Via S.Sofia 78, 95123 Catania, Italy
e-mail: alfio.bonanno@inaf.it

it is conceivable that in the bulk of the convection zone "cyclonic" turbulence (Parker, 1955) could be an efficient mechanism to produce a turbulent dynamo whose α-effect is dominated by the kinetic helicity (Steenbeck, Krause, and Rädler, 1966).

The discovery, due to the interpretation of helioseismic data (Schou, 1991), that the stable stratified region below the convection zone is characterized by the presence of a strong horizontal shear (Kosovichev, 1996) has suggested the possibility of an α-effect located just beneath the convection zone (Parker, 1993) in the region called the tachocline (Spiegel and Zahn, 1992). This fact has opened the door to the possibility that the source of the α-effect can have a magnetic origin, which can thus be attributed to various possible MHD instabilities (Ferriz-Mas, Schmitt, and Schüssler, 1994; Dikpati and Gilman, 2001; Chatterjee *et al.*, 2011). More recently, Bonanno and Urpin (2011, 2012) have further clarified a few aspects of kink and quasi-interchange instabilities in a stably stratified plasma. While kink instabilities are generated from a pure toroidal field, quasi-interchange instabilities are produced by a mixed combination of the poloidal and toroidal fields and their spectrum can be rather different from simple kink waves. It has also been shown that kink modes in spherical geometry are more effective near the Equator, a result that supports the possibility of producing a non-zero α-effect at low latitudes. In fact, although the presence of a poloidal field breaks the symmetry and creates a preferred helicity in the turbulence flow, Del Sordo *et al.* (2011) recently argued that kink instabilities alone can produce a preferred chirality in the turbulent plasma due to a symmetry-breaking effect at non-linear level.

In this article, after reviewing the basic properties of kink and quasi-kink instabilities in a stably stratified plasma, the possibility of explicitly constructing a solar dynamo with an α-effect localized at the bottom of the convective zone and at low latitudes will be discussed and compared to mean-field models where the α-effect is instead localized near the surface layer, as proposed by Brandenburg (2005).

2. Quasi-Interchange Instabilities in a Stably Stratified Plasma

The stability of a stably stratified column of plasma in the ideal MHD limit is the central problem of most of the controlled-fusion literature. In this context, the energy principle (Bernstein *et al.*, 1958) has extensively been used in the past to study the stability of poloidal or toroidal fields (Tayler, 1973a, 1973b) and also of mixed combinations of both (Tayler, 1980).

In cylindrical geometry, it can be proven that the plasma is stable for all azimuthal and vertical wave numbers (m and k), if it is stable for $m = 0$ in the $k \to 0$ limit, and for $m = 1$ for all k (Goedbloed and Poedts, 2004). On the other hand, to show that a generic configuration with a combination of axial field and non-homogeneous azimuthal field is stable against the $m = 1$ mode (for all k) is not an easy task in general and one has to resort either to a variational approach or to a numerical investigation of the full eigenvalue problem in the complex plane. In this respect, the "normal mode" approach can be more useful in astrophysics, as it is often important to know the growth rate of the instability and the properties of the spectrum of the unstable modes (Bonanno and Urpin, 2008a, 2008b). In particular, generic combinations of axial and azimuthal fields are subject to a class of resonant MHD waves that can never be stabilized for any value of the ratio of poloidal and toroidal fields. The instability of these waves has a mixed character, being both current- and pressure-driven. In this case the most rapidly growing unstable modes are resonant, *i.e.* the wave vector $\mathbf{k} = (m/s)\mathbf{e}_\theta + k_z\mathbf{e}_z$ is perpendicular to the magnetic field, $\mathbf{B}\cdot\mathbf{k} = 0$ where k_z is the wavevector in the axial direction, m is the azimuthal wavenumber, and s is the cylindrical radius. The length scale of this instability depends on the ratio of poloidal and azimuthal

field components and it can be very short, while the width of the resonance turns out to be extremely narrow. For this reason its excitation in simulations can be problematic.

It is interesting to have a qualitative understanding of the MHD spectrum for a simple cylindrical plasma equilibrium configuration consisting of a mixed configuration of an azimuthal B_ϕ and a constant axial field B_z. As was shown in Bonanno and Urpin (2011), for a generic disturbance of the form $e^{(\sigma t - ik_z z - im\varphi)}$ an approximate expression for the dimensionless growth rate [$\Gamma = \sigma/\omega_{A\phi}$] being $\omega_{A\phi}$, the Alfvén frequency in the azimuthal direction, is given by

$$\Gamma^2 = \frac{2m^2(\alpha - 1)}{m^2 + (p^2 + m^2)\varepsilon^2}, \tag{1}$$

where p is the dimensionless radial wavenumber, $\varepsilon = B_z/B_\varphi$, $\alpha = \partial \ln B_\varphi / \partial \ln s$, and s is the cylindrical radius. Moreover, the dimensionless vertical wavenumber [$q = k_z s$] is close to the resonance condition [$f \equiv q\varepsilon + m \approx 0$], which implies that the total Alfvén frequency is zero [$\omega_A = \mathbf{B}\cdot\mathbf{k}/\sqrt{4\pi\rho} = 0$] at the resonance. The instability is never suppressed for any finite value of ε and the growth rate is a rapidly increasing function of m in particular $\Gamma^2 \approx (1+\varepsilon^2)^{-1}$ in the limit $m \gg p^2$. If $\alpha < 1$, it is possible to show that

$$\Gamma^2 \approx f^2 \frac{1+\alpha}{1-\alpha}, \tag{2}$$

which implies instability if $\alpha > -1$. The profile with $\alpha < -1$ is stable in this approximation. Note that modes with q satisfying the resonance condition $\omega_A = 0$ (or $f = 0$) are marginally stable because $\Gamma = 0$ for them, but $\Gamma^2 > 0$ in a neighborhood of the resonance. Therefore, the dependence of Γ on q should have a two-peak structure for any m. As in the case $\alpha > 1$, the instability occurs for any value of ε. If $\alpha = 1$, then we have

$$\Gamma^2 \approx \mu f \left[\frac{2m}{m^2+q^2} \pm \sqrt{\frac{4m^2}{(m^2+q^2)^2} + 4\mu} \right], \tag{3}$$

where μ is a positive number of the order unity. In this case, the dependence $\Gamma^2(q)$ also has a two-peak structure because $\Gamma = 0$ at the resonance but $\Gamma^2 > 0$ in its neighborhood. The instability is always present for any finite value of ε. This explicit solution shows that, if $\alpha > -1$, the instability always occurs for disturbances with q and m close to the condition of magnetic resonance, $\omega_A = 0$. The axial field cannot suppress the instability which occurs even if B_z is significantly greater than B_φ.

3. Kink Instabilities Below the Tachocline

Can the instabilities described in the previous section be operative in the overshoot layer of the Sun, or even below? To address this question we must consider the problem in spherical geometry, including the stabilizing effect of gravity and the destabilizing effect of thermal diffusivity. In fact, the stability of the spherical magnetic configurations has been studied in much less detail and even the overall stability properties of stellar radiation zones are rather unclear. Braithwaite and Nordlund (2006) studied the stability of a random initial field in the stellar radiative zone by direct numerical simulations, and it was found that the stable magnetic configurations generally have the form of tori with comparable poloidal and toroidal-field strengths. The possible relaxation mechanism was further discussed by Duez

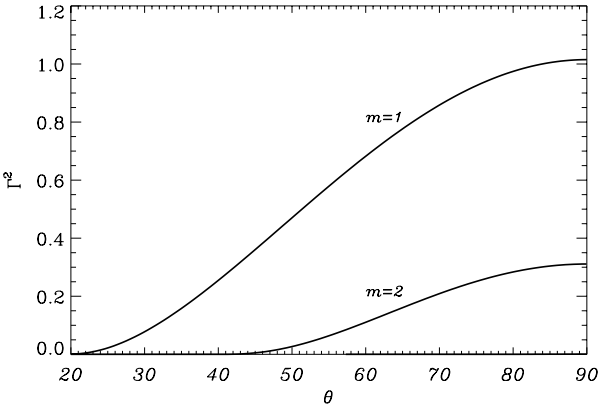


Figure 1 The dimensionless growth rate of the fundamental radial mode as a function of the polar angle, θ, for $\ell = 5$, $m = 1, 2$, and $d = 0.1$ in the case of a neutral stratification. Note that ℓ is the latitudinal "quantum" number of the perturbation. See Bonanno and Urpin (2012) for further details.

and Mathis (2010). The stability of azimuthal fields near the rotation axis has also been studied by Spruit (1999). The author used a heuristic approach to estimate the growth rate and criteria of instability. Unfortunately, many of these estimates and criteria are misleading because they do not apply in the main fraction of the volume of a radiation zone where the stability properties can be qualitatively different. The heuristic approach was criticized by Zahn, Brun, and Mathis (2007). The stability of the toroidal field in rotating stars has been considered by Kitchatinov (2008) and Kitchatinov and Rüdiger (2008) who argued that the magnetic instability is essentially three-dimensional and determined the threshold field strength at which the instability sets. Estimating this threshold in the solar radiation zone, the authors impose an upper limit on the magnetic field ≈ 600 G.

The problem has recently been investigated by Bonanno and Urpin (2012) where it was shown that the most unstable modes have low radial wavelengths at variance with the claim of Kitchatinov and Rüdiger (2008). Moreover if the thermal conductivity is considered, no threshold field is needed to trigger the instability. The most interesting result is illustrated by the angular dependence of the growth rate, as can be seen in Figure 1, where it appears that the instability is effective mostly near the Equator, and not along the axis, as originally supposed by Tayler (1973a). In general the location of the most unstable latitude depends also on the geometry of the basic state and cannot be determined only by local instability criteria.

Can this instability produce an α-effect? It is difficult to consistently compute the strength of the α-effect within the linear analysis alone, although a first estimation in this direction has been provided by Rüdiger, Kitchatinov, and Elstner (2012). However, recent investigations in cylindrical symmetry pointed out that it is possible to generate a non-zero helicity out of a kink instability due to non-linear effects (Del Sordo *et al.*, 2011; Bonanno *et al.*, 2012).

4. Flux-Transport Dynamo and Tachocline α-Effect

The previous discussion put forward that a likely location of the α-effect generated by current-driven instabilities is precisely in the overshoot layer with a strong concentration of the turbulence at low latitudes where the instability is more effective. Although only 3D numerical simulations can in principle determine the full structure of the α-tensor, as a first step in building a realistic model we can consider a latitudinal dependence of

the type $\cos\theta \sin^2\theta$ to model the suppression of the α-effect near the poles. The other essential ingredient is the inclusion of the meridional circulation. In fact, in the presence of a low eddy diffusivity (η_t) (as there is likely to be in the overshoot region), the magnetic Reynolds number becomes very large, and the dynamics of the mean-field flow becomes an essential ingredient of the dynamo process. In this regime the advection produced by the meridional circulation dominates the diffusion of the magnetic field which is then "transported" by the meridional circulation (Dikpati and Charbonneau, 1999; Dikpati and Gilman, 2001; Küker, Rüdiger, and Schultz, 2001; Bonanno *et al.*, 2002; Chatterjee, Nandy, and Choudhuri, 2004; Bonanno, Elstner, and Belvedere, 2006; Guerrero and de Gouveia Dal Pino, 2008).

The magnetic induction equation reads

$$\frac{\partial \mathbf{B}}{\partial t} = \boldsymbol{\nabla}(\mathbf{U} \times \mathbf{B} + \alpha \mathbf{B}) - \boldsymbol{\nabla} \times (\eta_T \boldsymbol{\nabla} \times \mathbf{B}), \tag{4}$$

where η_T is the turbulent diffusivity. Axisymmetry implies that relative to spherical coordinates the magnetic field [$\mathbf{B}$] and the mean flow field ($\mathbf{U}$), respectively, read

$$\mathbf{B} = B(r,\theta,t)\mathbf{e}_\phi + \boldsymbol{\nabla} \times \big[A(r,\theta,t)\mathbf{e}_\phi\big],$$

$$\mathbf{U} = \mathbf{u}(r,\theta) + r\sin\theta\,\Omega(r,\theta)\mathbf{e}_\phi$$

being $A(r,\theta,t)$ the vector potential. The meridional circulation [$\mathbf{u}(r,\theta)$] and differential rotation [$\Omega(r,\theta)$] are the poloidal and toroidal components of the global velocity flow field [$\mathbf{U}$]. In particular the poloidal and toroidal components of Equation (4), respectively, determine

$$\frac{\partial A}{\partial t} + \frac{1}{s}(\mathbf{u}\cdot\boldsymbol{\nabla})(sA) = \alpha B + \frac{\eta_T}{r}\frac{\partial^2(rA)}{\partial r^2} + \frac{\eta_T}{r^2}\frac{\partial}{\partial\theta}\left(\frac{1}{s}\frac{\partial(sA)}{\partial\theta}\right), \tag{5a}$$

$$\begin{aligned}\frac{\partial B}{\partial t} &+ s\rho(\mathbf{u}\cdot\boldsymbol{\nabla})\left(\frac{B}{s\rho}\right)\\ &= \frac{\partial\Omega}{\partial r}\frac{\partial(A\sin\theta)}{\partial\theta} - \frac{1}{r}\frac{\partial\Omega}{\partial\theta}\frac{\partial(sA)}{\partial r} + \frac{1}{r}\frac{\partial}{\partial r}\left(\eta_T\frac{\partial(rB)}{\partial r}\right)\\ &\quad + \frac{\eta_T}{r^2}\frac{\partial}{\partial\theta}\left(\frac{1}{s}\frac{\partial(sB)}{\partial\theta}\right) - \frac{1}{r}\frac{\partial}{\partial r}\left(\alpha\frac{\partial(rA)}{\partial r}\right) - \frac{\partial}{\partial\theta}\left(\frac{\alpha}{\sin\theta}\frac{\partial(A\sin\theta)}{\partial\theta}\right),\end{aligned} \tag{5b}$$

where $s = r\sin\theta$. The α-effect and the turbulent diffusivity are parametrized by means of

$$\begin{aligned}\alpha &= \frac{1}{4}\alpha_0\cos\theta\sin^2\theta\left[1 + \operatorname{erf}\left(\frac{x-a_1}{d}\right)\right]\left[1 - \operatorname{erf}\left(\frac{x-a_2}{d}\right)\right],\\ \eta &= \eta_c + \frac{1}{2}(\eta_t - \eta_c)\left[1 + \operatorname{erf}\left(\frac{r-r_\eta}{d_\eta}\right)\right],\end{aligned} \tag{6}$$

where α_0 is the amplitude of the α-effect, $x = r/R_\odot$ is the fractional radius, a_1, a_2 and d define the location and the thickness of the turbulent layer, η_t is the eddy diffusivity, η_c the magnetic diffusivity beneath the convection zone and d_η represents the width of this transition. In this investigation, the values $a_1 = 0.67$, $a_2 = 0.72$, and $d_\eta = 0.025$ have been used.

The components of the meridional circulation can be represented with the help of a stream function $\Psi(r,\theta) = -\sin^2\theta\cos\theta\psi(r)$ so that

$$u_r = \frac{1}{r^2\rho\sin\theta}\frac{\partial\Psi}{\partial\theta}, \qquad u_\theta = -\frac{1}{r\rho\sin\theta}\frac{\partial\Psi}{\partial r} \tag{7}$$

with the consequence that the condition $\nabla\cdot(\rho\mathbf{u}) = 0$ is automatically fulfilled. A strategy to constrain several properties of the meridional circulation is to assume the differential-rotation profile $\Omega(r,\theta)$ as a given ingredient, and deduce an approximation for the function ψ from the angular-momentum conservation along the azimuthal direction. An approximate expression for ψ is thus

$$\psi \approx \frac{5\rho r}{2\Omega_{\rm eq}}\int_0^\pi \langle u_r u_\theta\rangle\,{\rm d}\theta \tag{8}$$

where $\Omega_{\rm eq}$ is the equatorial angular velocity. In particular, for the standard, isotropic mixing-length theory, Equation (8) becomes (Durney, 2000)

$$\psi \approx -\frac{5\rho r}{2\Omega_{\rm eq}}\langle u_r^2\rangle. \tag{9}$$

In principle it would be possible to explicitly compute ψ and ψ' using the relation (9) knowing the convective velocities of the underlying stellar model. In practice this would be problematic, because the convective fluxes and their radial derivatives computed from standard mixing-length theory are discontinuous at the base of the convective zone. In a more realistic situation the presence of an overshoot layer implies that $\langle u_r^2\rangle \to 0$ smoothly so that u_θ is continuous at the inner boundary. Nevertheless, one can use the representation (9) to determine the stagnation point where $\psi' = 0$, which turns out to be around $x = 0.8$ in terms of the fractional radius. An explicit form of the function ψ which incorporates the following features reads:

$$\psi = C\left[1 - \exp\left(-\frac{(x-x_b)^2}{\sigma^2}\right)\right](x-1)x^2, \tag{10}$$

where C is a normalization factor, $x_b = 0.65$ defines the penetration of the flow, $\sigma = 0.08$ measures how rapidly $\langle u_r^2\rangle$ decays to zero in the overshoot layer and the location of the stagnation point. The density profile is taken to be $\rho = \rho_0(\frac{1}{x} - x_0)^m$ in which m is an index representing the stratification of the underlying solar model. Its value in the region of interest is approximately 1.5, and $x_0 = 0.85$, so that with these value the strength of the meridional circulation at low latitudes is of the same order as the surface flow, as discussed by Bonanno (2011); see also Pipin and Kosovichev (2011b) for a similar investigation. The radial profile of the α-effect, turbulent diffusivity, stream function, and meridional circulation used in the calculation are depicted in Figure 2.

The helioseismic profile for the differential rotation is taken so that

$$\Omega(r,\theta) = \Omega_{\rm c} + \delta\left(\frac{r-r_{\rm c}}{d_{\rm c}}\right)\big(\Omega_{\rm s}(\theta) - \Omega_{\rm c}\big), \tag{11}$$

where $\Omega_{\rm c}/2\pi = 432.8$ nHz is the uniform angular velocity of the radiative core, $\Omega_{\rm s}(\theta) = \Omega_{\rm eq} + a_2\cos^2\theta + a_4\cos^4\theta$ is the latitudinal differential rotation at the surface and $\delta(x) \equiv (1 + {\rm erf}(x))/2$. In particular, $\Omega_{\rm eq}/2\pi = 460.7$ nHz is the angular velocity at the Equator, $a_2/2\pi = -62.9$ nHz, and $a_4/2\pi = -67.13$ nHz. In this calculation the angular velocity

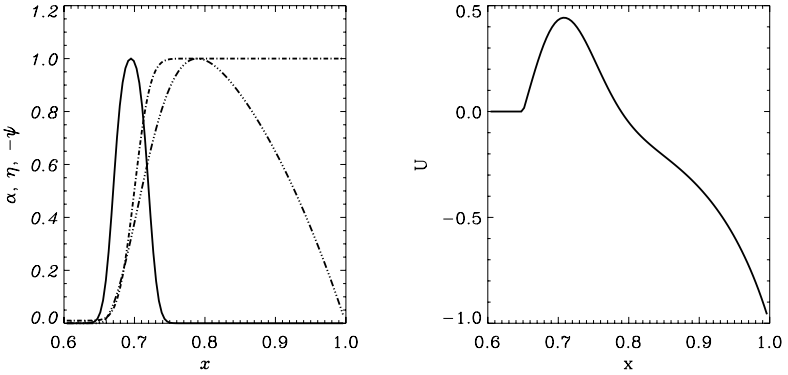


Figure 2 The α-effect (solid line), turbulent diffusivity (dot-dashed line), and (minus) the function $\psi(r)$ (dot-dot-dot-dashed line), are depicted in the left panel. The meridional circulation in units of the maximum surface value at a latitude of 45° is instead depicted in the right panel.

is normalized in terms of equatorial differential rotation $\Omega_{\rm eq}$, $r_{\rm c} = 0.71$, and $d_{\rm c} = 0.025$. As usual, the dynamo equations can be made dimensionless by introducing the dynamo numbers $C_\Omega = R_\odot^2 \Omega_{\rm eq}/\eta_{\rm T}$, $C_\alpha = R_\odot \alpha_0/\eta_{\rm T}$, $C_\omega = R_\odot \omega/\eta_{\rm T}$, $C_u = R_\odot U/\eta_{\rm T}$, where ω is the frequency of the dynamo wave, and $U = u_\theta(r = R_\odot, \theta = 45°)$. This linear dynamo problem is solved with a finite-difference scheme for the radial dependence and a polynomial expansion for the angular dependence by imposing a potential field as a boundary condition for the surface, and a perfect conductor for the inner boundary. Further details of the numerical approach can be found in Bonanno *et al.* (2002), Bonanno, Elstner, and Belvedere (2006), and also in Jouve *et al.* (2008).

A reference solution can then be obtained in the region of high C_u, and its main properties are the following: $C_\alpha = 2$, $C_\Omega = 4.5 \times 10^5$, and $C_u = 400$, which implies a period of about 24 years, a turbulent diffusivity $\eta_{\rm t} = 3.1 \times 10^{11}$ cm^2 s^{-1}, and a (poleward) surface flow of about 18 m s^{-1}. The basic properties of this solution are depicted in Figure 3 where the butterfly diagram for the toroidal field at the bottom of the convection zone and of the radial field at the surface are shown. It is interesting to notice that at low latitude the phase relation is consistent with the observations, namely $B_r B_\phi < 0$. The other property of this solution is that it shows a strong preference for dipolar modes. In particular the critical solution for the symmetric mode has $C_\alpha = 3.70$, significantly greater than the antisymmetric mode. On the contrary, if we had considered the case of an α-effect uniformly distributed throughout the whole of the convection zone, the result would have been $C_\alpha = 2.1$ for both symmetric and antisymmetric modes with a difference of about 1 % most probably due to numerics. The conclusion is that the parity of the solution is more likely to be a dipole if the α-effect is located at the bottom of the convection zone. In our simulations the precise value of the ratio $\eta_{\rm c}/\eta_{\rm t}$ was not crucial, as we obtained basically the dynamo solution for $\eta_{\rm c}/\eta_{\rm t} = 0.005$ and $\eta_{\rm c}/\eta_{\rm t} = 0.05$ although in all of the simulations presented in this work the value $\eta_{\rm c}/\eta_{\rm t} = 0.01$ has been chosen. For instance the period obtained for $\eta_{\rm c}/\eta_{\rm t} = 0.005$ was 24.9 years, and for $\eta_{\rm c}/\eta_{\rm t} = 0.05$ was 24.5 years. This is not surprising as the penetration of the meridional flow is very weak as showed in the right panel of Figure 2 and therefore the precise value of $\eta_{\rm c}$ cannot significantly change the type of dynamo action. In addition, changing the width of the α-effect and of the turbulent diffusivity also did not lead to significant changes in the solution, although we expect that the spatial extension of the turbulent layer should be of the order of the tachocline width $d_\eta \approx 0.05$ solar radii.

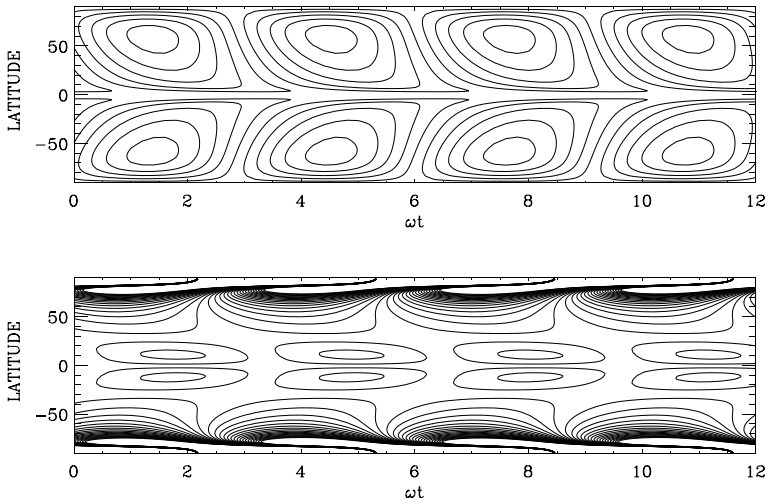


Figure 3 Butterfly diagram for a solution with meridional circulation included. The isocontours for the toroidal field (upper panel) at the base of the convection zone and the radial field on the top boundary are shown. Full and dashed lines show the positive and negative levels respectively. The solution is obtained with $C_\alpha = 2$, $C_\Omega = 4.5 \times 10^5$, and $C_u = 400$, which implies a period of about 24 years a turbulent diffusivity $\eta_t = 3.1 \times 10^{11}$ cm^2 s^{-1} and a (poleward) surface flow of about 18 m s^{-1}. It is interesting to notice that at low latitude the phase relation is consistent with the observations, namely $B_r B_\phi < 0$. This type of solution also shows a strong preference for dipolar modes. In particular the critical solution for the symmetric mode has $C_\alpha = 3.70$, significantly greater than the antisymmetric mode. An essential ingredient in order to get this property correct is the presence of a strong meridional circulation at the bottom of the convective zone. In particular in this case the bottom flow is about 8.5 m s^{-1}.

However, the serious, unsatisfactory aspect of an advection-dominated dynamo is the fact that the strength of the return flow is largely unknown, and the eddy diffusivity is about one order of magnitude greater than would be expected on the basis of the standard mixing-length theory.

5. Solar Dynamo from Subsurface Shear Instabilities

In recent years the possibility that the dynamo operates in the subphotospheric layers of the Sun mostly driven by the negative gradient of the angular velocity near the surface has been proposed (Brandenburg, 2005). It is interesting to see if a mean-field dynamo model can describe this case with the help of a more refined differential-rotation profile including the negative shear in the subsurface layers. In order to discuss this issue, a slightly modified version of the analytical approximation presented in Dikpati *et al.* (2002) has been used. The radial profile of the α-effect and turbulent diffusivity are depicted in Figure 4. In particular the rotation rate is taken constant in the radiative interior [Ω_c], and the tachocline is located at the same radius as in the model described in the previous session. It is assumed that it has constant width $\approx 0.05 R_\odot$, and at the top of the tachocline its rotation rate is given by

$$\Omega(r_{cz}, \theta) = \Omega_{cz} + a_2 \cos^2\theta + a_4 \cos^4\theta \tag{12}$$

where $\Omega_{cz} = -a_2/5 - 3a_4/35$, $a_2 = -61$ nHz, and $a_4 = -73.5$ nHz. It is thus assumed that there is a known negative gradient below the surface down to a radius $r_s = 0.95 R_\odot$ and the

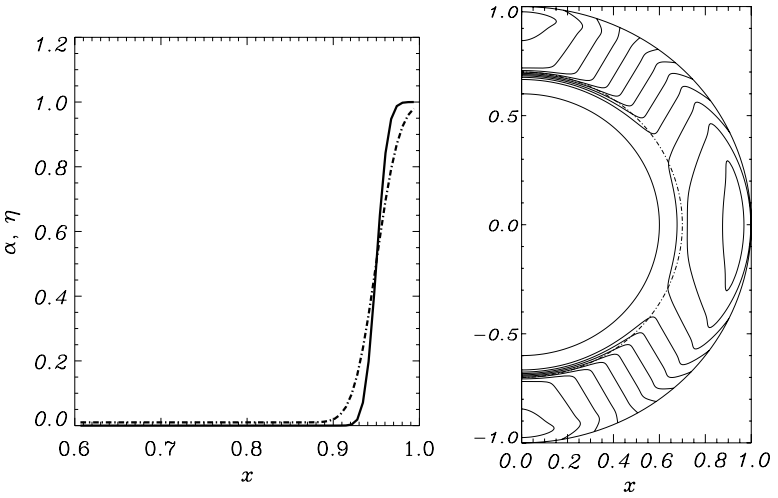


Figure 4 Left panel: the α-effect profile (solid line) and turbulent diffusivity (dot-dashed line) used for an $\alpha\Omega$ model with a dynamo action produced from subsurface shear instabilities. Right panel: the isocontour lines of the differential rotation $[\Omega(r,\theta)]$ in units of the equatorial rotation. Note the negative shear near the surface around $r/R_\odot \approx 0.95$.

latitudinal dependence of this shear layer is modeled by $P(\theta) = (p_0 + p_4 \cos^4\theta)/R_\odot$, so that

$$\Omega(r,\theta) = \Omega_c + \delta\left(\frac{r-r_c}{d_c}\right) Q(\theta)(r-r_{cz}) + \delta\left(\frac{r-r_s}{d_s}\right)\left[\Omega(r_{cz},\theta) - \Omega_c\right] + \delta\left(\frac{r-r_s}{d_s}\right)\left[\Omega_{eq} - \Omega_{cz} - P(\theta)(r-R_\odot) - Q(\theta)(r-r_{cz})\right] \tag{13}$$

and $Q(\theta) = (\Omega_{eq} - \Omega_{cz} + P(\theta)(R_\odot - r_s))/(r - r_{cz})$. For actual calculations, the values $p_0 = 437$ nHz and $p_4 = -722$ nHz have been chosen. In order to confine the magnetic field in the subsurface layers, it is assumed that η_t is also maximum near the surface, and it sharply decreases by about two orders of magnitude just below the supergranulation layer where the α-effect is located. The basic features of the (critical) dynamo solution in this case is depicted in Figure 5 for a solution with $C_\alpha = 18$, $C_\Omega = 1 \times 10^3$ which implies a period of about four years and a turbulent diffusivity $\eta_t = 1.4 \times 10^{13}$ cm^2 s^{-1}. The period is clearly too small, but this is not surprising as the dynamo wave is basically entirely confined in the surface layers as can be seen in Figure 6. The only possibility to match the solar period is to further increase the turbulent diffusivity in the supergranulation layer, but this would imply that the spatial extension of the dynamo wave propagates deeply within the convective zone, thus preventing the negative radial shear to produce the correct butterfly diagram. Moreover, the parity of the solution is clearly symmetric because $C_\alpha = 15$ for quadrupolar modes.

6. Conclusions

Where is the α-effect located in the Sun? Despite the difficulties present in models of flux-dominated dynamo (too low eddy diffusivity, unknown strength, and location of the return

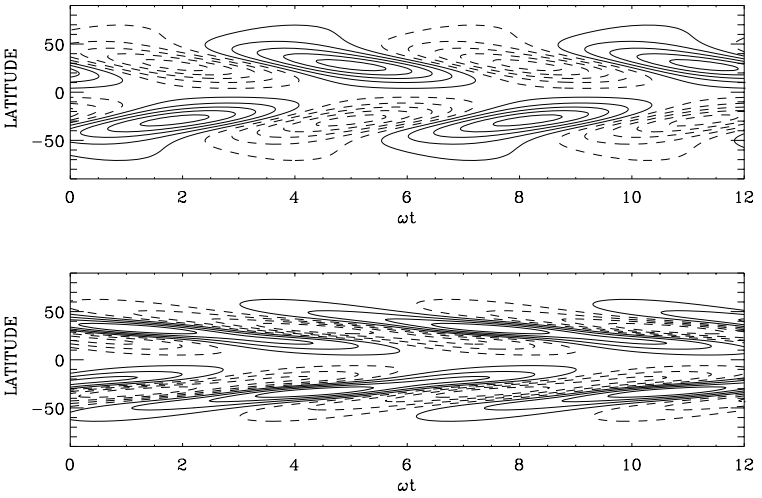
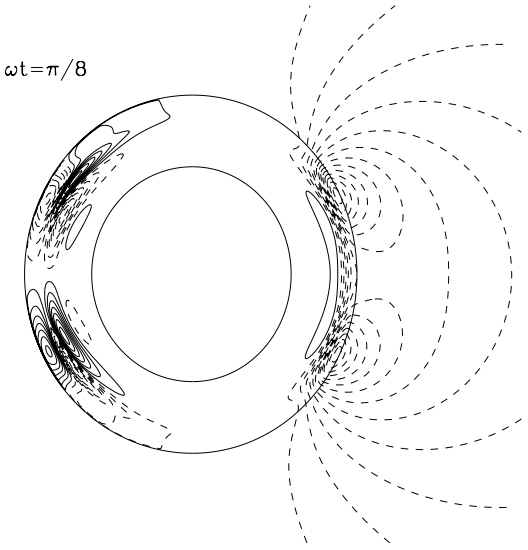


Figure 5 Butterfly diagram for a model with strong radial subsurface shear. The isocontours for the toroidal field (upper panel) at $r = 0.95R$ and the radial field on the top boundary are shown. Full and dashed lines show the positive and negative levels respectively. The solution is obtained with $C_\alpha = 18$ and $C_\Omega = 1 \times 10^3$, which implies a period of about four years and a turbulent diffusivity $\eta_t = 1.4 \times 10^{13}$ cm^2 s^{-1}.

Figure 6 Toroidal and poloidal field configuration for a typical dynamo solution with a strong subsurface shear, and $C_\alpha = 18$, $C_\Omega = 1 \times 10^3$. In the left part are the isocontour lines of the toroidal field, with solid line for negative B_ϕ and dashed line for positive value of the field. The right part represents of the streamlines of the poloidal field given by contours of $Ar\sin\theta$. Solid line are for negative values of A. Note that the dynamo action is confined around the region of strongest radial shear, and, for this reason, the period tends to be rather short in general (less than ten years in these types of model).



flow), mean-field models with an α-effect located at the bottom of the convective zone are successful in reproducing several aspects of the solar activity cycle. In this model the origin of the α-effect is due to the turbulent magnetic helicity and current helicity generated by the quasi-interchange instabilities below the tachocline and does not follow from mixing-length theory. The most striking result comes from the parity of the solution: only with an α-effect located at the bottom of the convective zone the dipolar modes are most easily excited, at least at the linear level. This result also applies for models with the α-effect generated by subsurface shear; the solutions are mostly symmetric, rather than antisymmetric. It would be

nice to see if a more realistic surface boundary condition can solve this problem as proposed by Pipin and Kosovichev (2011a).

Although the models discussed in the previous session are kinematic, it is difficult to believe that the discussion of the previous session will drastically change if a non-linearity via the α-quenching is included in the models. Jouve *et al.* (2008) present a code comparison between kinematic models and non-linear models that shows that the critical solutions are basically the same within the tolerance of different numerical schemes. On the other hand, the advantage of the kinematic approach is to provide a complete view of the spectrum of the dynamo waves that can be excited for a given set of dynamo numbers. This is clearly an important piece of information for the non-linear simulations, for instance for investigating degenerate configurations due to the degeneracy in the spectrum.

An important constraint on the viability of a solar dynamo with a tachocline α-effect is the fact that successful models must have a *positive* α-effect: no migration is present for a negative α-effect. It would be an important check for the theory to show that this is actually the case by means of 3D global numerical simulations of the kink and quasi-interchange instability in spherical symmetry. Recent investigations in this direction in cylindrical symmetry have questioned this possibility (Gellert, Rüdiger, and Hollerbach, 2011), although further numerical and analytical work is needed before a firm conclusion can be reached.

Acknowledgements I am grateful to Axel Brandenburg and Fabio Del Sordo for stimulating discussions. I would also like to thank the organizers of the SDO meeting (in particular Sasha Kosovichev) and ESA for financial support.

References

Bernstein, I.B., Frieman, E.A., Kruskal, M.D., Kulsrud, R.M.: 1958, *Proc. Roy. Soc. London A* **244**, 17.
Braithwaite, J., Nordlund, A.: 2006, *Astron. Astrophys.* **450**, 1077.
Bonanno, A.: 2011, arXiv:1112.4929.
Bonanno, A., Elstner, D., Belvedere, G.: 2006, *Astron. Nachr.* **327**, 680.
Bonanno, A., Urpin, V.: 2008a, *Astron. Astrophys.* **477**, 35.
Bonanno, A., Urpin, V.: 2008b, *Astron. Astrophys.* **488**, 1.
Bonanno, A., Urpin, V.: 2011, *Phys. Rev. E* **84**, 056310.
Bonanno, A., Urpin, V.: 2012, *Astrophys. J.* **747**, 137.
Bonanno, A., Elstner, D., Rüdiger, G., Belvedere, G.: 2002, *Astron. Astrophys.* **390**, 673.
Bonanno, A., Brandenburg, A., Del Sordo, F., Mitra, D.: 2012, *Phys. Rev. E* **86**, 016313. arXiv:1204.0081.
Brandenburg, A.: 2005, *Astrophys. J.* **625**, 539.
Chatterjee, P., Nandy, D., Choudhuri, A.R.: 2004, *Astron. Astrophys.* **427**, 1019.
Chatterjee, P., Mitra, D., Rheinhardt, M., Brandenburg, A.: 2011, *Astron. Astrophys.* **534**, 46.
Del Sordo, F., Bonanno, A., Brandenburg, A., Dhrubaditya, M.: 2011, In: Mandrini, C.H. (ed.) *Comparative Magnetic Minima: Characterizing Quiet Times in the Sun and Star*, *IAU Symp.* **286**, Cambridge University Press, Cambridge. arXiv:1111.1742.
Dikpati, M., Charbonneau, P.: 1999, *Astrophys. J.* **518**, 508.
Dikpati, M., Gilman, P.A.: 2001, *Astrophys. J.* **559**, 428.
Dikpati, M., Corbad, T., Thompson, M.J., Gilman, P.A.: 2002, *Astrophys. J.* **575**, L41.
Duez, V., Mathis, S.: 2010, *Astrophys. J.* **517**, 58.
Durney, B.R.: 2000, *Astrophys. J.* **528**, 486.
Gellert, M., Rüdiger, G., Hollerbach, R.: 2011, *Mon. Not. Roy. Astron. Soc.* **414**, 2696.
Ferriz-Mas, A., Schmitt, D., Schüssler, M.: 1994, *Astron. Astrophys.* **289**, 949.
Goedbloed, H., Poedts, S.: 2004, *Principles of Magnetohydrodynamics*, Cambridge University Press, Cambridge.
Gruzinov, A.V., Diamond, P.H.: 1994, *Phys. Rev. Lett.* **72**, 1651.
Guerrero, G., de Gouveia Dal Pino, E.M.: 2008, *Astron. Astrophys.* **485**, 267.
Jouve, L., *et al.*: 2008, *Astron. Astrophys.* **483**, 949960.
Küker, M., Rüdiger, G., Schultz, M.: 2001, *Astron. Astrophys.* **374**, 301.
Kitchatinov, L.: 2008, *Astron. Rep.* **52**, 247.

Kitchatinov, L., Rüdiger, G.: 2008, *Astron. Astrophys.* **478**, 1.
Kosovichev, A.G.: 1996, *Astrophys. J.* **269**, 61.
Parker, E.N.: 1955, *Astrophys. J.* **122**, 293.
Parker, E.N.: 1993, *Astrophys. J.* **408**, 707.
Pipin, V.V., Kosovichev, A.G.: 2011a, *Astrophys. J.* **727**, L45.
Pipin, V.V., Kosovichev, A.G.: 2011b, *Astrophys. J.* **738**, 104.
Rüdiger, G., Kitchatinov, L., Elstner, D.: 2012, *Mon. Not. Roy. Astron. Soc.* **425**, 2267. arXiv:1107.2548.
Schou, J.: 1991, In: Gough, D., Toomre, J. (eds.) *Challenges to Theories of the Structure of Moderate-Mass Stars*, *Lecture Notes Phys.* **388**, Springer, Berlin, 81.
Silant'ev, N.A.: 2000, *Astron. Astrophys.* **364**, 449.
Spiegel, E.A., Zahn, J.-P.: 1992, *Astron. Astrophys.* **265**, 106.
Spruit, H.: 1999, *Astron. Astrophys.* **349**, 189.
Steenbeck, M., Krause, F., Rädler, K.H.: 1966, *Z. Naturforsch.* **21a**, 369.
Tayler, R.J.: 1973a, *Mon. Not. Roy. Astron. Soc.* **161**, 365.
Tayler, R.J.: 1973b, *Mon. Not. Roy. Astron. Soc.* **163**, 77.
Tayler, R.J.: 1980, *Mon. Not. Roy. Astron. Soc.* **191**, 151.
Zahn, J.-P., Brun, A.S., Mathis, S.: 2007, *Astron. Astrophys.* **474**, 145.

DOI 10.1007/978-1-4899-8005-2_13
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-013-0345-1

A Comparison of Solar Cycle Variations in the Equatorial Rotation Rates of the Sun's Subsurface, Surface, Corona, and Sunspot Groups

J. Javaraiah

Received: 13 February 2012 / Accepted: 6 June 2013 / Published online: 18 July 2013

Abstract Using the *Solar Optical Observing Network* (SOON) sunspot-group data for the period 1985 – 2010, the variations in the annual mean equatorial-rotation rates of the sunspot groups are determined and compared with the known variations in the solar equatorial-rotation rates determined from the following data: i) the plasma rotation rates at $0.94R_\odot, 0.95R_\odot, \ldots, 1.0R_\odot$ measured by the *Global Oscillation Network Group* (GONG) during the period 1995 – 2010, ii) the data on the soft-X-ray corona determined from *Yohkoh*/SXT full-disk images for the years 1992 – 2001, iii) the data on small bright coronal structures (SBCS) that were traced in *Solar and Heliospheric Observatory* (SOHO)/EIT images during the period 1998 – 2006, and iv) the Mount Wilson Doppler-velocity measurements during the period 1986 – 2007. A large portion (up to $\approx 30^\circ$ latitude) of the mean differential-rotation profile of the sunspot groups lies between those of the internal differential-rotation rates at $0.94R_\odot$ and $0.98R_\odot$. The variation in the yearly mean equatorial-rotation rate of the sunspot groups seems to be lagging behind that of the equatorial-rotation rate determined from the GONG measurements by one to two years. The amplitude of the GONG measurements is very small. The solar-cycle variation in the equatorial-rotation rate of the solar corona closely matches that determined from the sunspot-group data. The variation in the equatorial-rotation rate determined from the Mount Wilson Doppler-velocity data closely resembles the corresponding variation in the equatorial-rotation rate determined from the sunspot-group data that included the values of the abnormal angular motions ($> |3^\circ|\,\mathrm{day}^{-1}$) of the sunspot groups. Implications of these results are pointed out.

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

J. Javaraiah (✉)
Indian Institute of Astrophysics, Bengaluru 560 034, India
e-mail: jj@iiap.res.in

1. Introduction

Studies on the solar-cycle variations in the solar differential rotation and meridional flow are important for understanding the physical system that generates solar activity and the solar cycle (Babcock, 1961; Ulrich and Boyden, 2005; Dikpati and Gilman, 2006; Karak, 2010). The solar differential rotation is well studied; in addition to using Doppler-velocity measurements, the data on different solar magnetic tracers are employed (mainly the sunspots and the sunspot groups). The results derived from the sunspot and the sunspot-group data show variations in both the equatorial-rotation rate and the latitude gradient of the rotation on several timescales, including timescales close to the 11-year solar cycle (Javaraiah and Gokhale, 1995; Javaraiah and Komm, 1999; Javaraiah, Bertello, and Ulrich, 2005a; Javaraiah, Bertello, and Ulrich, 2005b; Javaraiah and Ulrich, 2006; Brajša, Ruždjak, and Wöhl, 2006). Recently, Javaraiah (2011) detected variations on timescales of a few days, including 9 and 30 – 40 day quasi-periodicities, in the coefficients of differential rotation determined from the Mount Wilson Doppler-velocity data during Solar Cycle 22. The $\approx$ 11-year period torsional oscillation detected by Howard and LaBonte (1980) in the Mount Wilson Doppler-velocity measurements is confirmed by helioseismic studies (*e.g.* Howe *et al.*, 2000; Antia, Basu, and Chitre, 2008)

Javaraiah and Komm (1999) studied the variations in the coefficients of the differential rotation determined from the sunspot-group data (1879 – 1976) and the Mount Wilson Doppler-velocity data (1962 – 1994). They found considerable differences between the periodicities (> two years) in the rotational coefficients derived from the Doppler-velocity and the sunspot-group data. Javaraiah *et al.* (2009) analyzed the data on the solar surface equatorial-rotation rate derived from the more accurate Mount Wilson Doppler-velocity data during the period 1986 – 2007 and confirmed the short-term periodicities found by Javaraiah and Komm (1999) in the data before the year 1996, but found no statistically significant variation after 1996. That is, there is a very large difference between the temporal variations of the equatorial-rotation rate during Solar Cycle 22 (*i.e.* before 1995) and Cycle 23. Hence, it has been suspected that the frequent changes in the instrumentation of the Mount Wilson spectrograph may have made the data before 1995 erroneous and responsible for the variations in the equatorial-rotation rate derived from these data. To verify this, here we compare the variations in the equatorial-rotation rates derived from different rotational data and with different techniques.

2. Data and Analysis

2.1. Methodology

The solar differential rotation can be determined from the full-disk Doppler-velocity data using the traditional polynomial expansion

$$\omega(\phi) = A + B\sin^2\phi + C\sin^4\phi, \tag{1}$$

and from sunspot data by using the first two terms of the expansion, *i.e.*

$$\omega(\phi) = A + B\sin^2\phi, \tag{2}$$

where $\omega(\phi)$ is the solar sidereal angular velocity at heliographic latitude ϕ, the coefficient A represents the equatorial-rotation rate, and B and C measure the latitudinal gradient

in the rotation rate, with B mainly representing low latitudes and C mainly higher latitudes.

The sidereal rotation rate [ω, in degrees day^{-1}] of a sunspot or sunspot group is computed as (Carrington, 1863; Godoli and Mazzucconi, 1979; Balthasar, Vázquez, and Wöhl, 1986)

$$\omega = \frac{\Delta L}{\Delta t} + 14.18, \tag{3}$$

or as (*e.g.* Howard, Gilman, and Gilman, 1984; Kambry and Nishikawa, 1990; Gupta, Sivaraman, and Howard, 1999; Brajša, Ruždjak, and Wöhl, 2006)

$$\omega = \frac{\Delta D_{\mathrm{CM}}}{\Delta t} + 0.9856, \tag{4}$$

where ΔL, ΔD_{CM}, and Δt are the differences between the heliographic longitudes [L], the central meridian distances [D_{CM}], and the observation times [t] of two consecutive days of observations of the sunspot or sunspot group, respectively, and the values 14.18° day^{-1} and 0.9856 are the Carrington rigid-body rotation rate and the correction factor corresponding to the Earth's rotation, respectively. The quantity $\frac{\Delta D_{\mathrm{CM}}}{\Delta t}$ represents the synodic rotation rate. Within the uncertainties, both these methods yield the same result (the latter gives a relatively low value for the rotation rate; see Kambry and Nishikawa, 1990). Javaraiah and co-authors have used Equation (3) in all of their earlier articles (*e.g.* Javaraiah and Gokhale, 1995, 1997a, 1997b; Javaraiah and Komm, 1999; Javaraiah, 2003a, 2005; Javaraiah, Bertello, and Ulrich, 2005a, 2005b; Javaraiah and Ulrich, 2006). (Details on the method of conversion of the synodic to sidereal rotation rate can be found in some of the above cited articles, *e.g.* Brajša, Ruždjak, and Wöhl, 2006 and references therein.)

The solar differential rotation can also be studied by binning the data into latitude intervals of reasonably small size, *i.e.* without using the above equations (see the references above). This subject is reviewed by a number of authors (*e.g.* Javaraiah and Gokhale, 2002 and references therein).

2.2. Sunspot Group Data

Here the Greenwich sunspot-group data during the period of May 1874 – December 1976 and the SOON sunspot-group data during 1977 – 2010 were used. These data were taken from the website solarcience.msfc.nasa.gov/greenwich.shtml and consist of t, ϕ, L, D_{CM}, *etc.* for each sunspot group on each day of its observation. By using Equation (3), the ω values were determined from the data corresponding to each pair of the consecutive days' observations of the sunspot groups during the period 1986 – 2009. Each year's data were fitted to Equation (2). The latitudinal dependencies on the mean rotation rate of the sunspot groups over the whole period 1977 – 2011 (whole SOON data set) were determined by averaging the daily values of the rotational velocities in two-degree (and five-degree) latitude intervals and also by fitting the whole-period daily data to Equation (2). The data in both the northern and southern hemispheres were combined. These calculations were also made for the Greenwich data during the period May 1874 – December 1976.

Since the rotation rates of tracers depend on the lifetimes, sizes, and age of the tracers, the latitudinal dependence in the initial rotation rates (the first two days' heliographic positions of the sunspot groups were used) of the sunspot groups was also studied by classifying the sunspot groups on the basis of their lifetimes: one to three days, four to five days, six to eight days, and longer than eight days. For this, the combined Greenwich and SOON data were used.

The sunspot-group data that correspond to $|D_{CM}| > 75^\circ$ on any day of the sunspot group lifetime were excluded. To study the initial rotation rates of the sunspot groups, the data during the entire lifetime of a sunspot group were eliminated on any one day for which $|D_{CM}| > 75^\circ$ to avoid ambiguity in the identification of the first two days in the lifetimes of the sunspot groups. Since the cutoff of D_{CM} was applied to both the eastern and western sides, the sunspot groups that emerged at the visible side of the Sun and after a certain number of days moved to the other side were not taken into account in the data sample of any of the aforementioned classes (Note: each disk passage of a recurring sunspot group was treated as an independent sunspot group). As in our previous articles, we excluded the data correspond to $> 2^\circ$ day^{-1} latitudinal motions, and $> 3^\circ$ day^{-1} longitudinal motions (*i.e. abnormal ω values*). Ward (1965, 1966) found that this precaution substantially reduces the uncertainties in the results (see Javaraiah and Gokhale, 1995). However, the abnormal values of ω may have some physical significance. For example, the abnormal rotation rates of sunspot groups may play some role in the productions of solar flares (Hiremath and Suryanarayana, 2003; Suryanarayana, 2010). Hence, most of the above calculations were made for both the cases, *i.e.* the sets of the sunspot-group data with and without the abnormal ω values were used.

2.3. Mount Wilson Doppler-Velocity Data

The daily values of the equatorial-rotation rate [A] derived from the Mount Wilson Doppler measurements during the period 3 December 1985 to 5 March 2007 are available (Javaraiah *et al.*, 2009). This period covers Solar Cycle 22 and most of Cycle 23. The data have been corrected for scattered-light (for details see Ulrich, 2001). Javaraiah *et al.* (2009) used these data after removing the very large spikes (*i.e.* $> 2\sigma$, where σ is the standard deviation of the original time series) and studied short-term variations in the corrected data. The time series has data gaps that vary in size, with a maximum gap of 49 days during Carrington-Rotation numbers 1560 – 1608. Therefore, the daily data were binned to one-year consecutive intervals, and the annual average values of the equatorial-rotation rate (average of the daily values of A over each year) were determined. In Figure 1 of Javaraiah *et al.* (2009) both the original and the corrected time series of the equatorial-rotation rate were shown. In the present analysis the corrected time series was used.

2.4. Data on Subsurface and Coronal Rotation Rates

The Sun's internal-rotation rates determined from the GONG data for 147 intervals of three GONG-months (a GONG-month is 36 days), which began on 7 May 1995 and ended on 31 October 2009, for $0.005R_\odot, 0.01R_\odot, 0.015R_\odot, \ldots, 0.995R_\odot, 1.0R_\odot$, and for latitudes $0, 2, 4, \ldots, 88$ degrees are available (Antia and Basu, 2010). Figure 1 shows variations in the equatorial-rotation rate [Ω_0, angular velocity at latitude $\phi = 0$] at $0.94R_\odot, 0.95R_\odot, \ldots, 1.0R_\odot$. As can be seen in this figure, the temporal patterns of the equatorial-rotation rates at the different depths are mostly similar. However, there is a suggestion that some features are smoothed when going from the interior toward the surface and *vice versa*. Specifically, the maxima in 2002 and 2008 are smoothed in deeper layers while the minimum in 2005 is smoothed from the interior toward the surface.

The annual mean values of the equatorial-rotation rate of the soft X-ray corona determined from the *Yohkoh*/SXT solar full-disk images for the period 1992 – 2001 were taken from Table 2 of Chandra, Vats, and Iyer (2010), and those determined from the rotation rates of small bright coronal structures (SBCS) that were traced in SOHO/EIT images during the period 1998 – 2006 were taken from Figure 1 of Jurdana-Šepić *et al.* (2011).

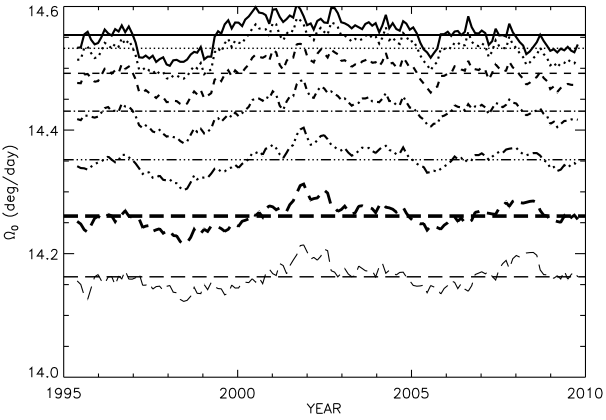


Figure 1 The solid, dotted, short-dashed, one dotted-dashed, three dotted-dashed, thick long-dashed, and thin long-dashed curves represent the variations in the equatorial-rotation rate [Ω_0] at $0.94R_\odot, 0.95R_\odot, \ldots, 1.0R_\odot$ (in the order of decreasing mean values of Ω_0), respectively, determined from the GONG data for 147 intervals of three GONG-months (for details about these data see Antia and Basu, 2010 and references therein).

3. Results

3.1. Depth Dependence in the Differential Rotation of Sunspot Groups

Figure 2 shows the latitudinal dependence on the mean rotation rate of sunspot groups determined from the SOON sunspot-group data during the period 1977 – 2011 separately from the data sets with and without the abnormal ω values. In the same figure the latitudinal dependencies on the mean (over the whole period 1995 – 2009) rotation rates of the plasma at different depths of the solar convection zone deduced from the GONG data are shown as well. As can be seen in this figure, a large portion (up to $\approx 30^\circ$ latitude) of the rotational profile that is obtained from the sunspot-group data that do not include the abnormal ω values lies between the corresponding profiles at $0.94R_\odot - 0.98R_\odot$ determined from the GONG data. Note that the value of the correlation coefficient shown in these figures is almost the same for any profile of the internal rotation shown in these figures. As can be seen in Figure 3, a similar conclusion can be drawn from the analysis of the large set of Greenwich data during May 1874 – December 1976 (however, the helioseismic measurements did not exist before 1995). The latitude dependence in the rotation determined from the sunspot-group data that included the abnormal ω values is obviously highly unrealistic. Since the equatorial-rotation rate determined from these data seems to be somewhat closer to the Ω_0 at $1.0R_\odot$, the comparison between its variation and that of the surface is studied here.

There are more abnormal values in the SOON (5.8 %) than in Greenwich data (2.9 %). In the Greenwich data the observation time contains the date with the fraction of a day. In the SOON data the fraction is rounded to 0.5 day. This might to some extent increase the number of abnormal ω values in the SOON data. Figure 4 shows the latitudinal dependence of the rotation determined by binning the sunspot-group data that included the abnormal ω values into five-degree latitude intervals. Obviously, the results show that the errors are much smaller than those shown in Figures 2(b) and 3(b). However, the latitudinal patterns of the data in five-degree latitude intervals (represented by the open circles) are mainly the same as the corresponding patterns of the data in the two-degree latitude intervals (*cf.* Figures 2(b) and 3(b)). The Greenwich data that included the abnormal ω values show rigid-body rotation (this is clearer in Figure 4(b) than in Figure 3(b)).

On average, the rotation rate determined from the magnetic tracers is higher than that derived from the Doppler-velocity measurements. This difference can be interpreted as the Doppler-velocity measurements representing the surface gas motion and the tracer motions

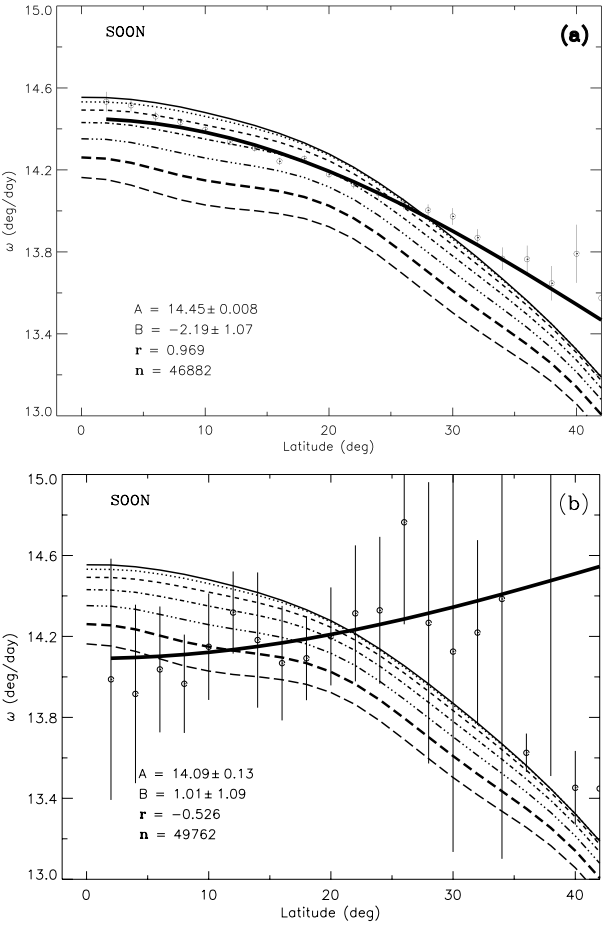


Figure 2 The latitudinal dependence of the mean rotation rates of sunspot groups determined by averaging the daily values of ω obtained from the SOON sunspot-group data during 1977 – 2011 over two-degree latitude intervals: $1° - 3°, 3° - 5°, 5° - 7°, \ldots, 40° - 42°$ (plotted at $2°, 4°, 6°, \ldots, 42°$). The error bars represent the standard errors. Panels (a) and (b) represent the data with and without the abnormal values of ω, respectively. The thick solid curve represents the corresponding mean profile deduced from the values (also shown) of the coefficients A and B of Equation (2) obtained from the total number of daily data [n]. The different curves represent the latitudinal dependencies on the mean (over the whole period 1995 – 2009) plasma rotation rates deduced from the GONG data at the same depths as in Figure 1. r represents the correlation coefficient between the sunspot group and the GONG data around $0.96R_{\odot}$ in $2°, 4°, 6°, \ldots, 36°$ latitudes.

representing the motions of the deeper layers (Foukal, 1972; Meunier, 2005), where the magnetic structures of the tracers are anchored. In addition to the concept of anchoring the magnetic structures of sunspot groups to the subsurface layers, the faster rotation rates of sunspots (faster than the photospheric plasma) have been explained in several ways: i) the emerging motion of the magnetic flux loop driven by buoyancy (*e.g.* Moreno-Insertis, 1986; Chou and Fisher, 1989; Shibata *et al.*, 1990), ii) geometrical projection effects (van Driel-Gesztelyi and Petrovay, 1990), iii) the drag due to ambient flows (Meyer *et al.*, 1979; Petrovay *et al.*, 1990), and iv) the interaction between magnetic buoyancy, drag, and Coriolis forces acting on the rising flux tubes (D'Silva and Howard, 1994). However, in spite of these effects, the similarity between the variation in the initial rotation rates of the sunspot groups with their lifetimes and the radial variation of the internal rotation rate determined from helioseismology suggests that the magnetic structures of sunspot groups with successively longer life times (2 – 12 days) are initially anchored in successively deeper layers of the Sun's convection zone (Javaraiah and Gokhale, 1997b). Hiremath (2002) confirmed this result by analyzing a large set of the Greenwich sunspot-group data, and Sivaraman *et al.* (2003) confirmed this by analyzing the sunspot-group data measured at the Mount Wilson and Kodaikanal Observatories. Short-lived and small sunspot groups predominate in a given

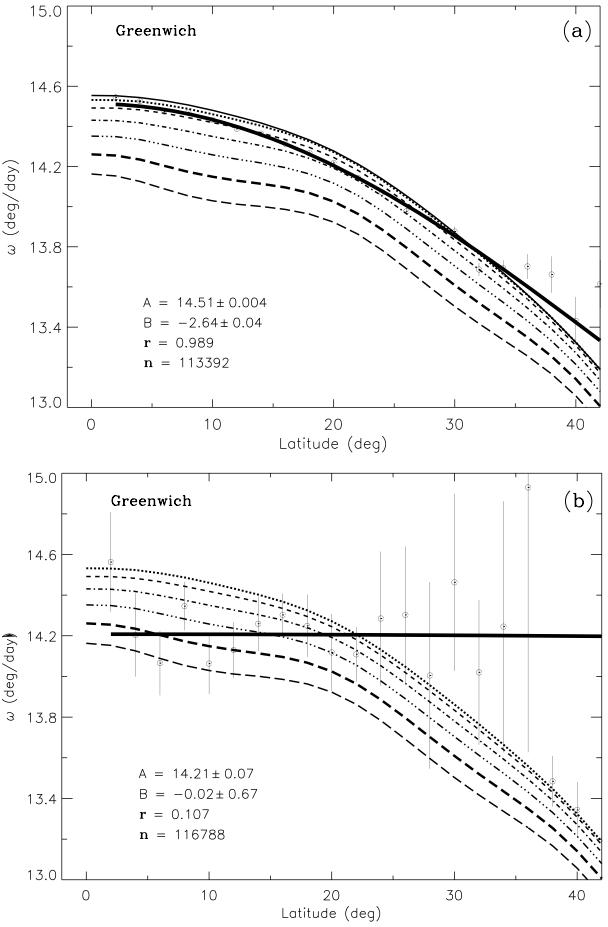


Figure 3 The latitudinal dependence of the mean rotation rates of sunspot groups determined by averaging the daily values of ω obtained from the Greenwich data during the period May 1874 – December 1976 over two-degree latitude intervals $1° - 3°, 3° - 5°, 5° - 7°, \ldots, 40° - 42°$ (plotted at $2°, 4°, 6°, \ldots, 42°$). The error bars represent the standard errors. Panels (a) and (b) represent the data with and without the abnormal values of ω, respectively. The thick solid curve represents the corresponding mean profile deduced from the values (also shown) of the coefficients A and B of Equation (2) obtained from the total number of daily data [n]. The different curves represent the latitudinal dependencies on the mean (over the whole period 1995 – 2009) plasma rotation rates deduced from the GONG data at the same depths as in Figure 1. r represents the coefficient of correlation between the sunspot group and the GONG data around $0.96R_\odot$ at $2°, 4°, 6°, \ldots, 36°$ latitude.

time interval (Javaraiah, 2012). Moreover, in the low latitudes near the base of the convection zone to near $0.8R_\odot$, the internal-rotation rate steeply increases and then gradually increases (or remains almost constant) up to near $0.95R_\odot$. The magnetic structures of the sunspot groups whose lifetimes exceed eight days seem to initially anchor in the layers below $0.8R_\odot$, and those of the sunspot groups with lifetimes of up to eight days seem to initially anchor in the layers above $0.8R_\odot$, as suggested by the pattern of the variation in the initial rotation rates of the sunspot groups with their lifetimes (see Javaraiah and Gokhale, 1997b). In Figure 5 the latitudinal dependencies on the mean initial rotation rates of the four different classes of sunspot groups whose lifetimes are in the ranges one to three, four to five, six to eight, and longer than eight days are compared with the latitudinal dependencies in the internal-rotation rates at depths $0.75R_\odot, 0.8R_\odot, 0.85R_\odot, 0.9R_\odot, 0.95R_\odot$, and $1.0R_\odot$. As can be seen in this figure, the portions up to 25° latitude in the profiles of the mean initial rotation rates of the shorter than and longer than eight days living sunspot groups are somewhat closer to those of the internal rotation at $0.94R_\odot - 0.96R_\odot$ and $0.8R_\odot$, respectively. These results are in general consistent with the results/suggestions in Javaraiah and Gokhale (1997b). The profile of the very short-lived sunspot groups is even higher than the rotation profile at $0.94R_\odot$. The angular motions of super-granules may influence the rotation rates of the very short-

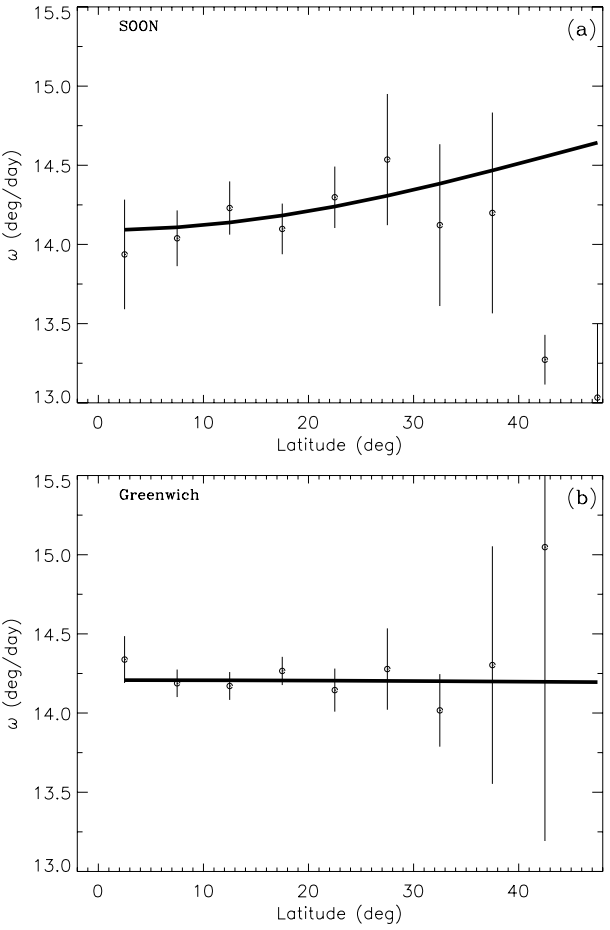


Figure 4 The latitudinal dependence of the mean rotation rates of sunspot groups determined by averaging the daily values of ω, obtained from the SOON (1977 – 2011) and Greenwich (1874 – 1976) sunspot-group data that included the abnormal values of ω, over five-degree latitude intervals, $0° – 5°$, $5° – 10°$, $10° – 15°$, ..., $45° – 50°$ (plotted at 2.5°, 7.5°, 12.5°, ..., 47.5°). The error bars represent the standard errors. The thick solid curve represents the corresponding mean profile deduced from the values of the coefficients A and B of Equation (2) obtained from the total number of daily data (*i.e.* the corresponding values that are given in Figures 2(b) and 3(b)).

lived sunspot groups. In this context it may be worth noting that the magnetohydrodynamic drag force may be strong on the small magnetic structures (D'Silva and Howard, 1994). The results (the profiles of the mean rotation rates of sunspot groups) shown in Figures 2 and 3 are determined from the combined data of all sunspot groups during all days in their respective lifetimes. These results as well as the results (the mean profiles of the initial rotation rates of sunspot groups) shown in Figures 5(a – c) are similar to the results found in most of the previous studies in which the sunspot groups were not sorted according to their lifetimes (Tuominen, 1962; Tuominen and Virtanen, 1988). Ruždjak *et al.* (2004) found that the initial velocity of recurring sunspot groups is higher than that of non-recurring sunspot groups and suggested that the recurring sunspot groups initially anchor at $0.93R_\odot$.

The range of the internal-rotation rate that corresponds to the increase in depth from $0.94R_\odot – 0.75R_\odot$ is approximately the same as that corresponding to the decrease in depth from $0.94R_\odot – 1.0R_\odot$. In view of the result that the magnetic structures of the sunspot groups with successively longer lifetimes (2 – 12 days) are initially anchored in successively deeper layers throughout the Sun's convection zone (Javaraiah and Gokhale, 1997b), in Figure 5 the internal-rotation rates are shown for a wide range of depths. In Figures 2 and 3 the mean rotation rates of the sunspot groups are compared with the internal-rotation rates at

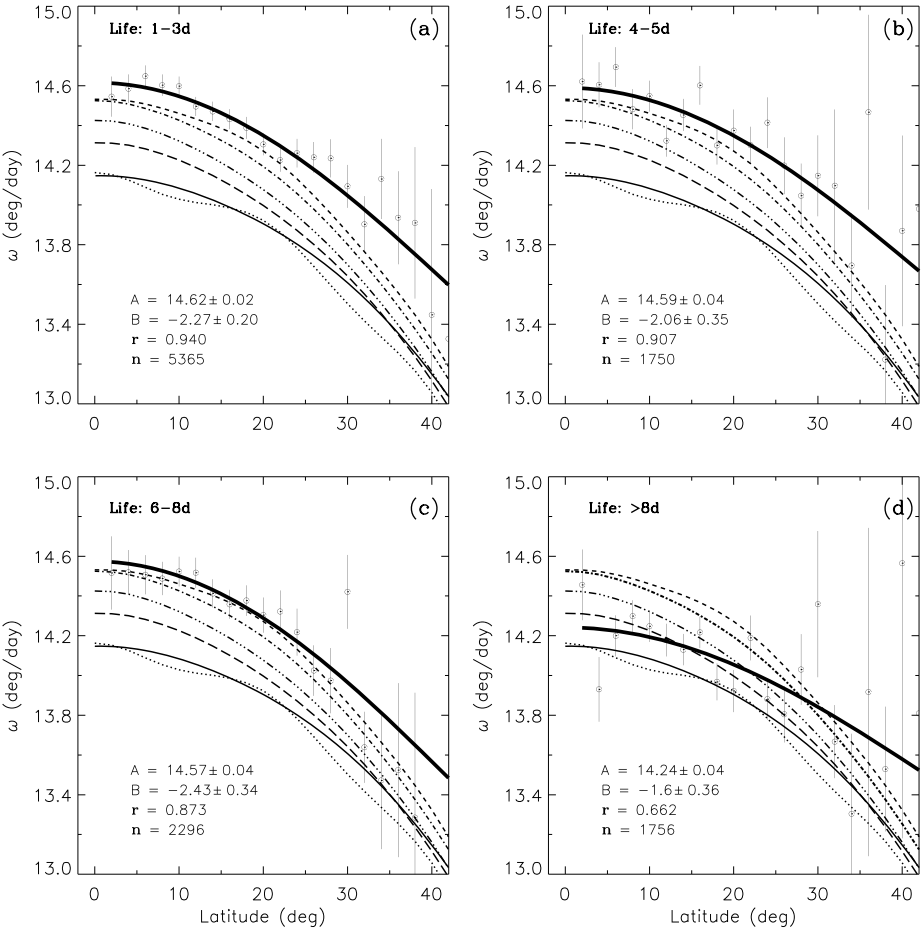


Figure 5 The latitudinal dependence of the mean initial rotation rates of sunspot groups whose life times are in the ranges (a) one to three days, (b) four to five days, (c) six to eight days, and (d) longer than eight days, determined from the combined Greenwich and SOON data during the period May 1874 – December 2011 by averaging the first-day values of ω over two-degree latitude intervals, $1^\circ - 3^\circ, 3^\circ - 5^\circ, 5^\circ - 7^\circ, \ldots, 40^\circ - 42^\circ$ (plotted at $2^\circ, 4^\circ, 6^\circ, \ldots, 42^\circ$). The error bars represent the standard errors. The thick solid curve represents the corresponding mean profile deduced from the values (also shown) of the coefficients A and B of Equation (2) obtained from the total number of daily data [n]. Here the internal-rotational profiles correspond to the different depths: thin-solid, long-dashed, three-dotted-dashed, one-dotted-dashed, dashed, and dotted curves represent the profiles at depths $0.75R_\odot$, $0.80R_\odot$, $0.85R_\odot$, $0.90R_\odot$, $0.95R_\odot$, and $1.0R_\odot$, respectively (sunspot-group data corresponding to $|D_{\rm CM}| \leq 75^\circ$ were used).

a narrow and somewhat shallower region, because the average rotation profile of sunspot groups is mostly contributed by the rotation rates of long-lived (lifetime longer than eight days) sunspot groups during the last few days of their lifetimes (Ruždjak *et al.*, 2004). The rotation rates of the long-lived and large sunspot groups are considerably lower during the initial and final days of their lifetimes.

The derived mean rotation rate of sunspot groups depends to some extent on the cutoff of $D_{\rm CM}$. It decreases for higher $D_{\rm CM}$ values (Ruždjak *et al.*, 2004 and references therein). In all their studies, Javaraiah and co-authors allowed $|D_{\rm CM}| \leq 75^\circ$. Ward (1965, 1966) allowed

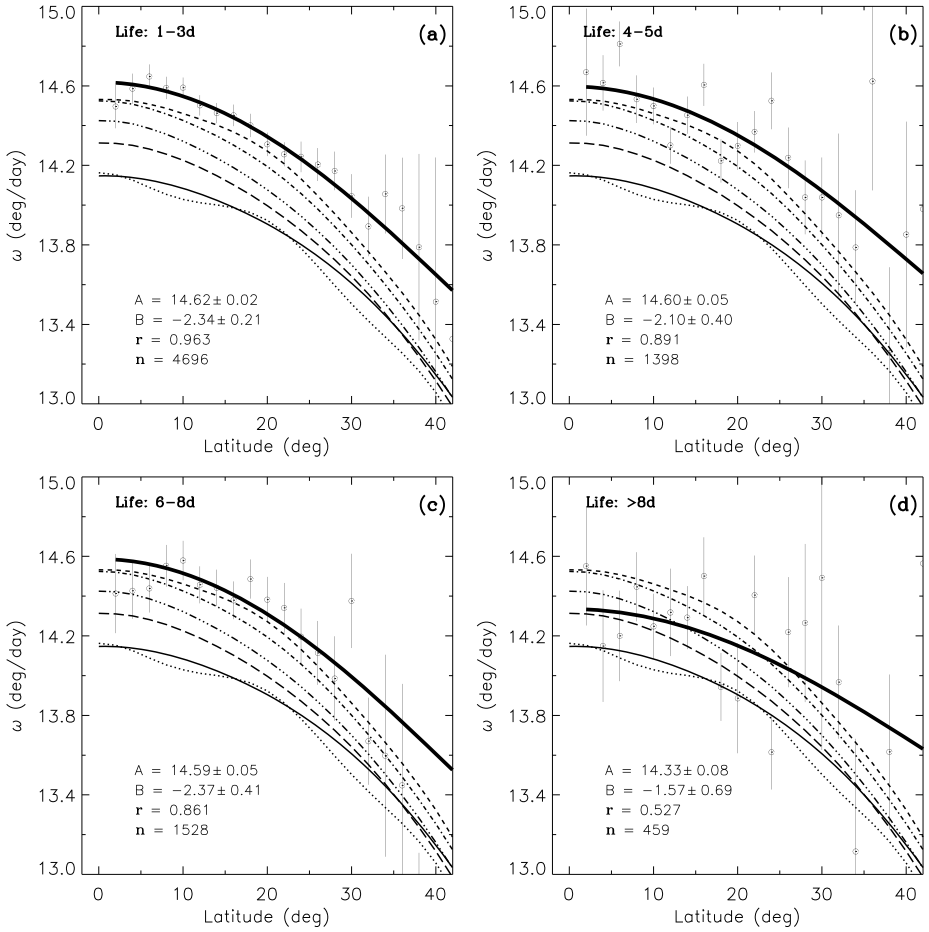


Figure 6 The same as Figure 5, but for the sunspot-group data corresponding to $|D_{\rm CM}| \leq 70^{\circ}$.

$|D_{\rm CM}|$ up to 80°, whereas many others allowed it up to $60^{\circ}-70^{\circ}$ only. A higher cutoff of $D_{\rm CM}$ reduces the contributions of long-lived sunspot groups somewhat, particularly contributions made during their initial and final days, for the derived mean rotation rate. Hence, the derived value of the mean rotation rate of sunspot groups is relatively high. However, it is necessary to verify that the first-day data sample is not contaminated with the sunspot groups arriving from other side of the Sun. To do this, we recalculated the results shown in Figure 5 by restricting $D_{\rm CM}$ up to $\pm 70^{\circ}$ only; the results are shown in Figures 6 and 7. These results are generally the same as those shown in Figure 5. Hence, the conclusions and inferences drawn above from the results shown in Figure 5 also hold for the results shown in Figures 6 and 7.

It is widely believed that magnetic flux, in the form of large flux tubes, emerges at the surface presumably from near the base of the convection zone (where the dynamo process is believed to be taking place) and is responsible for sunspots and other solar activity phenomena (see Rosner and Weiss, 1992; Gough, 2010). There are also suggestions and arguments that the sunspots form immediately beneath the surface (Kosovichev, Duvall, and

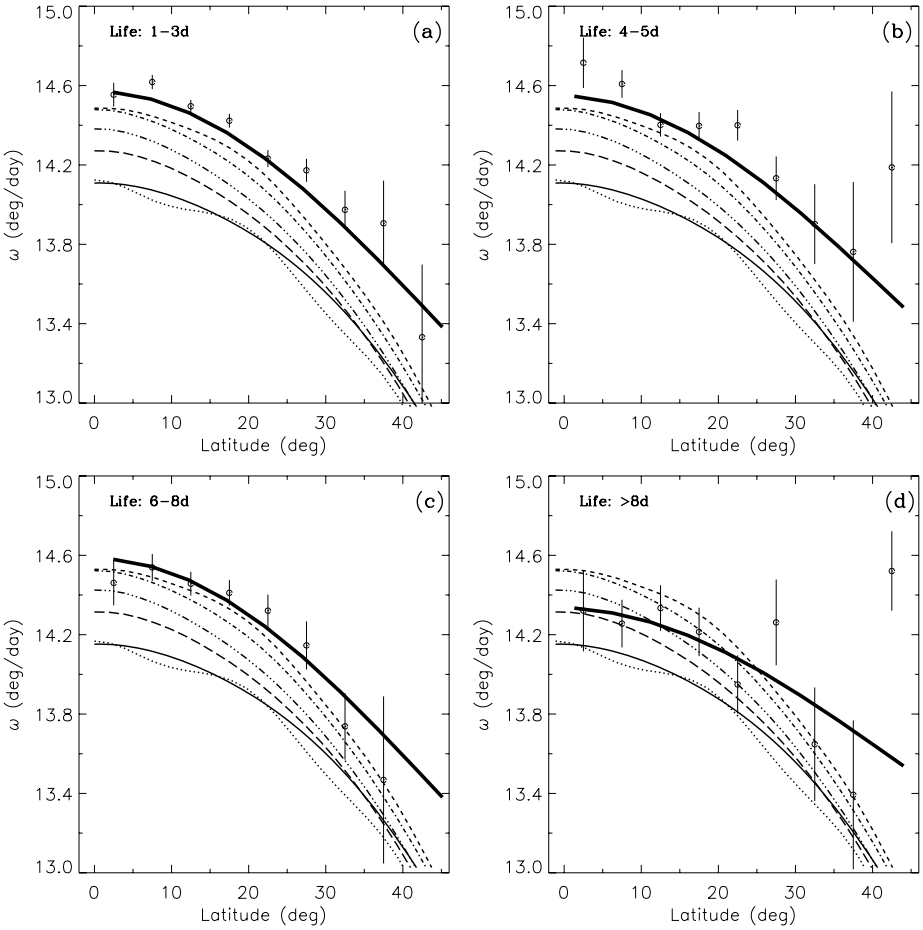


Figure 7 The latitudinal dependence of the mean initial rotation rates of sunspot groups whose life times are in the ranges of (a) one to three days, (b) four to five days, (c) six to eight days, and (d) longer than eight days, determined from the combined Greenwich and SOON data during the period May 1874 – December 2011 by averaging the first-day values of ω over five-degree latitude intervals, $0°-5°, 5°-10°, 10°-15°, \ldots, 45°-45°$ (plotted at $2.5°, 7.5°, 12.5°, \ldots, 47.5°$). The error bars represent the standard errors. Here the internal-rotational profiles correspond to the different depths: thin-solid, long–dashed, three-dotted-dashed, one-dotted-dashed, dashed, and dotted curves represent the profiles at depths $0.75R_\odot, 0.80R_\odot, 0.85R_\odot, 0.90R_\odot, 0.95R_\odot$, and $1.0R_\odot$, respectively. The sunspot-group data corresponding to $|D_{CM}| \leq 70°$ were used. The thick solid curve represents the corresponding mean profile deduced from the values of the coefficients A and B of Equation (2) obtained from the total number of daily data (*i.e.* the corresponding values that are given in Figure 6).

Scherrer, 2000; Kosovichev, 2002) and in different layers throughout the convection zone (Brandenburg, 2005). The results above are in general consistent with these suggestions and arguments, but they also support the idea that large magnetic structures might be generated near the base of the solar convection zone; many of the large magnetic structures may be fragmenting or branching into smaller structures while buoyantly rising through the solar convection zone, *i.e.* small magnetic structures may be fragmented or branched parts of the large magnetic structures (Javaraiah, 2003b). Schüssler and Rempel (2005) argued that the

dynamical disconnection of sunspots from their magnetic roots probably takes place during the final phases of their magnetic structures' buoyant ascent toward the surface.

3.2. Variation in the Equatorial-Rotation Rate

Figure 8 shows the variations in the annual mean values of the equatorial-rotation rate determined from the Doppler-velocity data and the sunspot-group data that did not include the abnormal values of ω. In this figure we also show the variations of the equatorial-rotation rates at $0.96R_\odot$ and $1.0R_\odot$ determined from GONG data for each of 147 intervals of three GONG-months during 1995 – 2009, the variation in the equatorial-rotation rate of the soft X-ray corona determined from the *Yohkoh*/SXT solar full-disk images for the period 1992 – 2001 and from the rotation rates of SBCS traced in SOHO/EIT images during the period 1998 – 2006 (Jurdana-Šepić *et al.*, 2011), and the annual mean sunspot number (Figure 8(b)) to study solar-cycle behavior of the equatorial-rotation rate. Since the mean value ($\approx 14.5^\circ$ day^{-1}) of the equatorial-rotation rate determined from the sunspot-group data during the period 1985 – 2010 is close to the mean Ω_0 at $0.96R_\odot$ (*cf.* Figures 2 and 3), the variation in the former is compared with that of the latter in addition to comparing it with the variation at $1.0R_\odot$ (due to the scaling problem the variation in Ω_0 at $1.0R_\odot$ is shown in Figure 8(a) instead of in Figure 8(b)). However, as can be seen in Figure 1, the patterns of the variations of the equatorial rotation rates at different depths are largely similar.

As can be seen in Figure 8, the mean value of the equatorial-rotation rate determined from the sunspot-group data is substantially higher (although it is lower during Cycle 22 than in the last 11 solar cycles (Javaraiah, Bertello, and Ulrich, 2005b; Suzuki, 2012)), and the amplitude of its variation is also about ten times higher than that determined from the Doppler-velocity measurements. The variation in the equatorial-rotation rate determined from the sunspot-group data steeply decreased during the rising phase of the Solar Cycle 22. It attained minimum at the maximum of this cycle and remained at approximately the same level up to the end of the cycle, then it steeply increased in the beginning of Cycle 23. The overall pattern suggests that there exists a quasi-11-year cycle in the equatorial-rotation rate (correlation coefficient ≈ -0.4, between sunspot number and A). At the beginning of Cycle 24 the equatorial-rotation rate is considerably higher than it was during the last about twenty years (including the beginnings of Cycles 22 and 23). The variation pattern in the equatorial-rotation rate is consistent with the well-known result of a higher rotation rate during the cycle minimum than during the maximum, which has been found in many studies (*e.g.* Brajša, Ruždjak, and Wöhl, 2006). However, it should be noted that the minimum years of the solar cycles contain mainly small sunspot groups (Javaraiah, 2012). These rotate faster than large sunspot groups (*e.g.* Howard, Gilman, and Gilman, 1984; Javaraiah and Gokhale, 1997b, see also Section 3.1 above). This property of the sunspot groups most probably has an influence on the aforementioned pattern of the solar cycle variation in the surface equatorial-rotation rate determined from the sunspot-group data.

The increases of rotational velocities in 2001 – 2002 and 2008 – 2009 are most pronounced in GONG results at $1.0R_\odot$. Similar increases also exist in the result obtained by using sunspot-group data (see Figure 8(a)). In fact, there is a reasonable agreement between the pattern of the variations in the equatorial-rotation rates determined from the sunspot-group data and the GONG data [Ω_0]. However, there is a considerable difference in the positions of the minima of these variations, giving the impression that the internal equatorial-rotation rate leads the equatorial-rotation rate of the sunspot groups by one to two years (the reason for this is not known to us).

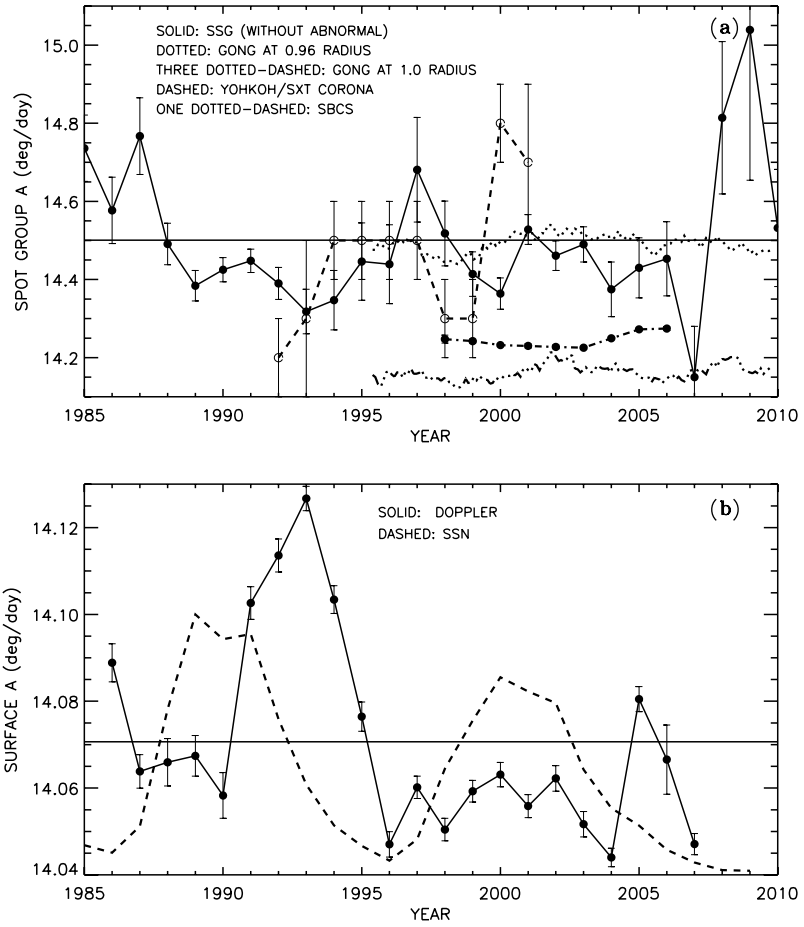


Figure 8 (a) Filled circle-solid curve: the variation in the annual equatorial-rotation rate A determined from the sunspot group (SSG) data after excluding the abnormal values of ω, which were calculated by using Equation (1). Open circle-dashed curve: the variations in the equatorial-rotation rate of the soft X-ray corona determined from *Yohkoh*/SXT full-disk images for the years 1992 – 2001 by Chandra, Vats, and Iyer (2010). Filled circle-one dotted-dashed curve: the variation in the annual mean equatorial-rotation rate determined from the data of SBCS, which were traced in SOHO/EIT images during the period 1998 – 2006 (Jurdana-Šepić *et al.*, 2011). In each of these cases the error-bars represent the values of 1σ obtained from the linear least-squares fits of the data (for SBCS σ has very low values). The dotted and three dotted-dashed curves are the variations in the equatorial-rotation rates at $0.96R_{\odot}$ and $1.0R_{\odot}$, determined from GONG data for each of the 147 intervals of the three GONG-months during 1995 – 2009 (Antia and Basu, 2010). The solid horizontal line is drawn at the mean values of $14.5° \pm 0.2°$ day^{-1}, determined from the yearly values of the sunspot-group data. (b) The solid curve represents the variation in the annual mean A determined from the corrected Mount Wilson Doppler-velocity data (Javaraiah *et al.*, 2009). The error bars represent one standard error (σ has very high values). The dashed curve represents the variation in the annual mean international sunspot number (SSN), which is normalized to the scale of A. The solid horizontal line is drawn at the mean value of A, $14.07° \pm 0.02°$ day^{-1}, determined from the yearly values. The Mount Wilson Doppler-velocity data in 2007 are available only up to March.

The annual variation in the equatorial-rotation rate determined from the Doppler-velocity data also steeply decreased during the rising phase and attained minimum at the sunspot maximum of Solar Cycle 22, but it increased sharply from 1990 to 1993 and suddenly

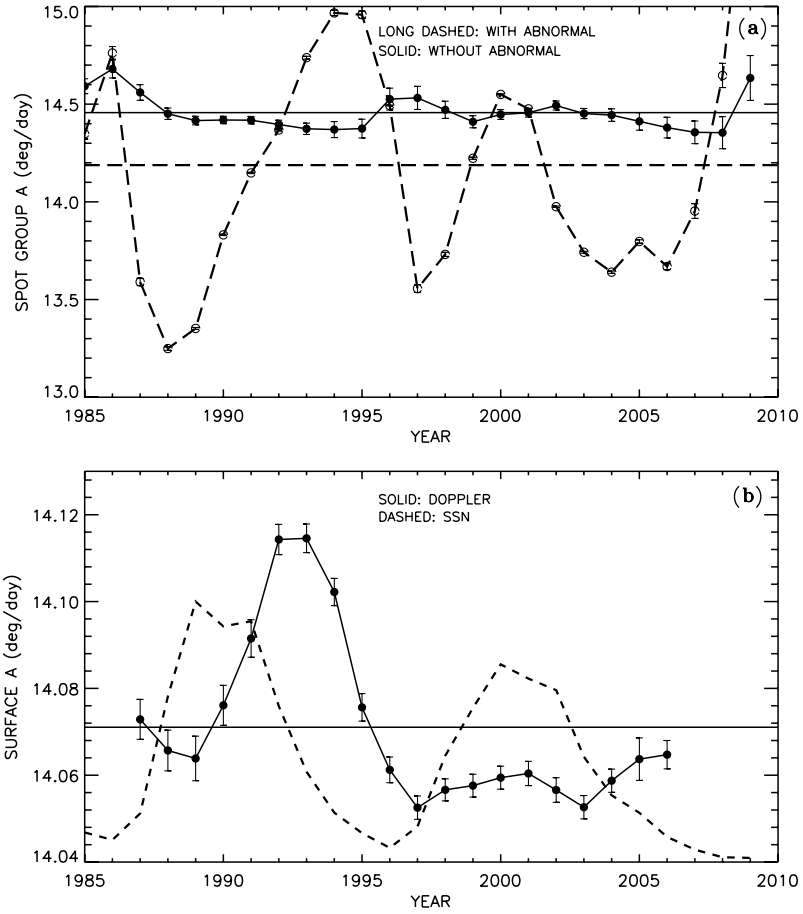


Figure 9 (a) The open circle-long-dashed and the filled circle-solid curves represent the variations in A derived from the sunspot group (SSG) data with and without abnormal ω values ($> 3^\circ$ day^{-1}), respectively, in three-year MTIs successively shifted by one year. In the first case the error-bars represent the standard error (because of the high σ value) and in the other case they represent the values of 1σ. The solid and long-dashed horizontal lines are drawn at the corresponding mean values of A, $14.46^\circ \pm 0.09^\circ$ day^{-1} and $14.19^\circ \pm 0.61^\circ$ day^{-1}. (b) The solid curve represents the three-year smoothed A derived from the annual mean values of A (shown in Figure 8(b)) determined from the Mount Wilson Doppler-velocity data (error bars represent one standard error). The solid horizontal line is drawn at the corresponding mean value of A, 14.07° day^{-1}. The dashed curve represents the variation in the annual mean international sunspot number (SSN), which is normalized to the scale of A. The Mount Wilson Doppler-velocity data in 2007 are available only up to March.

dropped from 1993 to 1994. Overall, the equatorial-rotation ration rate is low during Cycle 23, but there is an indication that the variation in the equatorial-rotation rate is in phase with the sunspot number in the interval 1998 – 2004 (however, the value of A suddenly increased around 2005 and then decreased). This behavior is opposite to that in the previous cycle. The overall pattern of the annual variation indicates that there is a five- to six-year periodicity in the equatorial-rotation rate during Cycle 22.

In Figure 9 we compare the variations in the equatorial-rotation rate determined from the Doppler-velocity data and the sunspot-group data that included the abnormal values of ω (*i.e.* $|\omega| > 3^\circ\,\mathrm{day}^{-1}$) calculated by using Equation (3). As already mentioned above, including the abnormal values of ω increases the uncertainties of the derived coefficients A and B, although the size of the data increased by 4 % – 5 % (the highest is 8.4 % around 2009). Hence, with the abnormal ω values the errors [σ] of the individual yearly averages are so large that they are not significantly (statistically) different from each other, so that the variations cannot be examined using them. Therefore, the values of A determined by binning the data into three-year moving time intervals (MTIs) were used. In Figure 9 the variation in A determined from the data that do not include the abnormal values and were binned into three-year MTIs is also shown, to indicate the difference between the yearly (*cf.* Figure 8(a)) and the three-year MTIs data. The equatorial-rotation rate determined from the Doppler-velocity data is shown for the three-year smoothed time series obtained from the annual values (*cf.* Figure 8(b)). As can be seen in Figure 9, the variation patterns in the equatorial-rotation rates derived from the sunspot-group data and the Doppler-velocity data closely resemble one another (correlation coefficient ≈ 0.5). In fact, during Cycle 23 the shapes of the curves are almost the same (the Doppler-velocity measurements are more accurate after 1995). This result indicates that the temporal variation in the equatorial-rotation rate determined from the Doppler-velocity measurements is of solar origin, which has been doubtful so far for the reason mentioned in Section 1. However, this result and conclusion is only suggestive and not compelling, because the uncertainties in A determined from the sunspot-group data that included the abnormal angular motions of the sunspot groups are large (in the figure the error bar represents one standard error level; the corresponding σ is very high). In addition, the abnormal angular motions of the sunspot groups may not represent the angular motions of the surface plasma. These motions may be the abnormal proper motions of the sunspot groups that took place at the locations of eruptive solar phenomena such as flares. The Doppler-velocity measurements may also be contaminated with the contributions from these plasma motions (the spikes in the daily Mt. Wilson velocity data were removed. Further trimming will create more data gaps and make the data more unrealistic).

On the other hand, the mean values $14.09^\circ \pm 0.13^\circ$ and $14.21^\circ \pm 0.07^\circ\,\mathrm{day}^{-1}$ (see Figures 2 and 3) of the equatorial-rotation rates determined from the SOON and Greenwich daily sunspot-group data, respectively, which included the abnormal ω values are significantly (more than 2σ level) lower than the corresponding values $14.45^\circ \pm 0.008^\circ$ and $14.51^\circ \pm 0.004^\circ$ determined from the data that did not include the abnormal values. Overall, it seems that mean values determined from the SOON data that included the abnormal ω values is somewhat closer to the mean value ($14.07^\circ \pm 0.02^\circ\,\mathrm{day}^{-1}$) of the equatorial-rotation rate determined from the Doppler-velocity data. In Figure 10 (the histogram of the distribution of the abnormal angular velocity values) there is a suggestion that the peaks of the open circle-long dashed curve in Figure 9 included the majority of the abnormal velocity values (in the range $10^\circ - 20^\circ$) that are somewhat consistent with the rotation rate of the Sun. Therefore, the similarity in the variations of the Doppler-velocity data and the sunspot-group data that included the abnormal motions suggest that the former may be largely of solar origin and are not spurious. Nevertheless, the highly unrealistic latitudinal gradient of the rotation (see Figures 2 and 3) that is obtained from the sunspot-group data that included the abnormal ω-values may have a strong influence on the corresponding equatorial-rotation rate, making the latter unreliable. For this reason, and because of the wide range and the large uncertainties in annual mean values (including all errors of measurements both random and systematic), the reality and reliability of the variation in the sunspot-group data that included the abnormal motions is doubtful. Hence, the similarity in the solar-cycle variations in the Doppler-velocity data and the sunspot-group data that included the abnormal

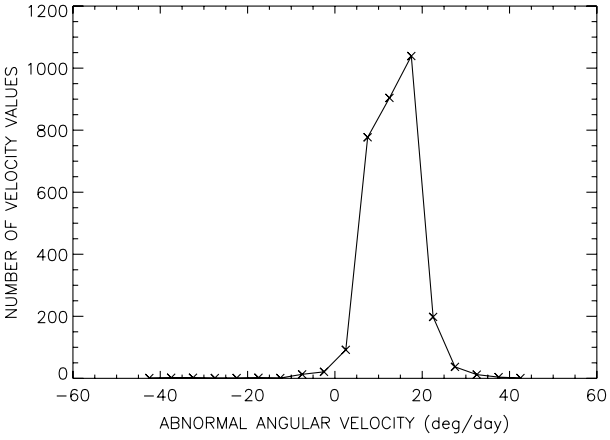


Figure 10 Distribution of the abnormal angular velocity values determined from the SOON sunspot-group data during the period 1 January 1977 to 31 January 2012

motions may instead suggest that the variation in the Doppler-velocity data (before 1995) could be contributed by the large uncertainties in this data.

As can be seen in Figure 8, the equatorial-rotation rate determined from the *Yohkoh*/SXT full-disk images for the period 1992 – 2001 by Chandra, Vats, and Iyer (2010) closely matches that determined from the sunspot-group data for all years except for the year 2000, where it has a high value. Close to this year there are humps in the variation of A determined from the sunspot group and the Doppler-velocity data as well, although they are small. Thus, the coronal rotation is strongly related to the angular motions of the surface magnetic features. (The phases of Ω_0 and A of the corona apparently agree even better, the agreement is particularly strong in minima.) However, the pattern of A determined from the SBCS seems to be considerably different from those of other data sets, particularly in the period 2001 – 2003. It even seems to be anti-correlated with that of the GONG data. The reason for this is not known. The similarity between the variation in the equatorial-rotation rate of sunspot groups and the corona and the variation in the internal equatorial-rotation rate is only in shape. The amplitude of the variation in the internal equatorial-rotation rate is very small.

4. Conclusions

The following conclusions can be drawn from our analysis:

i) A large portion (up to $\approx 30^\circ$ latitude) of the mean differential-rotation profile of the sunspot groups lies between those of the internal differential-rotation rates at $0.94R_\odot$ and $0.98R_\odot$. The portions up to 25° latitude on the mean profiles of the initial rotation rates of the up to eight days and longer than eight days living sunspot groups are close to those of the internal rotation near $0.96R_\odot$ and $0.8R_\odot$, respectively.

ii) At the end of Cycle 23 and at the beginning of Cycle 24, the value of the equatorial-rotation rate determined from the sunspot-group data is considerably higher than that at the beginning of Cycle 23. Overall, the variation pattern in the equatorial-rotation rate during Cycles 22 and 23 resembles the pattern of the known solar-cycle variation in the equatorial-rotation rate determined from sunspot data.

iii) There is a reasonable agreement between the variation pattern in the equatorial-rotation rate determined from the sunspot-group data and that determined from the GONG data [Ω_0]. Moreover, the variation in the internal equatorial-rotation rate apparently

leads the equatorial-rotation rate of sunspot groups by one to two years (the reason for this is not known to us).

iv) The pattern of the known solar-cycle variation in the equatorial-rotation rate of the solar corona determined from the *Yohkoh*/SXT full-disk images and SBCS closely matches that determined from the sunspot-group data (except during the period 2001 – 2003). This indicates that the coronal rotation is strongly related to the rotational motion of the surface magnetic features.

v) The variation in the equatorial-rotation rate determined from the Mount Wilson Doppler-velocity data substantially differs from the corresponding variation in the equatorial-rotation rate determined from the sunspot-group data without the values of the abnormal angular motions ($> |3^\circ|\,\mathrm{day}^{-1}$) of the sunspot groups, whereas it closely resembles the corresponding variation determined from the sunspot-group data with the values of the abnormal angular motions.

vi) Conclusion v) above may suggest that the solar-cycle variation in the surface equatorial-rotation rate determined from the Doppler-velocity measurements (before 1995) is caused by the inconsistency and uncertainties in the data (but this needs more studies to find a definite answer).

Acknowledgements The author thanks the anonymous referee for the detailed comments and useful suggestions, and H.M. Antia for providing the values of the internal-rotation rates that he has determined from GONG data. The author also thanks Luca Bertello for helpful comments and suggestions, B.A. Varghese for his help in making the figures, and the organizers of the LWS/SDO-3/SOHO26/GONG-2011 workshop for kindly providing a partial financial support to attend the workshop. This work used data obtained by the Global Oscillation Network Group (GONG) Program, managed by the National Solar Observatory (NSO), which is operated by AURA, Inc. under a cooperative agreement with the National Science Foundation. The data were acquired by instruments operated by the Big Bear Solar Observatory, High Altitude Observatory, Learmonth Solar Observatory, Udaipur Solar Observatory, Instituto de Astrofísico de Canarias, and Cerro Tololo Interamerican Observatory.

References

Antia, H.M., Basu, S., Chitre, S.M.: 2008, *Astrophys. J.* **681**, 680. doi:10.1086/588523.
Antia, H.M., Basu, S.: 2010, *Astrophys. J.* **720**, 494. doi:10.1088/0004-637X/720/1/494.
Babcock, H.W.: 1961, *Astrophys. J.* **133**, 572. doi:10.1086/147060.
Balthasar, H., Vázquez, M., Wöhl, H.: 1986, *Astron. Astrophys.* **155**, 87.
Brandenburg, A.: 2005, *Astrophys. J.* **625**, 539. doi:10.1086/429584.
Brajša, R., Ruždjak, D., Wöhl, H.: 2006, *Solar Phys.* **237**, 365. doi:10.1007/s11207-006-0076-7.
Carrington, R.C.: 1863 *Observations of the Spots on the Sun*, Williams and Norgate, London.
Chandra, S., Vats, H.O., Iyer, K.N.: 2010, *Mon. Not. Roy. Astron. Soc.* **407**, 1108. doi:10.1111/j.1365-2966.2010.16947.x.
Chou, D.-Y., Fisher, G.H.: 1989, *Astrophys. J.* **341**, 533. doi:10.1086/167514.
Dikpati, M., Gilman, P.A.: 2006, *Astrophys. J.* **649**, 498. doi:10.1086/506314.
D'Silva, S., Howard, R.F.: 1994, *Solar Phys.* **151**, 213. doi:10.1007/BF00679072.
Foukal, P.: 1972, *Astrophys. J.* **173**, 439. doi:10.1086/151435.
Godoli, G., Mazzucconi, F.: 1979, *Solar Phys.* **64**, 247. doi:10.1007/BF00151436.
Gough, D.O.: 2010, In: Hasan, S.S., Rutten, R.J. (eds.) *Magnetic Coupling Between the Interior and the Atmosphere of the Sun*, *Astrophys. Space Sci. Proc.* **SSA 1570-6591**, Springer, Berlin, 37. doi:10.1007/978-3-642-02859-5_4.
Gupta, S.S., Sivaraman, K.R., Howard, R.: 1999, *Solar Phys.* **188**, 225. doi:10.1023/A:1005229124554.
Hiremath, K.M.: 2002, *Astron. Astrophys.* **386**, 674. doi:10.1051/0004-6361:20020276.
Hiremath, K.M., Suryanarayana, G.S.: 2003, *Astron. Astrophys.* **411**, L497. doi:10.1051/0004-6361:20031618.
Howard, R., LaBonte, B.J.: 1980, *Astron. Astrophys.* **239**, L33. doi:10.1086/183286.
Howard, R., Gilman, P.I., Gilman, P.A.: 1984, *Astrophys. J.* **283**, 373. doi:10.1086/162315.

Howe, R., Christensen-Dalsgaard, J., Hill, F., Komm, R.W., Larsen, R.M., Schou, J., Thompson, M.J., Toomre, J.: 2000, *Astrophys. J. Lett.* **533**, L163. doi:10.1086/312623.
Javaraiah, J.: 2003a, *Solar Phys.* **212**, 23. doi:10.1023/A:1022912430585.
Javaraiah, J.: 2003b, In: Pevtsov, A.A., Uitenbroek, H. (eds.) *Current Theoretical Models and Future High Resolution Solar Observations: Preparing for ATST* **CS-286**, Astron. Soc. Pac., San Francisco, 325.
Javaraiah, J.: 2005, *Mon. Not. Roy. Astron. Soc.* **362**, 1311. doi:10.1111/j.1365-2966.2005.09403.x.
Javaraiah, J.: 2011, *Adv. Space Res.* **48**, 1032. doi:10.1016/j.asr.2011.05.004.
Javaraiah, J.: 2012, *Solar Phys.* **281**, 827. doi:10.1007/s11207-012-0106-6.
Javaraiah, J., Bertello, L., Ulrich, R.K.: 2005a, *Astrophys. J.* **626**, 579. doi:10.1086/429898.
Javaraiah, J., Bertello, L., Ulrich, R.K.: 2005b, *Solar Phys.* **232**, 25. doi:10.1007/s11207-005-8776-y.
Javaraiah, J., Gokhale, M.H.: 1995, *Solar Phys.* **158**, 173. doi:10.1007/BF00680841.
Javaraiah, J., Gokhale, M.H.: 1997a, *Solar Phys.* **170**, 389. doi:10.1023/A:1004928020737.
Javaraiah, J., Gokhale, M.H.: 1997b, *Astron. Astrophys.* **327**, 795.
Javaraiah, J., Gokhale, M.H.: 2002, *The Sun's Rotation*, Nova Science, New York.
Javaraiah, J., Komm, R.W.: 1999, *Solar Phys.* **184**, 41. doi:10.1023/A:1005028128077.
Javaraiah, J., Ulrich, R.K.: 2006, *Solar Phys.* **237**, 245. doi:10.1007/s11207-006-0130-5.
Javaraiah, J., Ulrich, R.K., Bertello, L., Boyden, J.E.: 2009, *Solar Phys.* **257**, 61. doi:10.1007/s11207-009-9342-9.
Jurdana-Šepić, R., Brajša, R., Wöhl, H., Hanslmeier, A., Poljančić, I., Svalgaard, L., Gissot, S.F.: 2011, *Astron. Astrophys.* **534**, A17. doi:10.1051/0004-6361/201014357.
Kambry, M.A., Nishikawa, J.: 1990, *Solar Phys.* **126**, 89. doi:10.1007/BF00158300.
Karak, B.B.: 2010, *Astrophys. J.* **724**, 1021. doi:10.1088/0004-637X/724/2/1021.
Kosovichev, A.G.: 2002, *Astron. Nachr.* **323**, 186. doi:10.1002/1521-3994(200208)323:3/4<186::AID-ASNA186>3.0.CO;2-I.
Kosovichev, A.G., Duvall, T.L. Jr., Scherrer, P.H.: 2000, *Solar Phys.* **192**, 159. doi:10.1023/A:1005251208431.
Meunier, N.: 2005, *Astron. Astrophys.* **436**, 1075. doi:10.1051/0004-6361:20042414.
Meyer, F., Schmidt, H.U., Simon, G.V., Weiss, N.O.: 1979, *Astron. Astrophys.* **76**, 35.
Moreno-Insertis, F.: 1986, *Astron. Astrophys.* **166**, 291.
Petrovay, K., Brown, J.C., van Driel-Gesztelyi, L., Fletcher, L., Marik, M., Stewart, G.: 1990, *Solar Phys.* **127**, 51. doi:10.1007/BF00158513.
Rosner, R., Weiss, N.O.: 1992, In: Harvey, K.L. (ed.) *The Solar Cycle* **CS-27**, Astron. Soc. Pac., San Francisco, 511.
Ruždjak, D., Ruždjak, V., Brajša, R., Wöhl, H.: 2004, *Solar Phys.* **221**, 225. doi:10.1023/B:SOLA.0000035066.96031.4f.
Schüssler, M., Rempel, M.: 2005, *Astron. Astrophys.* **441**, 337. doi:10.1051/0004-6361:20052962.
Shibata, K., Nozawa, R., Matsumoto, R., Sterling, A.C., Tajima, T.: 1990, *Astrophys. J.* **351**, L25. doi:10.1086/185671.
Sivaraman, K.R., Sivaraman, H., Gupta, S.S., Howard, R.: 2003, *Solar Phys.* **214**, 65. doi:10.1023/A:1024075100667.
Suryanarayana, G.S.: 2010, *New Astron.* **15**, 313. doi:10.1016/j.newast.2009.09.004.
Suzuki, M.: 2012, *Solar Phys.* **278**, 257. doi:10.1007/s11207-012-9946-3.
Tuominen, J.: 1962, *Z. Astrophys.* **55**, 110.
Tuominen, I., Virtanen, H.: 1988, *Adv. Space Res.* **8**, 141. doi:10.1016/0273-1177(88)90183-4.
Ulrich, R.K.: 2001, *Astrophys. J.* **560**, 466. doi:10.1086/322524.
Ulrich, R.K., Boyden, J.E.: 2005, *Astrophys. J. Lett.* **620**, L123. doi:10.1086/428724.
van Driel-Gesztelyi, L., Petrovay, K.: 1990, *Solar Phys.* **126**, 285. doi:10.1007/BF00153051.
Ward, F.: 1965, *Astrophys. J.* **141**, 534. doi:10.1086/148143.
Ward, F.: 1966, *Astrophys. J.* **145**, 416. doi:10.1086/148783.

DOI 10.1007/978-1-4899-8005-2_14
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-013-0338-0

Recovering Joy's Law as a Function of Solar Cycle, Hemisphere, and Longitude

B.H. McClintock · A.A. Norton

Received: 21 February 2012 / Accepted: 28 May 2013 / Published online: 10 July 2013

Abstract Bipolar active regions in both hemispheres tend to be tilted with respect to the East–West Equator of the Sun in accordance with Joy's law, which describes the average tilt angle as a function of latitude. Mt. Wilson Observatory data from 1917 – 1985 are used to analyze the active-region tilt angle as a function of solar cycle, hemisphere, and longitude, in addition to the more common dependence on latitude. Our main results are as follows: i) We recommend a revision of Joy's law towards a weaker dependence on latitude (slope of 0.13 – 0.26) and without forcing the tilt to zero at the Equator. ii) We determine that the hemispheric mean tilt value of active regions varies with each solar cycle, although the noise from a stochastic process dominates and does not allow for a determination of the slope of Joy's law on an 11-year time scale. iii) The hemispheric difference in mean tilt angles, $1.1^\circ \pm 0.27$, over Cycles 16 to 21 was significant to a three-σ level, with average tilt angles in the Northern and Southern hemispheres of $4.7^\circ \pm 0.26$ and $3.6^\circ \pm 0.27$, respectively. iv) Area-weighted mean tilt angles normalized by latitude for Cycles 15 to 21 anticorrelate with cycle strength for the southern hemisphere and whole-Sun data, confirming previous results by Dasi-Espuig *et al.* (*Astron. Astrophys.* **518**, A7, 2010). The Northern Hemispheric mean tilt angles do not show a dependence on cycle strength. v) Mean tilt angles do not show a dependence on longitude for any hemisphere or cycle. In addition, the standard deviation of the mean tilt is 29 – 31° for all cycles and hemispheres, indicating that the scatter is due to the same consistent process even if the mean tilt angles vary.

Keywords Joy's law · Sunspots

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

B.H. McClintock (✉)
Centre for Astronomy, James Cook University, Townsville, QLD 4810, Australia
e-mail: u1049686@umail.usq.edu.au

A.A. Norton
HEPL, Stanford University, Stanford, CA 94305, USA
e-mail: aanorton@stanford.edu

1. Introduction

It is believed that magnetic fields generated at the base of the convective zone become buoyant and rise as toroidal flux tubes. Oriented in the East–West direction, flux-tube loops emerge from the solar surface to form sunspots. Observations of bipolar sunspots, on average, show leading spots closer to the Equator than following spots. Known as Joy's law, this was first published by Hale *et al.* (1919) after statistical analysis showed that the mean tilt angle of bipolar sunspots increased with latitude in both hemispheres. Joy's law has traditionally been interpreted as the Coriolis force operating in the separate hemispheres on motion in the rising magnetic flux tubes. Coriolis forces dissipate once flux tube emergence ends and tilt should relax to zero, but observations made by Howard (2000) showed tilt trending towards average, nonzero values after emergence. Babcock (1961) proposed that tilt is due to a spiral orientation of initial magnetic-field lines prior to emergence. Tilt angle dependence on the latitude has been confirmed by many authors (Howard, 1991; Wang and Sheeley, 1991; Sivaraman, Gupta, and Howard, 1993, 1999) and provides constraints on the magnetic-field strength of the flux tubes which emerge to form the observed active regions (D'Silva and Howard, 1993; Schüssler *et al.*, 1994).

We analyze the tilt angles independently by hemisphere. Since the transport of magnetic fields in Babcock–Leighton dynamo models (Babcock, 1961; Leighton 1964, 1969) is partly achieved through a meridional-circulation cell seated in an individual hemisphere, the Northern and Southern Hemispheres can become decoupled to some degree (Dikpati and Gilman, 2001; Chatterjee, Nandy, and Choudhuri, 2004). While it is obvious from the butterfly diagram that some degree of cross-hemispheric coupling prevents the hemispheres from becoming grossly out of phase (at least for solar cycles observed since the late 1800s), nevertheless, hemispheric phase lags are observed. For example, the polar-field reversals in the Northern and Southern Hemispheres occurred half a year apart during Cycle 23 (Durrant and Wilson, 2002; Norton and Gallagher, 2010), and the Northern Hemisphere led the Southern by 19 months in the declining phase of Cycle 20 (Norton and Gallagher, 2010), while hemispheres have been observed to be up to two years out of phase.

In addition to temporal phase lags between the hemispheres, it is common that one hemisphere dominates the other in the production of sunspot numbers and sunspot area (Temmer *et al.*, 2006). McIntosh *et al.* (2013) suggest that hemispheric asymmetry is a normal ingredient of the solar cycle and has important consequences in the structuring of the heliosphere. Charbonneau (2007) finds a "rich variety of behavior characterizing the two-hemisphere dynamo solution" including intermittency (a cessation of sunspot production similar to the Maunder Minimum) operating independently in separate hemispheres. Data analysis separated into hemispheres is critical to avoid blurring a signal that may be distinct in isolated hemispheres.

A tipping ($m = 1$ mode) or warping ($m > 1$ mode) of the toroidal magnetic band in the solar interior with respect to the equatorial plane in one or both hemispheres due to an MHD instability, as proposed by Cally, Dikpati, and Gilman (2003) and observed by Norton and Gilman (2005), would impart initial tilt angles dependent on longitude prior to a flux rope's rise through the convection zone. An $m = 0$ instability is expected for toroidal fields stronger than 50 kG on average, whereas $m > 0$ is more likely for weaker toroidal fields. The growth of the tipping or deformation, and whether it is symmetric or asymmetric across the Equator, depends in part upon the width of the toroidal band (Cally, Dikpati, and Gilman, 2003). A toroidal field tipped with respect to the Equator would not produce a different mean tilt angle averaged over longitude and latitude for a given cycle, but it would increase the scatter of the mean tilt angle. It could also explain why the tilt does not relax to zero

after the active region has fully emerged, as observed by Kosovichev and Stenflo (2008) and summarized nicely as follows: "It may be that Joy's law reflects not the dynamics of the rising flux tube, but the spiral orientation of the toroidal magnetic field lines below the surface as suggested by Babcock (1961)." We argue that Joy's law is due to a combination of both the Coriolis force's acting on the rising flux as it rises as well as an initial tilt imparted to the flux rope from the toroidal geometry that it retains. We search for a dependence of tilt angle on longitude as well as a dependence of noise in the mean tilt angle as a function of solar-cycle strength. It also appears that the tilt angle is inherently noisy, presumably due to the turbulent convection that is encountered by the flux ropes during their rise. However, Stenflo and Kosovichev (2012) argue that the many examples of large bipolar active regions with tilts that differ from the expected Joy's law angle by 90° are not simply regions buffeted by turbulent convection, but instead are regions from a different flux system that coexists at any given latitude.

2. Recovering Joy's Law

Furthering work by Howard (1996) and others, we examine bipolar active region tilt angles observed at the Mt. Wilson Observatory. We also record tilt angle dependence on hemisphere, solar cycle, latitude, and longitude (dependence on longitude is discussed in Section 4). In Figure 1, mean tilt-angle values as a function of latitude for each hemisphere are shown averaged over Solar Cycles 16 to 21 for data collected at the Mt. Wilson Observatory between 1923 and 1985. Cycle 15 began in 1913, but Mt. Wilson observations for this data did not begin until 1917, near solar maximum of this cycle. We excluded Cycle 15 from this analysis, as it is an incomplete representation of a solar cycle. Mt. Wilson data are not available for the end of Solar Cycle 21 from January 1986 to September 1986. However, the monthly smoothed sunspot number had dropped to around 12.2 by January 1986. At most, this would have amounted to approximately 110 spots *versus* the 4000 pairs in this cycle. After removing single sunspots from analysis, the effect on overall results would have been negligible; therefore, Cycle 21 is included. The only regions excluded were individual spots, *i.e.* groups that did not have at least one sunspot in both the leading and following portions of a group. These were indexed in the Mt. Wilson data with a tilt angle of zero. Dasi-Espuig *et al.* (2010) thoroughly investigated entries with a zero tilt angle and found only one data point that corresponded to a true tilt value of zero; all others were single sunspots whose tilt angle could not be defined. The sample standard deviation of each latitudinal bin is divided by the square root of the bin population number and overplotted as standard error bars.

Empirical Joy's law equations from previous works are also plotted in Figure 1 as described by Wang and Sheeley (1991) as Equation (1), Leighton (1969) as Equation (2), Norton and Gilman (2005) as Equation (3), and Dasi-Espuig *et al.* (2010) as Equation (4):

$$\sin\gamma = 0.48\sin\theta + 0.03 \tag{1}$$

$$\sin\gamma = 0.5\sin\theta \tag{2}$$

$$\gamma = 0.2\theta + 2.0 \tag{3}$$

$$\gamma = (0.26 \pm 0.05)\theta \tag{4}$$

where γ is tilt angle and θ is latitude. Southern Hemisphere latitudes are considered positive for plotting purposes. Tilt angles in both hemispheres are considered positive if the leading spot is closer to the Equator than the following spot.

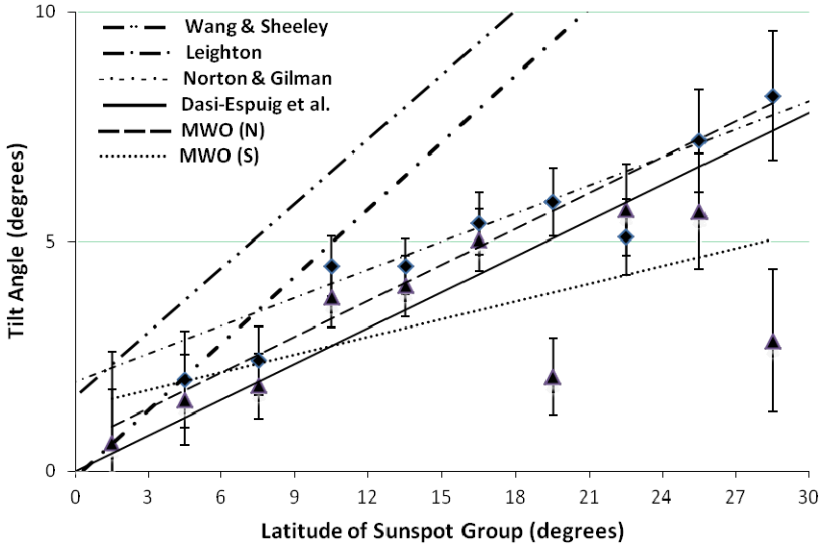


Figure 1 Tilt angle as a function of latitude, Northern (diamond) and Southern (triangle) Hemispheres, for Solar Cycles 16 to 21. Data were binned in 3° of latitude. Standard errors of the mean are overplotted as error bars. Common Joy's law equations are plotted for reference: Wang and Sheeley (1991), Equation (1), dashed-dot-dot; Leighton (1969), Equation (2), large dash-dot; Norton and Gilman (2005), Equation (3), small dash-dot; Dasi-Espuig *et al.* (2010), Equation (4), solid. Linear fit Equation (5) for Northern Hemisphere data (dash) and Equation (6) for Southern Hemisphere data (dot) are also shown.

In order to understand Equations (1) – (4), some background on data and analysis is in order. Equation (1) was formulated by Wang and Sheeley (1991) after an analysis of National Solar Observatory/Kitt Peak data, utilizing 2710 magnetograms of bipolar magnetic regions (BMRs) collected during Solar Cycle 21. Tilt angles were determined by hand, analyzing magnetogram prints at a time of approximate peak flux for each BMR. Averages were flux-weighted and taken over sine-latitude bins of width 0.05 (approximately 3°). Equation (2) was formulated by Leighton (1969), who approximated Joy's law from measurements by Brunner (1930). Norton and Gilman (2005) implemented Joy's law as part of a sunspot-behavior model, and Equation (3) is the best fit to an average of tilt angle as a function of latitude for over 650 active regions observed in *Michelson Doppler Imager* (MDI) data from 1996 – 2004 (Norton and Gilman, 2004). Equation (4) was determined by Dasi-Espuig *et al.* (2010) using available Mt. Wilson data, including the latter part of Solar Cycle 15 to most of Cycle 21. The data were binned by 5° latitude and area-weighted in an effort to reduce scatter, and linear fits were forced through the origin for Equation (4).

We find a linear fit for the relationship of the northern and southern average tilt angles as a function of latitude to be:

$$\gamma_{\mathrm{N}} = 0.26\theta + 0.58 \quad (\text{or } \sin\gamma_{\mathrm{N}} = 0.271\sin\theta + 0.010) \tag{5}$$

$$\gamma_{\mathrm{S}} = 0.13\theta + 1.38 \quad (\text{or } \sin\gamma_{\mathrm{S}} = 0.425\sin\theta + 0.024) \tag{6}$$

The values of the binned, average tilt angles observed at the higher latitudes are not well fit by the Wang and Sheeley (1991) or previous historical Joy's law equations. We propose an updated Joy's law with a lower slope between 0.13 – 0.26, as seen in Equations (5) and (6) for the Northern and the Southern Hemisphere determined using Mt. Wilson data from Cycles 16 to 21. The linear correlation coefficient for Equations (5) and (6) are 0.96 and

0.65 for the Northern and Southern Hemispheres, respectively. The correlation coefficient ($-1 \leq r \leq 1$) measures the strength of the linear relationship between two variables and is defined as the covariance of the two variables divided by the product of their standard deviations. Values of -1 and 1 indicate a perfect inverse or direct relationship, respectively. We correlate the binned, mean tilt angle and latitude.

The hemispheric, linear fits to Joy's law are more consistent with the result of Dasi-Espuig *et al.* (2010), who report a lower slope value of 0.26 – 0.28, than with the equations from the 1990s and prior that had higher slopes. We also propose that the Joy's law equation should not be forced through the origin. It is reasonable that the slope reported by Dasi-Espuig *et al.* (2010) is higher than the slope reported here because they force the fit through the origin, which we do not. We justify our approach as being purely observational. If we did force the fit through the origin, our slopes in Equations (5) and (6) would increase to be 0.29 and 0.20 in the Northern and Southern Hemispheres, respectively. The data here consistently demonstrate that Joy's law does vary by hemisphere. It is possible that the mechanisms responsible for tilt angles in each hemisphere have a canceling effect on tilt near the Equator and therefore are not an accurate indication of Joy's law by hemisphere.

The results of Dasi-Espuig *et al.* (2010), showing that tilt angle is variable as a function of solar cycle, are noteworthy. Our initial attempts to recover Joy's law for each hemisphere and solar cycle were frustrating due to the fact that Joy's law only appears weakly (Figure 2). Cycle 16 had a low population (< 25) in the first ($0-3^\circ$) and last ($27-30^\circ$) bins, resulting in large error bars. These bins were subsequently removed from all individual cycle plots for consistency. The data are poorly fit by a linear function in most cases. The linear correlation coefficients range from $r = 0.18$ (Cycle 17 North, Cycle 19 South) to $r = 0.86$ (Cycle 20 North). The large amount of scatter and high noise apparent in Joy's law is interesting, because it indicates that a stochastic process is competing with the mechanism that determines the tilt angles. The stochastic process dominating Joy's law on the short time scale is considered to be turbulent convection imparting random tilt angles to the rising flux tubes (Fisher, Fan, and Howard, 1995; Weber, Fan, and Miesch, 2012). We agree with Dasi-Espuig *et al.* (2010), who state that "no clear difference could be determined between the slopes of Joy's law from cycle to cycle," as can be seen in Figure 2; therefore, we use the mean tilt value from each hemisphere for each cycle to analyze the hemispheric differences.

It is possible that the recovery of a mean bipolar region tilt angle and scatter for a given solar cycle can be used as a diagnostic for that cycle, *i.e.* the strength of the cycle as indicated by Dasi-Espuig *et al.* (2010) or the geometry/orientation of the toroidal fields from which the flux ropes begin their initial rise (Babcock, 1961; Norton and Gilman, 2004). Simulations by Weber, Fan, and Miesch (2012) of thin flux tubes rising through solar-like turbulent convection show how much the tilt-angle scatter increases with decreasing flux and field strength. Therefore, quantifying the scatter in Joy's law can constrain the flux and field strength within the context of their model; *i.e.* a larger scatter is indicative of flux tubes dominated by convection instead of magnetic buoyancy. In addition, since smaller average tilt angles minimize the amount of active-region flux that becomes the poloidal field, a smaller average tilt angle leads to a weaker polar cap mean field strength (Petrie, 2012).

We are uncertain why specific bins in the Southern Hemisphere showed such different behavior from the other bins. We found that late in all solar cycles (except Solar Cycle 20) aberrant activity occurred at the $18-21^\circ$ latitudes. In particular, the Southern Hemisphere during Cycle 19 is very disorganized, with the high-latitude bins of $18-21^\circ$ (304 regions, 24 %) and $24-27^\circ$ (150 regions, 12 %) having negative mean tilt values, meaning that these bipolar regions have a following spot closer to the Equator than the leading spot. The

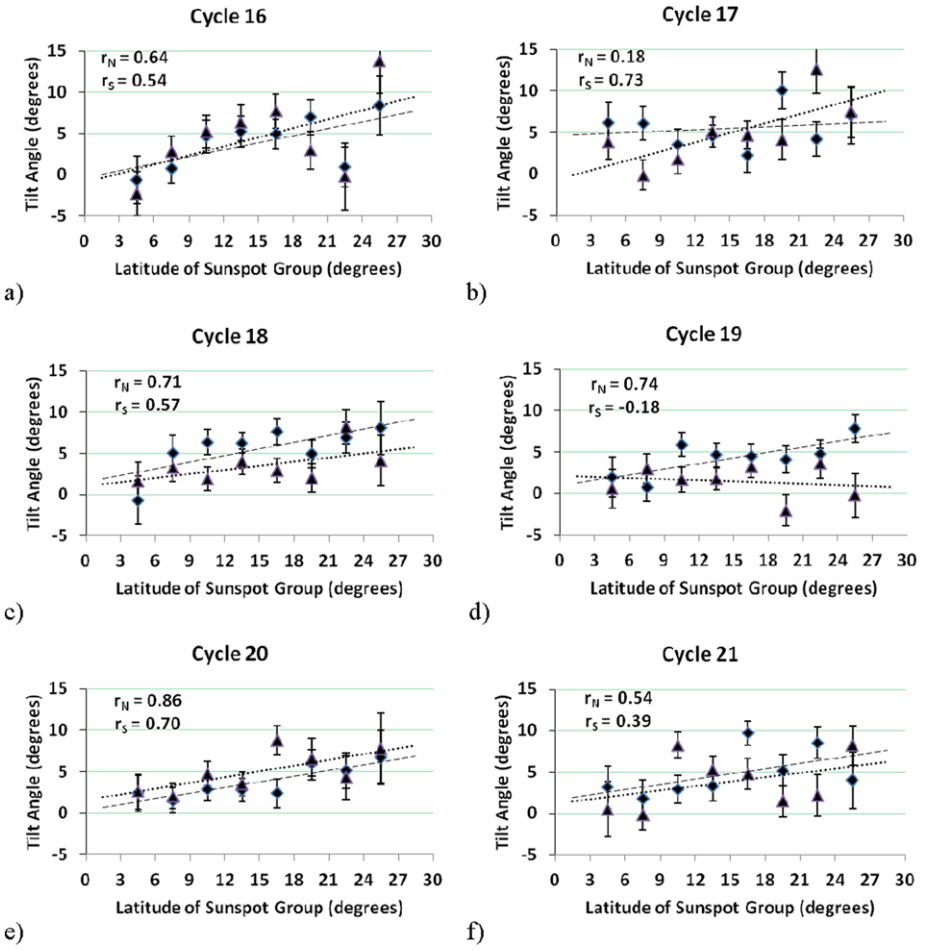


Figure 2 Tilt angle as a function of latitude for the Northern (diamond) and Southern (triangle) Hemispheres for Solar Cycles 16 – 21 are shown in panels (a)–(f), respectively. Data were binned in 3° of latitude. Standard errors of the mean are overplotted as error bars. Linear fits to Northern (dash) and Southern (dot) Hemisphere data are shown with linear correlation coefficients [r_N, r_S] included in the legends.

Southern Hemispheric tilt angles for Cycle 19 are responsible for the low mean tilt angles for the whole Sun in Cycle 19 as reported by Dasi Espuig *et al.* (2010, see their Table 1). It would be of interest to study this in more detail and better understand the conditions favorable for aberrant configurations, *i.e.* anti-Hale and negative tilt angles, to occur.

3. Joy's Law as a Function of Hemisphere

The average tilt angle and standard error of the mean for each hemisphere for Solar Cycles 16 to 21 are given in Table 1. The standard error of the mean was calculated as the sample standard deviation divided by the square root of the sample number. Differences in Joy's law between hemispheres are poorly determined (below the two-σ level) for Cycles 16, 17, 20, and 21. However, Cycles 18 and 19 as well as the data averaged over all

Table 1 Mean tilt angle in degrees for Northern Hemisphere [$\overline{\gamma}_{\rm N}$] and Southern Hemisphere [$\overline{\gamma}_{\rm S}$] with standard deviation of mean [$\sigma_{\overline{\gamma}}$] for Solar Cycles 16 to 21. In addition, difference of hemispheric mean tilt angle and statistical significance are shown in last two columns.

Solar Cycle	$(\overline{\gamma} \pm \sigma_{\overline{\gamma}})_{\rm N}$	$(\overline{\gamma} \pm \sigma_{\overline{\gamma}})_{\rm S}$	$\Delta\overline{\gamma} = \lvert\overline{\gamma}_{\rm N} - \overline{\gamma}_{\rm S}\rvert$	$\frac{\Delta\overline{\gamma}}{\sqrt{\sigma^2_{\overline{\gamma}_{\rm N}} + \sigma^2_{\overline{\gamma}_{\rm S}}}}$
16	$3.8° \pm 0.73$	$4.4° \pm 0.81$	0.6	0.3
17	$5.4° \pm 0.70$	$4.0° \pm 0.71$	1.4	1.4
18	$5.7° \pm 0.61$	$2.9° \pm 0.60$	2.8	3.3
19	$4.6° \pm 0.53$	$1.8° \pm 0.59$	2.8	3.5
20	$3.5° \pm 0.60$	$4.8° \pm 0.66$	1.3	1.5
21	$5.0° \pm 0.67$	$4.4° \pm 0.68$	0.6	0.6
16–21	$4.7° \pm 0.26$	$3.6° \pm 0.27$	1.1	3.0

cycles show a significant difference between hemispheres. These findings are indicative that Joy's law varies by hemisphere and by solar cycle.

Using the results in Table 1, we attempt to answer the following questions: Is there a significant difference between Northern and Southern Hemispheric mean tilt? The last row of Table 1 indicates that, yes, there is a significant difference of mean tilt at a three-σ level. Do the hemispheric differences in mean tilt values change from cycle to cycle? We find an average value of $\Delta\overline{\gamma}$ over all six cycles equal to 1.5 with a statistical significance of nearly four-σ (3.9). Therefore, we are convinced that there is significant variation in hemispheric mean tilts from cycle to cycle.

We agree with Dasi-Espuig *et al.* (2010) that a revision of Joy's law is necessary. Their conclusion that a relationship exists between cycle strength and mean tilt is intriguing, and we attempted to confirm this result. We used the values reported by Goel and Choudhuri (2009) of total sunspot area in microhemispheres by solar cycle and hemisphere [$A_{\rm N}$, $A_{\rm S}$] for Cycles 15 through 21. Cycle 15 data were only available from just prior to solar maximum until the end of the cycle. The minimal effects of data missing from the last nine months of Cycle 21 are discussed in Section 2. Sunspot area is used as a proxy for cycle strength (Solanki and Schmidt, 1993). Areas were calculated from Royal Greenwich Observatory data. We compare total sunspot area to mean tilt separated by hemisphere and solar cycle in Table 2. Assuming that a larger total sunspot area indicates a stronger cycle and that hemispheric differences exist within each cycle, we find evidence of the same inverse relationship as Dasi-Espuig *et al.* (2010) such that a stronger cycle produces less average tilt.

In Figure 3, the area-weighted mean tilt values normalized by mean latitude (see $(\overline{\gamma}/\bar{\lambda})_{\omega}$ in Table 2) are plotted as a function of total sunspot area for Solar Cycles 15 to 21 for the Northern and Southern Hemispheres as well as the total Sun. Dasi-Espuig *et al.* (2010) used area-weighting to give larger and, therefore, less scattered groups more influence on mean tilt. The mean latitude of sunspot emergence decreases and approaches zero as the solar cycle progresses. Normalizing by latitude removes that latitudinal bias and allows for the inclusion of incomplete cycles in our analysis. Linear regression lines are fit to normalized mean tilt and sunspot area for each hemisphere. Correlation coefficients [r] are found to be $r_{\rm N} = -0.29$, $r_{\rm S} = -0.83$, $r_{\rm tot} = -0.75$ for the Northern Hemisphere, Southern Hemisphere, and total-Sun values.

There is an inverse correlation of area-weighted mean tilt to sunspot area and, by proxy, cycle strength in the Southern Hemisphere. The probability is 2.1 % that the linear correla-

Table 2 Sunspot area [$10^4\mu$-hemispheres], mean tilt angle [$\overline{\gamma}$], mean tilt angle normalized by mean latitude $\overline{\gamma}/\bar{\lambda}$, and area-weighted mean tilt angle normalized by mean latitude $(\overline{\gamma}/\bar{\lambda})_\omega$ values are provided for Northern and Southern Hemispheres and total Sun for Solar Cycles 15 – 21. The strength of the correlation of mean tilt with sunspot area was measured as the correlation coefficient [r] for each hemisphere and total-Sun values. Cycle 15 data were only available after solar maximum.

	Solar Cycle							
	15*	16	17	18	19	20	21	r
A_N	4.3	4.7	6.0	7.4	10.6	6.9	7.5	
$\overline{\gamma}_N$	4.1°	3.8°	5.4°	5.7°	4.6°	3.5°	5.0°	0.25
$\overline{\gamma}_N/\bar{\lambda}$	0.35	0.27	0.36	0.37	0.26	0.24	0.33	−0.17
$(\overline{\gamma}_N/\bar{\lambda})_\omega$	0.45	0.29	0.50	0.46	0.30	0.32	0.40	−0.29
A_S	3.6	3.9	6.0	7.0	7.4	4.9	7.8	
$\overline{\gamma}_S$	3.7°	4.4°	4.0°	2.9°	1.8°	4.8°	4.4°	−0.45
$\overline{\gamma}_S/\bar{\lambda}$	0.27	0.32	0.28	0.20	0.12	0.35	0.29	−0.67
$(\overline{\gamma}_S/\bar{\lambda})_\omega$	0.43	0.43	0.27	0.29	0.11	0.33	0.28	−0.83
A_{tot}	7.9	8.6	12.0	14.5	18.0	11.9	15.3	
$\overline{\gamma}_{tot}$	3.9°	4.2°	4.7°	4.3°	3.4°	4.1°	4.7°	−0.16
$\overline{\gamma}_{tot}/\bar{\lambda}$	0.31	0.30	0.32	0.29	0.20	0.29	0.31	−0.64
$(\overline{\gamma}_{tot}/\bar{\lambda})_\omega$	0.44	0.35	0.39	0.38	0.23	0.32	0.34	−0.75

tion coefficient of $r_S = -0.83$ in the South is due to chance. Total-Sun values also suggest an inverse relationship of area-weighted mean tilt angle values with cycle strength, the correlation coefficient $r_{tot} = -0.75$ having a 5.0 % probability of chance. The correlation between mean tilt and cycle strength in the Northern Hemisphere is insignificant. The smallest chance probabilities of 2.1 % and 5.0 % for the Southern Hemisphere and total-Sun correlations are at or below the usual significance level of 5 %, and therefore we confirm a statistically significant negative correlation between area-weighted mean tilt value and cycle strength as measured by sunspot area in the Southern Hemisphere and the whole-Sun data.

4. Joy's Law as a Function of Longitude: Searching for Evidence of a Tipped Toroidal Field in Tilt Angle Data

If toroidal magnetic fields at the base of the convection zone in each hemisphere were tipped with respect to the equatorial plane, as proposed in theory by Cally, Dikpati, and Gilman (2003) and observations (Norton and Gilman, 2005), then flux tubes would begin their rise through the convection zone with a tilt dependent on longitude. This might be observable as a pattern when tilt angles in each hemisphere are studied as a function of longitude. It is well established that active longitudes appear during each solar cycle and certain longitudes host active regions repeatedly over time (De Toma, White, and Harvey, 2000). If an $m = 1$ instability were present, we would expect to see a sinusoidal pattern.

To reveal longitudinal structure, possibly relating to the orientation of the toroidal field in each hemisphere, we separated tilt data by hemisphere and solar cycle. Active-region tilt angles as a function of longitude were plotted for the Northern and Southern Hemispheres for all solar cycles, with data binned into 20° longitudes, then averaged. Plots for Cycles 18 – 20 are presented in Figure 4(a) – (f). We expected an $m = 1$ sinusoidal pattern suggestive of a

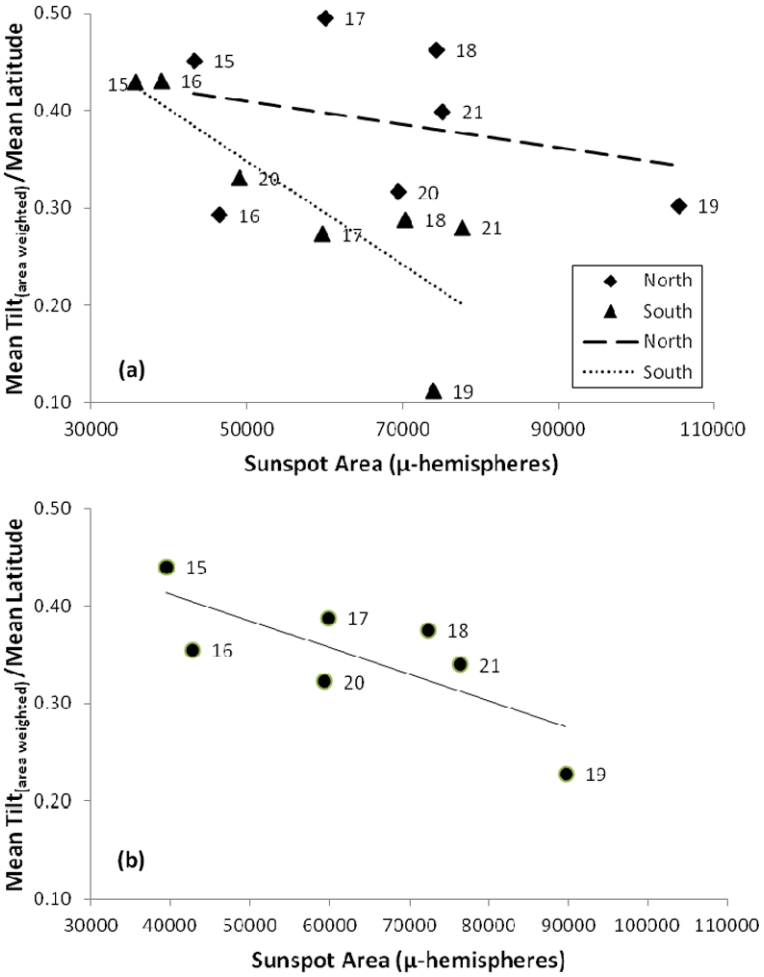


Figure 3 Area-weighted mean tilt angle normalized by mean latitude (degrees) as a function of sunspot area [μ-hemispheres] for Cycles 15 to 21. Panel (a) shows Northern Hemisphere (diamond, dashed line) and Southern Hemisphere (triangle, dotted line), and panel (b) shows the total-Sun sunspot area divided by two. Linear correlation coefficients [r] for each hemisphere and total Sun are $r_{\mathrm{N}} = -0.29$, $r_{\mathrm{S}} = -0.83$, $r_{\mathrm{tot}} = -0.75$.

tipped toroidal field in each hemisphere. We attempted to fit the data with sinusoidal curves representing $m = 1$ through $m = 8$ patterns with various amplitudes. No fit to the data was statistically significant. Therefore, we report no longitudinal dependence in Joy's law.

If a tipped toroidal field were only present for one to two years during a solar cycle, this might prevent a tilt-angle dependence on longitude to be decipherable when averaging over ≈ 11 years. However, it may be possible to see increased scatter in the tilt-angle values for a cycle that has a tipped toroidal field compared to a cycle without one. For this reason, we determined the standard deviation (not the standard deviation of the mean) for the average tilt angle as a function of hemisphere and cycle (see Table 3). The standard-deviation values have a very small range, from 29.3–31.2°, even though the strength of the cycle, shown as sunspot area, varies a great deal. The errors of the standard deviation values shown in

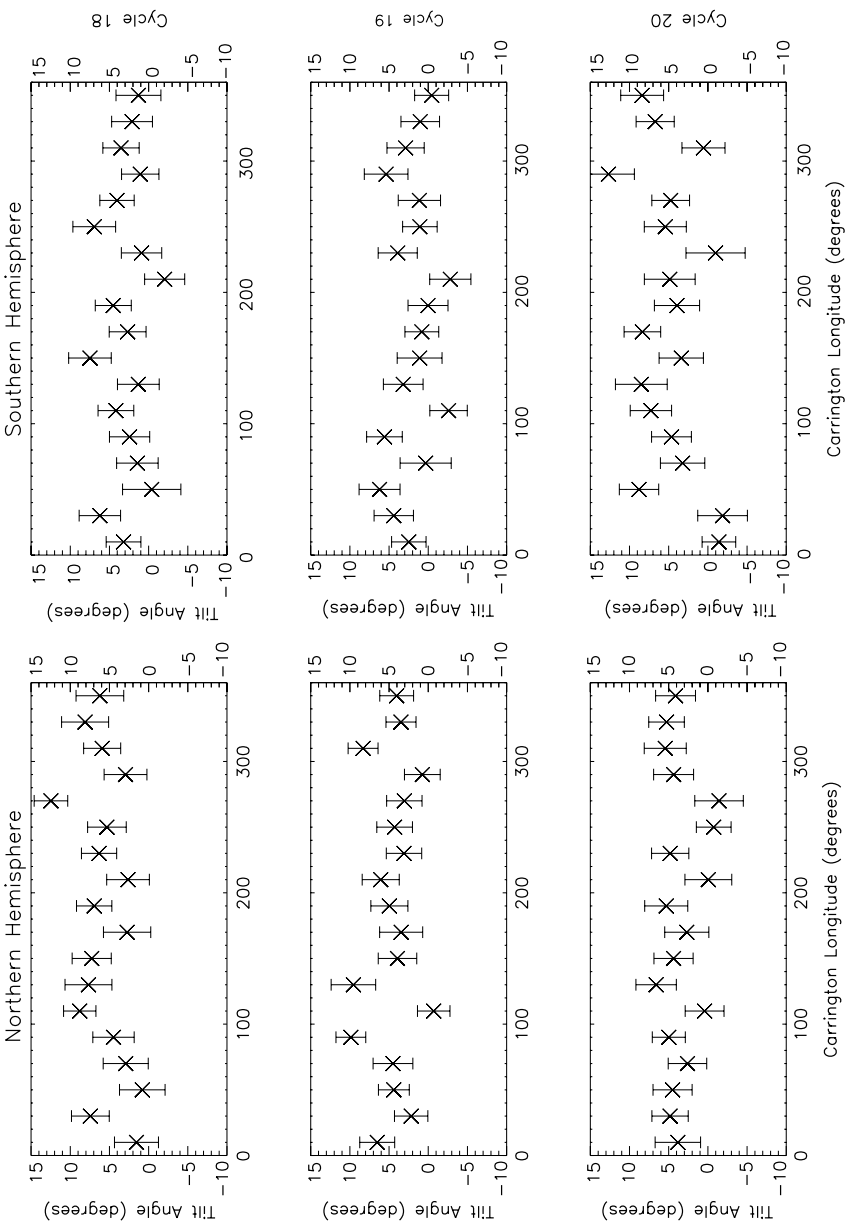


Figure 4 Tilt angle as a function of longitude, Solar Cycles 18 – 20, Northern (a) – (c) and Southern Hemispheres (d) – (f). Data were binned 20° in longitude. Standard deviations of tilt angle for all sunspot groups in each longitudinal bin are overplotted as error bars.

Table 3 Cycle strength in terms of sunspot area [$10^4\mu$-hemispheres] and the standard deviation [σ_γ] (not the standard deviation of the mean) in the mean tilt angle for the bipolar regions of the Northern and Southern Hemispheres for Cycles 16 – 21.

Cycle strength and standard deviation of average tilt angle						
Cycle	16	17	18	19	20	21
A_N	4.7	6.0	7.4	10.6	6.9	7.5
σ_γ	29.6°	30.5°	30.0°	31.2°	29.6°	29.8°
A_S	3.9	6.0	7.0	7.4	4.9	7.8
σ_γ	30.6°	31.0°	30.4°	29.6°	29.3°	29.7°

Table 3 range from 0.53 – 0.81°. The small range of standard deviation values indicates that the source of scatter in the tilt angles is due to a process that is nearly identical from one cycle and hemisphere to the next. The values shown in Table 3 do not support the presence of a tilting or deformation of the toroidal band in the solar interior since there is no difference in scatter of observed tilt angles at the surface between one cycle and the next. We agree with Fisher, Fan, and Howard (1995) that the very small range of the standard deviations of the tilt angle (referred to as rms tilt in their article) are consistent with a process such as the buffeting by convection that is persistent in scale as a function of longitude and latitude and similar from one cycle to the next.

5. Summary

We determined that the mean tilt angle observed in Solar Cycles 16 to 21 was significantly different in the Northern and Southern Hemispheres. Hemispheric differences up to 2.8° in average tilt angle persist across solar cycles. We suggest a revision to Joy's law equations with a weaker dependence on latitude (slopes of 0.26 and 0.13 for the Northern and Southern Hemispheres were found) and more attention paid to the differences between hemispheres and cycles. We did not force the linear fit through the origin as Dasi-Espuig *et al.* (2010) did in their analysis. It is possible that bipolar active regions at the Equator have mean tilt angles of zero because the sampling is an aggregate of flux activity from both hemispheres. If we do force the fits through the origin, we find slopes of 0.29 and 0.20 for the Northern and Southern Hemispheres, compared to 0.26 reported by Dasi-Espuig *et al.* (2010). Weber, Fan, and Miesch (2012) simulate rising flux tubes, including the effect of convection, and produce an expected slope for Joy's law dependent upon the strength of the source toroidal field and total flux in the tube. Our slope values of 0.29 and 0.20 in the Northern and Southern Hemispheres are consistent with field strengths of 15 kG in the interior and flux ropes containing between 10^{20} and 10^{21} Mx.

We confirm the results of Dasi-Espuig *et al.* (2010) that whole-Sun mean tilt angles, weighted by area and normalized by latitude, for Cycles 16 to 21 show a statistically significant negative correlation with cycle strength (see Figure 3b). A tilt-angle dependence upon cycle strength is a feedback mechanism in which the Sun can regulate sunspot-cycle amplitudes; *i.e.* a stronger cycle produces a smaller tilt angle and therefore a weaker poloidal seed field for the $n + 1$ cycle (Cameron and Schüssler, 2012). Jiang *et al.* (2010) study the effect of meridional-flow perturbations and find that larger perturbations reduce the tilt angle of bipolar magnetic regions and thus diminish its contribution to the polar field. The perturbations are caused by near-surface inflows towards the active-region band in each hemisphere, and the perturbation amplitude increases with stronger magnetic cycles. This mechanism may explain the observed anticorrelation between tilt angle and cycle strength. However,

some doubts are cast on the results, because the Northern Hemisphere did not exhibit a statistically significant negative correlation with cycle strength, while the Southern Hemisphere did (see Figure 3a). We hope that this result reinforces the importance of isolating data by hemisphere.

We searched for a non-axisymmetric mechanism at work by analyzing tilt angles as a function of longitude (see Figure 4 for Cycles 18 – 20). We attempted to fit the data with sinusoidal curves representing $m = 1$ through $m = 8$ patterns with various amplitudes. No fit to the data was statistically significant. Therefore, we find no evidence that tilt angles vary regularly in longitude. A toroidal field tipped with respect to the East–West direction would introduce a significant scatter into Joy's law if the flux rope retained some of the original tilt imparted to it from the source toroidal field. Therefore, we calculated the standard deviation of the average tilt angle from each cycle and hemisphere. The values exhibited a narrow range from 29.3 – 31.2° even though the cycle strengths varied greatly (see Table 3). This does not support the presence of a tilting or deformation of the toroidal field but is consistent with a process such as the buffeting by convection that is persistent in scale in latitude and longitude and similar from one cycle to the next.

Moreover, a bias towards reporting positive results regarding Joy's law may have impeded progress on this topic that would benefit from identifying time periods in which Joy's law cannot be recovered. These would be times in which the stochastic processes of turbulent convection dominate the tilt-producing mechanism thought to be the Coriolis force. The work by Weber, Fan, and Miesch (2012) is a great step towards the ability to interpret the scatter of bipolar region tilt angles in any period of the solar cycle to constrain the toroidal field strength in the interior and the flux residing in the thin flux tubes. The standard deviation values of the average tilt angle shown in Table 3 are consistent with Weber, Fan, and Miesch (2012) simulations of flux tubes containing 10^{21} Mx and forming from a toroidal field with a strength of 50 kG.

References

Babcock, H.W.: 1961, *Astrophys. J.* **133**, 572.
Brunner, W.: 1930, *Astron. Mitt. Zurich* **124**, 67.
Cally, P., Dikpati, M., Gilman, P.A.: 2003, *Astrophys. J.* **582**, 1190.
Cameron, R.H., Schüssler, M.: 2012, *Astron. Astrophys.* **548**, A57.
Charbonneau, P.: 2007, *Adv. Space Res.* **40**, 899.
Chatterjee, P., Nandy, D., Choudhuri, A.R.: 2004, *Astron. Astrophys.* **427**, 1019.
Dasi-Espuig, M., Solanki, S.K., Krivova, N.A., Cameron, R.H., Penuela, T.: 2010, *Astron. Astrophys.* **518**, A7.
De Toma, G., White, O.R., Harvey, K.L.: 2000, *Astrophys. J.* **529**, 1101.
Dikpati, M., Gilman, P.A.: 2001, *Astrophys. J.* **559**, 428.
D'Silva, S., Howard, R.F.: 1993, *Solar Phys.* **148**, 1. doi:10.1007/BF00675531.
Durrant, C.J., Wilson, P.R.: 2002, *Solar Phys.* **214**, 23. doi:10.1023/A:1024042918007.
Fisher, G.H., Fan, Y., Howard, R.F.: 1995, *Astrophys. J.* **438**, 463.
Goel, A., Choudhuri, A.R.: 2009, *Res. Astron. Astrophys.* **9**, 115.
Hale, G.E., Ellerman, F., Nicholson, S.B., Joy, A.H.: 1919, *Astrophys. J.* **49**, 153.
Howard, R.F.: 1991, *Solar Phys.* **136**, 251. doi:10.1007/BF00146534.
Howard, R.F.: 1996, *Annu. Rev. Astron. Astrophys.* **34**, 75.
Howard, R.F.: 2000, *Astron. Astrophys.* **21**, 119.
Jiang, J., Isik, E., Cameron, R.H., Schmitt, D., Schüssler, M.: 2010, *Astrophys. J.* **717**, 597.
Kosovichev, A.G., Stenflo, J.O.: 2008, *Astrophys. J. Lett.* **688**, L115.
Leighton, R.B.: 1964, *Astrophys. J.* **140**, 1547.
Leighton, R.B.: 1969, *Astrophys. J.* **156**, 1.
McIntosh, S.W., Leamon, R.J., Gurman, J.R., Olive, J.P., Cirtain, J.W., Hathaway, D.H., Burkepile, J., Miesch, M., Markel, R.S., Sitongia, L.: 2013, *Astrophys. J.* **765**, 146.

Norton, A.A., Gallagher, J.C.: 2010, *Solar Phys.* **261**, 193. doi:10.1009/s11207-009-9479-6.
Norton, A.A., Gilman, P.A.: 2004, *Astrophys. J.* **603**, 348.
Norton, A.A., Gilman, P.A.: 2005, *Astrophys. J.* **630**, 1194.
Petrie, G.J.D.: 2012, *Solar Phys.* **281**, 577. doi:10.1007/s11207-012-0117-3.
Schüssler, M., Caligari, P., Ferriz-Mas, A., Moreno-Insertis, F.: 1994, *Astron. Astrophys.* **281**, L69.
Sivaraman, K.R., Gupta, S.S., Howard, R.F.: 1993, *Solar Phys.* **146**, 27. doi:10.1007/BF00662168.
Sivaraman, K.R., Gupta, S.S., Howard, R.F.: 1999, *Solar Phys.* **189**, 69. doi:10.1023/A:1005277515551.
Solanki, S.K., Schmidt, H.U.: 1993, *Astron. Astrophys.* **267**, 287.
Stenflo, J.O., Kosovichev, A.G.: 2012, *Astrophys. J.* **745**, 129.
Temmer, M., Rybák, J., Bendík, P., Veronig, A., Vogler, F., Otruba, W., Pötzi, W., Hanslmeier, A.: 2006, *Astron. Astrophys.* **447**, 735.
Wang, Y.-M., Sheeley, N.R.: 1991, *Astrophys. J.* **375**, 761.
Weber, M.A., Fan, Y., Miesch, M.S.: 2012, *Solar Phys.* doi:10.1007/s11207-012-0093-7.

DOI 10.1007/978-1-4899-8005-2_15
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-013-0321-9

A Search for Helioseismic Signature of Emerging Active Regions

S. Kholikov

Received: 23 April 2012 / Accepted: 7 May 2013 / Published online: 5 June 2013

Abstract Using SOHO/MDI and GONG observations we present time–distance deep-focusing measurements to examine the deeper layers of the solar convective zone. The constructed travel-time maps show 10 – 15 second perturbations at depths of 40 – 75 Mm around active region locations before their emergence to the solar surface. The majority of the active regions used in this study were the same as those used in the recent work published by Ilonidis, Zhao, and Kosovichev (*Science* **333**, 993, 2011). In order to confirm the capability of time–distance measurements to detect emerging active regions, we used a technique similar to their time–distance scheme. Our measurements only in some cases show a similar travel-time anomaly. Additionally, we have shown that the technique utilized in our study can provide more spatial details of the emerging flux configurations.

Keywords Time–distance helioseismology · Active regions

1. Introduction

Solar surface activity is a product of magnetic flux emerged from the Sun's interior. Emerging strong magnetic fields form active regions of various configurations. The physical properties as well as the evolution of some aspects of active regions have been well studied by various observational and theoretical techniques (Fan, 2008; Birch *et al.*, 2013). However, the detection of magnetic flux prior to its emergence to the surface remains as one of the most important subjects in this field. Specifically, space weather and solar cycle predictions depend on the outcome of this investigation. Local helioseismology based on time–distance analysis introduced by Duvall *et al.* (1993) became one of the powerful tools for probing solar subsurface properties. The first attempt to detect acoustic wave perturbations resulting from the presence of strong magnetic flux deep in the solar convection zone was carried out

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

S. Kholikov (✉)
National Solar Observatory, Tucson, AZ 85719, USA
e-mail: kholikov@noao.edu

by Chang, Chou, and Sun (1999). Using TON and MDI data they found that an emerging active region (NOAA 7978) generates phase changes of acoustic waves at depths of up to 30 Mm before appearing on the solar surface. Another analysis, reported by Kosovichev, Duvall, and Scherrer (2000), indicated that emerging active regions introduce a complex pattern of perturbations at depths of approximately 18 Mm. They also estimated the speed of emergence to be 1.3 $\mathrm{km\,s^{-1}}$. Relatively similar estimates of emerging subsurface flux were reported by Zharkov and Thompson (2008). Komm, Howe, and Hill (2012) used a ring-diagram analysis (Hill, 1988) to find evidence of emerging magnetic flux in vorticity measurements in subsurface layers of the convective zone.

Ilonidis, Zhao, and Kosovichev (2011) applied a deep-focusing time–distance technique to detect travel-time perturbations of emerging magnetic flux. They found large travel-time shifts at depths of 40 – 75 Mm. The magnitude of travel-time perturbations found in this study were very large, 10 – 15 seconds. However, Braun (2012) reported that acoustic holography technique measurements of the travel times of the same active regions did not show any travel-time anomalies. According to theoretical simulations, expected travel-time shifts due to the mass-flow caused by emerging Ω-loops at the deep layers should not exceed a few seconds in magnitude (Fan, 2008). Modeling buoyant magnetic flux, Birch, Braun, and Fan (2010) also suggested that the travel-time perturbations are approximately one second for flow speeds of 40 $\mathrm{m\,s^{-1}}$ at depths of 30 Mm. Since the sensitivity of travel times at lower turning points is higher for horizontal flows, the large time-shifts observed by Ilonidis, Zhao, and Kosovichev (2011) are not related to strong magnetic field changes. Moreover, magnetic field perturbations alone may not be responsible for such large travel-time perturbations. According to Chou and Serebryanskiy (2002), an estimated travel-time shift of about 0.015 second at the base of the convective zone ($\approx$ 200 Mm below the surface) is equivalent to a magnetic field strength of $(4-7) \times 10^5$ G.

This study is focused on reproducing the results of Ilonidis, Zhao, and Kosovichev (2011) by using an independently developed technique and additional measurements from the GONG helioseismology network. In addition, we followed the Ilonidis, Zhao, and Kosovichev (2011) analysis technique and used several active regions including two of the regions described in their paper to compare our results.

2. Data and Analysis Technique

We used MDI and GONG Doppler velocity images for three active regions (AR): NOAA 10488, 8164 and 10132. They were observed in different time periods between 1996 and 2003. For each AR, we analyzed data for the period of 2 – 3 days preceding their emergence. Although the processing steps we used differ from those used by Ilonidis, Zhao, and Kosovichev (2011) in detail, the basic outcome remained the same. Phase velocity filtering is one of the most important parts of many time–distance measurements (Duvall *et al.*, 1997).

In the standard time–distance analysis, the data cube is filtered by transforming it into Fourier domain, applying a filter to select acoustic waves with certain phase-speed parameters (phase speed and FWHM), and transforming back to the time–space domain. The size of data cubes is usually 30 × 30 degrees in latitude and longitude with durations of nearly 8 hours. The phase-speed filtering with specific parameters isolates acoustic waves with approximately the same raypath and lower turning point at a selected depth range. In order to increase the signal-to-noise ratio, Ilonidis, Zhao, and Kosovichev (2011) used a modified phase-speed filter, replacing the Gaussian shape filter by a Π-shape filter, which is flat at the

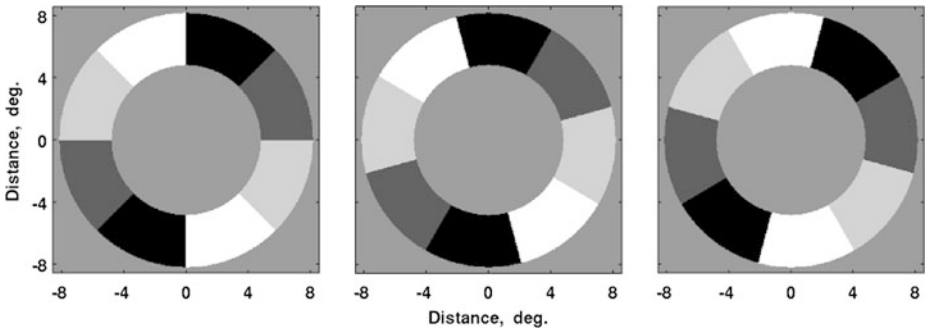


Figure 1 Configurations of the 45° arcs for the deep focusing cross-correlation measurements. Each arc is divided into 22 separation distances relative to the central point. The cross correlations of solar oscillation signals are computed between averaged diametrically opposite sub-arc locations.

center and Gaussian at the edges; in our study the same phase-speed filter was employed. The flat part of the filter corresponds to the phase velocity range of $21-29\,\mathrm{km\,s^{-1}}$. The main steps of the filtering procedure are as follows:

- Decompose individual Doppler velocity images into spherical harmonics (SH) coefficients using the central 120×120 degree in latitude and longitude part of the disk.
- Apply the phase-speed filter to all individual (l, m) SH coefficients of the observed time series.
- Reconstruct the oscillation velocities using the filtered coefficients; they can be reconstructed using the relation

$$V(\theta,\phi)=\sum_{l=0}^{l_{\max}}\sum_{m=0}^{l}C_l^m P_l^m(\theta)\mathrm{e}^{\mathrm{i}m\phi+\delta\phi}. \tag{1}$$

  Here C_l^m are SH coefficients, P_l^m is the associated Legendre polynomial of degree l and order m, θ and ϕ are latitude and longitude, respectively. $\delta\phi$ is a differential rotation profile (Libbrecht and Morrow, 1991), function of latitude and time (relative to the middle of the data cube) to remove the surface differential rotation. The range of l for the chosen phase-speed parameters is $l=70-200$, implying that it is not necessary to use very high degrees of SH coefficients, therefore $l_{\max}=300$ covers all our needs.
- Tracking of the reconstructed images is done by introducing the surface differential rotation profile angular velocity into the argument of the complex exponential in the right-hand side of Equation (1).

Deep focusing cross-correlation calculations of these reconstructed data are produced using the scheme discussed in Ilonidis, Zhao, and Kosovichev (2011). The maximum peak of the cross-correlation function for the phase speeds ranging in $21-29\,\mathrm{km\,s^{-1}}$ corresponds to separation distance $\Delta=12-13°$. For any given central target region position, the annulus diameters used in these calculations are $D1=9.6°$, $D2=16.32°$. The annulus is divided into an even number of arcs. For this analysis we use the arc sizes of 60°, 45°, 36°, 30°, 25.71°, 22.5°, 20°. In each case, we have an even number of arcs to cross correlate the signal between the diametrically opposite pairs. Each arc configuration is rotated left and right by 1/3 of its length which makes 21 independent configurations. Figure 1 demonstrates a diagram of three configurations of the 45° arcs. The annulus width is divided into 22 equidistant sub-annuli relative to the central position. The cross correlations are computed between

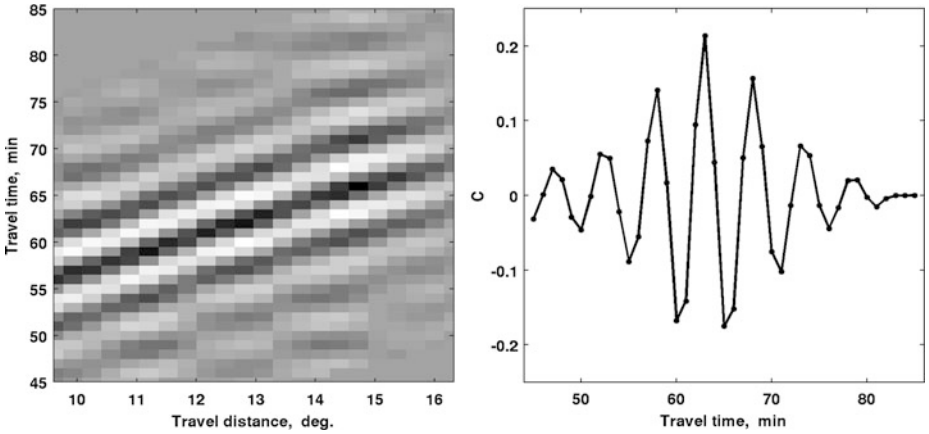


Figure 2 An example of cross correlation of solar oscillations between two particular pixels as a function of separation distance (left). An average cross-correlation function after time-shifted to the central travel distance location (right). The resulting cross correlations are fitted by Gabor wavelet to obtain travel times.

the opposite sub-annuli for each arc configuration. The individual cross correlations are difficult to fit with the Gabor wavelet, therefore, instead of fitting every cross correlation, we combined the cross correlations for the different travel distances. Before averaging these 22 cross correlations, we shifted them into one of the central travel distance locations using the mean time–distance relation which is obtained by averaging measurements over the entire target region. This operation can be done by interpolation. To avoid interpolation artifacts, we used the Fourier shift theorem:

$$\psi(t-\tau_\Delta) = \mathcal{F}^{-1}\left\{e^{-i\omega\tau_\Delta}\Psi(\omega)\right\}, \tag{2}$$

where ψ and Ψ are the cross-correlation function and its Fourier transform, respectively, and τ_Δ is the corresponding time shift. Taking an inverse Fourier transform shifts the cross correlations in the positive or negative time direction relative to the central travel time. Time shifts (τ_Δ) relative to the middle of the 22 travel distances are obtained from the average time–distance relation, and cross correlations are shifted to the central travel-time location. An example of this operation is presented in Figure 2.

Once cross correlations are computed and averaged over all travel distances, travel-time maps for selected target regions are constructed by the standard Gabor wavelet fitting procedure (Duvall *et al.*, 1997). In order to cover a typical active region in size, $13^\circ \times 13^\circ$, travel-time maps centered at AR locations where emergence was expected were produced. Measurements were carried out for the 2 – 3 days time periods that preceded the emergence of the AR to the solar surface.

3. Results and Discussion

The main purpose of this work was to measure travel-time perturbations at depths of 40 – 75 Mm described in Ilonidis, Zhao, and Kosovichev (2011) for detection of emerging active regions. Two active regions presented in their study were also analyzed. AR 10488 is recognized as one of the fastest emerging flux regions in the descending phase of the previous solar activity cycle. Initial evidence of the emergence of this AR was observed on the solar

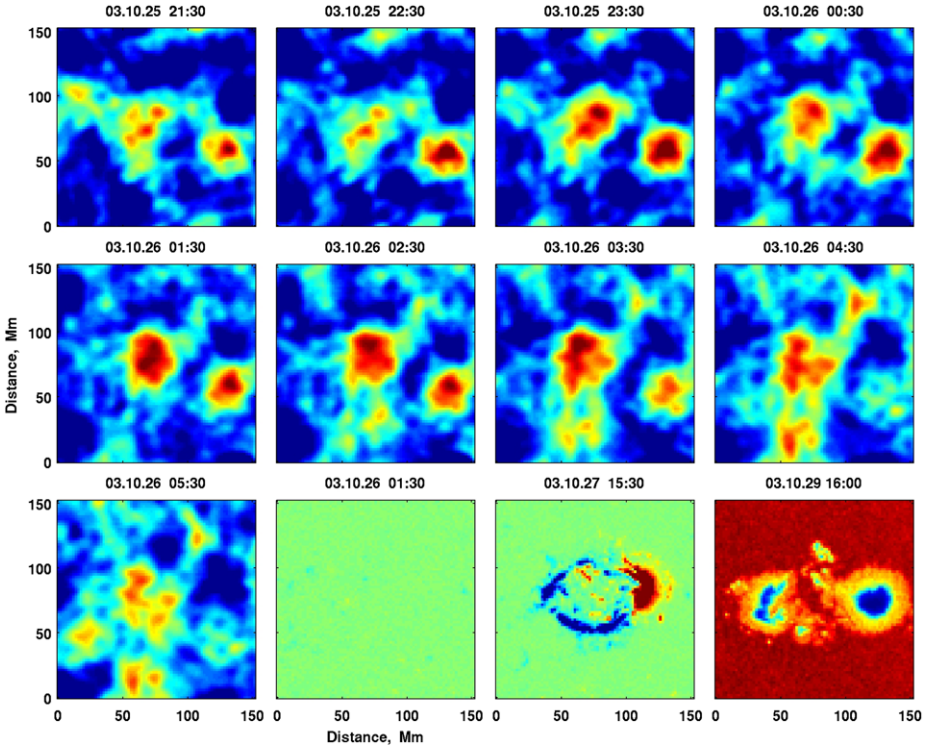


Figure 3 Deep focusing travel-time maps of the Active Region 10488 (top two rows and left image in bottom row). The focus depth for these measurements covers 40 – 70 Mm. Each map is created using 8-hour duration MDI data cubes centered at the emerging active region location (dates and times are displayed at the top). The color scale in the travel-time maps covers the perturbation range 0 (blue) to −15 (red) seconds. The maximum time shift magnitude is observed at around 1:30 UT on 26 October. The magnetic fields before and after the emergence time and a continuum image for a later time are presented in the last three panels of the bottom row. The active region location shows significant double structure travel-time perturbation on below the surface prior the emergence on 25 – 26 October, and after emergence on the surface on 29 October.

surface at 9:30 UT on 26 October 2003. Since the location was not far from the center of the solar disk, analysis can be done without distortion and projection artifacts. Starting from 21:30 UT on 25 October, travel-time maps show a clear decrease in two locations separated by about 40 Mm from each other. Nine consecutive travel-time maps prior to the emergence time are presented in Figure 3; each map is produced using an 8-hour data cube at the times shown on each panel. The travel-time perturbations on these maps range from 0 (blue) to −15 (red) seconds. Despite a few minor features, the strongest perturbations occur at two locations and remain for 7 hours. It should also be mentioned that the first map is 11 hours away from the time when the AR started appearing on the surface. MDI magnetogram and continuum images of corresponding AR locations are shown in the last three panels of the bottom row. There is no significant surface magnetic field during the time period when the travel times are computed. In the magnetogram of 15:30 UT on 27 October, the AR is clearly seen at the surface and starts to evolve. The continuum image taken two days later displays two separate components of the AR. In the quiet areas of the maps travel times are about −(5 – 7) second. Another active region analyzed in Ilonidis, Zhao, and Kosovichev (2011) was AR 8164, which was smaller in size and less active. It emerged to the solar surface at

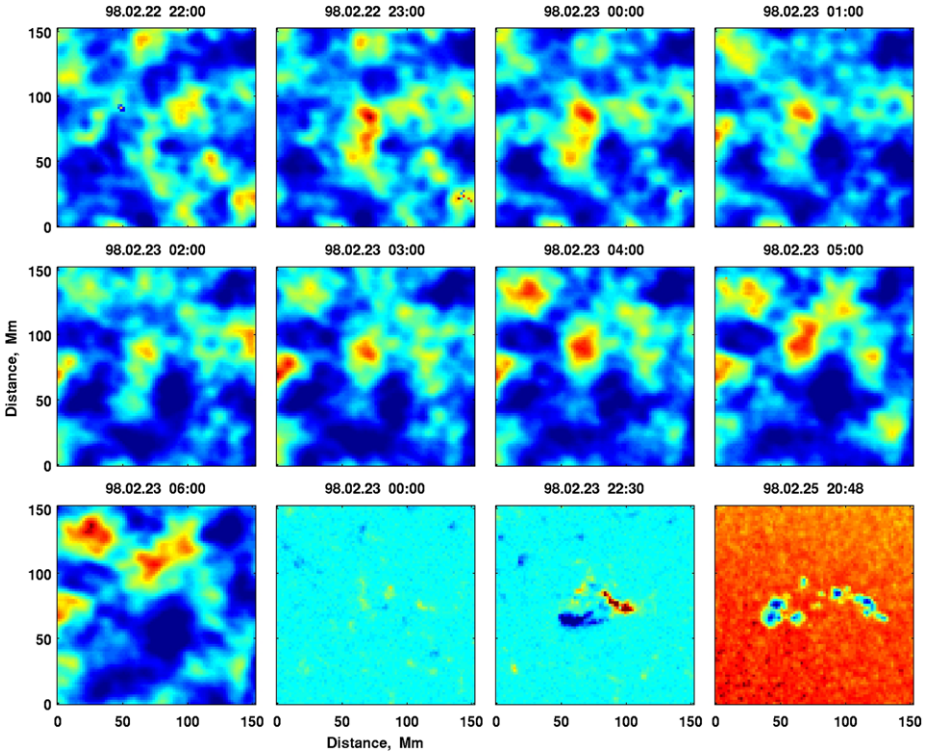


Figure 4 Travel time maps of the Active Region 8164 obtained from MDI SH time series. The color scale in travel-time maps covers the range from 3 (blue) and −15 (red) seconds (dates and times are displayed at the top). The travel-time anomaly is not as strong as in the case of AR 10488. There is a significant travel-time anomaly around the central part of the maps before the AR emerged to the surface. This target region had some surface magnetic activity before emergence (second panel in bottom row). Some of the arc configuration endpoints are possibly affected by this activity which increases additional time shifts around the target region. The continuum picture (last panel) shows a complex configuration of sunspots. Travel time perturbations are most likely very sensitive to the surface activity and possibly to the magnetic field configuration.

about 5:00 UT on 23 February 1998. Figure 4 presents the travel-time perturbation maps of this active region. Unlike AR 10488, the location of this region had some magnetic field before its emergence time. A 10 – 15 second shift is visible in the travel-time maps during the last few hours prior the emergence and at start of the emergence. However, there are several locations where the time shifts take place. The continuum image of the target region two days later shows a complex pattern of sunspot configuration. Perhaps, parts of the arc configurations are affected by the presence of magnetic elements on the surface, which may lead to additional time shifts in the travel-time maps.

Similar complicated patterns were found in the travel-time map of AR 10132. In Figure 5 travel-time perturbation maps of this active region are shown. Both GONG and MDI data showed identical features in size and location, suggesting that time shifts for emergence and other locations have a solar origin. It should also be mentioned that from the 12 selected active regions, only three showed detectable time shifts before emerging. Most of the ARs did not show significant time shifts. ARs which did not show significant travel-time perturbations are listed in Table 1.

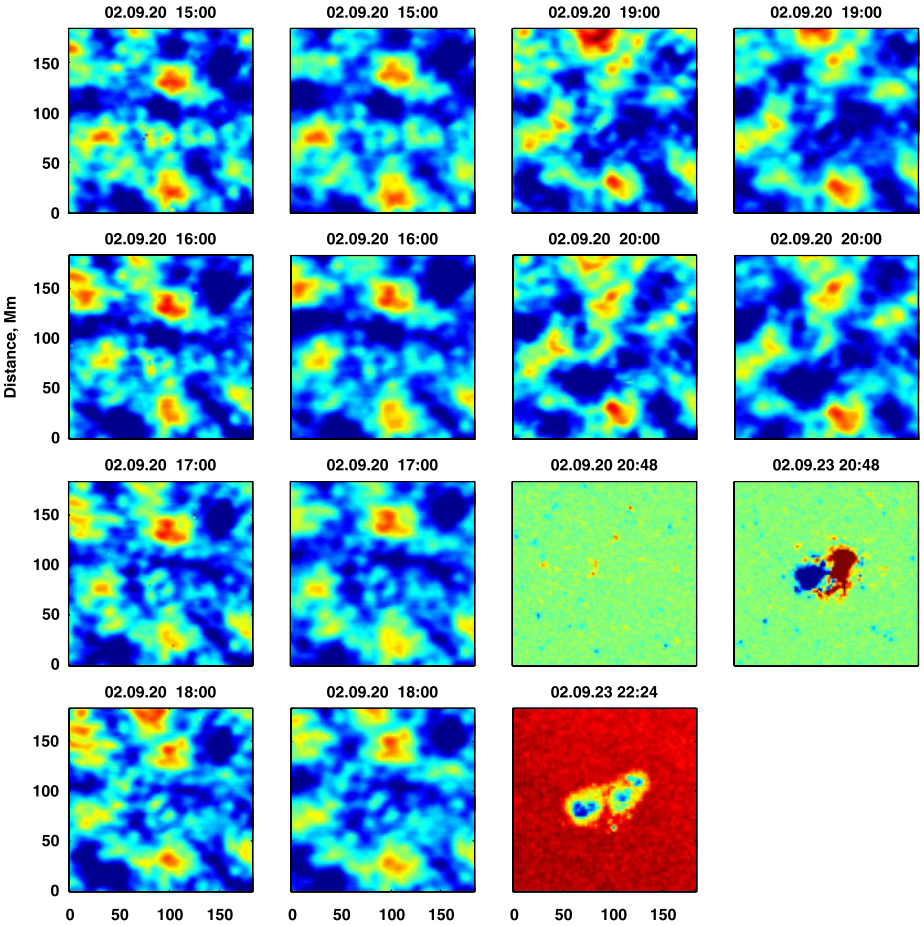


Figure 5 Travel time maps of the Active Region 10132. The color scale in this maps covers the range from 5 (blue) and −13 (red) second (date and time of each map are displayed on the top of each map). Two cotemporal columns represent maps obtained from MDI SH time series (left) and from GONG velocity images (right). This AR emerged to the surface in 21 September 2002, at about 17:00. In the last pair of panels corresponding MDI magnetograms before and after emergence are shown (±400 G). Presence of some magnetic elements around the emergence region can be seen, which may be responsible to part of travel-time perturbations on the maps. The last bottom panel shows an intensity power map two hours later showing three separate sunspots in the active region. Travel time perturbations from the two independent instruments are very similar.

4. Conclusions

Our study shows that the deep focusing time–distance helioseismology measurements indeed provide evidence of some emerging ARs, with time shift magnitudes significantly larger than estimated predictions from simulations (Fan, 2008).

We were able to reproduce the travel-time maps for emerging active regions, similar to Ilonidis, Zhao, and Kosovichev (2011), with one significant difference in our analysis technique: we used solar oscillation signals from a much larger fraction of the solar disk to generate spherical harmonic coefficient time series for selecting acoustic ray paths focused

Table 1 Analyzed active regions which did not show travel-time anomalies at emerged locations prior to their appearance on the Sun's surface. Heliographic coordinates correspond to the center of the tracked regions.

NOAA	Date	Latitude	Longitude
08171	1998.02.27	−25	4
08226	1998.05.24	15	5
08404	1998.12.07	−23	−22
08857	2000.02.06	6	−9
09058	2000.06.23	−14	5
09165	2000.09.13	13	−40
09396	2001.03.25	−6	12
09611	2001.09.07	9	3
10314	2003.03.15	−15	−22

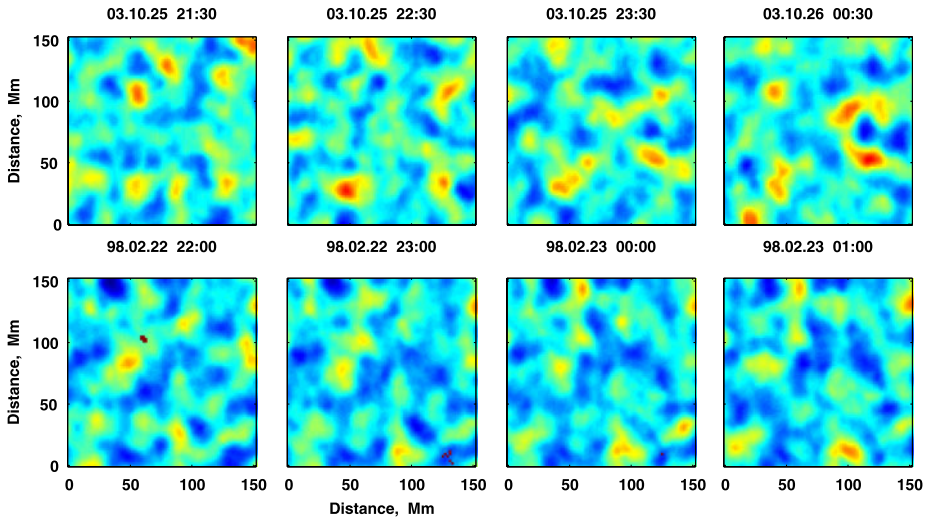


Figure 6 Standard error maps corresponding to the Active Regions 10488 (top row) and 8164 (bottom row). Only first four consecutive maps are provided. The color scale in these maps covers the range 1 (blue) to 3.6 (red) second. Errors in our measurements are obtained from averaging 21 arc configuration measurements. Arc configurations are not completely independent, because each original 7 arc configurations were rotated left and right by 1/3 of arc's length to produce another 14 configurations. Due to this procedure, the magnitude of errors may be underestimated by a factor of 1.7.

at depths of 40 – 75 Mm. We believe that the this approach enables us to resolve a more detailed structure of the emerging flux in the travel-time maps. The similarity of our findings with those reported by Ilonidis, Zhao, and Kosovichev (2011) despite the use of different data-processing steps, leads to an independent confirmation (only in three cases from the analyzed 12 ARs) of their study. In this point we cannot explain why only these particular ARs show travel-time anomalies, while all others do not show. Error-bar maps in Figure 6 show specific spatial pattern, which may be related to Π-shape of the used phase-velocity filter. A more detailed study of the non-standard phase-speed filtering measurements has to be done before making any further interpretations of the time shifts. Our future goal is to apply this technique to a larger number of active regions and try to design different phase-speed filters in order to increase successful detections.

Acknowledgements This work utilized data obtained by the Global Oscillation Network Group (GONG) program, managed by the National Solar Observatory, which is operated by AURA, Inc. under a cooperative agreement with the National Science Foundation. The data were acquired by instruments operated by the Big Bear Solar Observatory, High Altitude Observatory, Learmonth Solar Observatory, Udaipur Solar Observatory, Instituto de Astrofísica de Canarias, and Cerro Tololo Interamerican Observatory. SOHO is a mission of international cooperation between NASA and ESA. Many thanks to A. Kosovichev for suggestions and useful comments which helped to improve the readability of the manuscript.

References

Birch, A.C., Braun, D.C., Fan, Y.: 2010, *Astrophys. J. Lett.* **723**, L190.
Birch, A.C., Braun, D.C., Leka, K.D., Barnes, G., Javornik, B.: 2013, *Astrophys. J.* **762**, 12.
Braun, D.C.: 2012, *Science* **336**, 296.
Chang, H.-K., Chou, D.-Y., Sun, M.-T.: 1999, *Astrophys. J. Lett.* **526**, L53.
Chou, D.-Y., Serebryanskiy, A.: 2002, *Astrophys. J. Lett.* **578**, L157.
Duvall, T.L. Jr., Jefferies, S.M., Harvey, J.W., Pomerantz, M.A.: 1993, *Nature* **362**, 430.
Duvall, T.L., Kosovichev, A.G., Scherrer, P.H., Bogart, R.S., Bush, R.I., Dec Forest, C., *et al.*: 1997, *Solar Phys.* **170**, 63.
Fan, Y.: 2008, *Astrophys. J.* **676**, 680.
Hill, F.: 1988, *Astrophys. J.* **333**, 996.
Ilonidis, S., Zhao, J., Kosovichev, A.: 2011, *Science* **333**, 993.
Komm, R., Howe, R., Hill, F.: 2012, *Solar Phys.* **277**, 205.
Kosovichev, A., Duvall, T.L. Jr., Scherrer, P.H.: 2000, *Solar Phys.* **192**, 159.
Libbrecht, K.G., Morrow, C.A.: 1991, In: Cox, A.N., *et al.* (eds.) *Solar Interior and Atmosphere*, Univ. Arizona Press, Tucson, 479.
Zharkov, S., Thompson, M.J.: 2008, *Solar Phys.* **251**, 369.

DOI 10.1007/978-1-4899-8005-2_16
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0093-7

Comparing Simulations of Rising Flux Tubes Through the Solar Convection Zone with Observations of Solar Active Regions: Constraining the Dynamo Field Strength

M.A. Weber · Y. Fan · M.S. Miesch

Received: 15 February 2012 / Accepted: 27 July 2012 / Published online: 7 September 2012

Abstract We study how active-region-scale flux tubes rise buoyantly from the base of the convection zone to near the solar surface by embedding a thin flux tube model in a rotating spherical shell of solar-like turbulent convection. These toroidal flux tubes that we simulate range in magnetic field strength from 15 kG to 100 kG at initial latitudes of 1° to 40° in both hemispheres. This article expands upon Weber, Fan, and Miesch (*Astrophys. J.* **741**, 11, 2011) (Article 1) with the inclusion of tubes with magnetic flux of 10^{20} Mx and 10^{21} Mx, and more simulations of the previously investigated case of 10^{22} Mx, sampling more convective flows than the previous article, greatly improving statistics. Observed properties of active regions are compared to properties of the simulated emerging flux tubes, including: the tilt of active regions in accordance with Joy's Law as in Article 1, and in addition the scatter of tilt angles about the Joy's Law trend, the most commonly occurring tilt angle, the rotation rate of the emerging loops with respect to the surrounding plasma, and the nature of the magnetic field at the flux tube apex. We discuss how these diagnostic properties constrain the initial field strength of the active-region flux tubes at the bottom of the solar convection zone, and suggest that flux tubes of initial magnetic field strengths of ≥ 40 kG are good candidates for the progenitors of large (10^{21} Mx to 10^{22} Mx) solar active regions, which agrees with the results from Article 1 for flux tubes of 10^{22} Mx. With the addition of more magnetic flux values and more simulations, we find that for all magnetic field strengths, the emerging tubes show a positive Joy's Law trend, and that this trend does not show a statistically significant dependence on the magnetic flux.

Keywords Interior, convection zone · Magnetic fields, models

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

M.A. Weber · Y. Fan · M.S. Miesch
High Altitude Observatory, National Center for Atmospheric Research, Boulder, CO, USA

M.A. Weber
e-mail: mariaw@ucar.edu

M.A. Weber (✉)
Colorado State University, Fort Collins, CO, USA
e-mail: mariaw@lamar.colostate.edu

1. Introduction

The toroidal magnetic field responsible for the emergence of solar active regions is believed to be generated by a dynamo mechanism at or near the base of the convection zone (*e.g.* Spiegel and Weiss, 1980; van Ballegooijen, 1982; Moreno-Insertis, Schüssler, and Ferriz-Mas, 1992; Gilman, 2000; Charbonneau, 2010). Of essential importance in the understanding of solar dynamo theory, and indeed all of solar physics, since flux emergence also regulates solar variability, is addressing how active-region flux tubes rise dynamically through a turbulent convection zone. Also of interest is identifying the dynamo-generated magnetic field strength at the base of the convection zone required to produce the observed properties of solar active regions.

Valuable insights into the nature and evolution of rising magnetic loops in the solar convective envelope have been gained through the use of the thin flux tube approximation by a plethora of authors (*e.g.* Spruit, 1981; Moreno-Insertis, 1986; Ferriz-Mas and Schüssler, 1993; Longcope and Klapper, 1997; Fan, 2009). It is found that in order for these simulated flux tubes to exhibit tilt angles and latitudes of emergence that agree well with observed solar active regions, the toroidal magnetic field at the base of the convection zone needs to be in the range of ≈ 30 kG to ≈ 100 kG (Choudhuri and Gilman, 1987; D'Silva and Choudhuri, 1993; Schüssler *et al.*, 1994; Caligari, Moreno-Insertis, and Schüssler, 1995). These studies also indicate that the Coriolis force is responsible for several observed asymmetries that exist between the leading (in the direction of solar rotation) and following polarities of solar active regions, such as: the tilt angles in accordance with Joy's Law (D'Silva and Choudhuri, 1993; Caligari, Moreno-Insertis, and Schüssler, 1995), the apparent faster proper motion of the leading polarity of an emerging active region on the solar surface (van Driel-Gesztelyi and Petrovay, 1990; Moreno-Insertis, Caligari, and Schüssler, 1994; Caligari, Moreno-Insertis, and Schüssler, 1995), and an asymmetry where the leading polarity shows a more coherent morphology (Fan, Fisher, and DeLuca, 1993; Caligari, Moreno-Insertis, and Schüssler, 1995; Caligari, Schüssler, and Moreno-Insertis, 1998). However, as these previous studies neglect the effects of turbulent convection on rising flux tubes, it is possible that convective downdrafts can pin portions of the flux tube down to the base of the convection zone, while helical upflows between the downdrafts can boost the rise of the flux tube (Fan, Abbett, and Fisher, 2003; Weber, Fan, and Miesch, 2011). Therefore, resulting emerging loops may exhibit different properties from those produced in the absence of convection and pose an intriguing topic of study.

As of yet, the magnetic field strength at which the solar dynamo operates is not well known, nor is it directly accessible via observations. However, solar cycle dynamo models that incorporate the Lorentz force from large scale mean fields indicate that the magnetic field strength generated and amplified at the base of the convection zone is ≈ 15 kG, and most likely cannot exceed 30 kG (Rempel, 2006a, 2006b). Recent simulations of solar-like stars that rotate three times the current solar rate have shown that a rotating convective envelope can generate a dynamo that consists of opposite polarity magnetic wreaths in two hemispheres, which span the depth of the convection zone (Brown *et al.*, 2010). When portions of these wreaths become strong enough, ≈ 35 kG or greater, a buoyant magnetic loop develops, which then rises through the convecting fluid in which it is embedded (Nelson *et al.*, 2011). While these dynamo-producing convection simulations are not meant to reproduce the solar dynamo directly, they do demonstrate that persistent toroidal fields can coexist with convection at moderate magnetic field strengths. In light of these studies, it is important to understand how toroidal flux tubes of weak to moderate field strengths, ≈ 15 kG to ≈ 50 kG, behave as they rise through a turbulent solar convective envelope. Some studies

have been performed that investigate the buoyant rise of fully three-dimensional isolated flux tubes in a turbulent convective velocity field (*e.g.* Fan, Abbett, and Fisher, 2003; Jouve and Brun, 2009). However, due in part to the limited numerical resolution of these simulations, large values of magnetic flux must be used, which are greater than typical active-region flux, in order to keep the tube from dissipating as it rises.

Recently, Weber, Fan, and Miesch, 2011 (hereafter Article 1) incorporate a thin flux tube model in a (separately computed) rotating three-dimensional convective velocity field representative of the solar convective envelope, in an effort to study the effects of turbulent solar-like convection on the dynamic evolution of active-region scale emerging flux tubes. Although the thin flux tube model cannot capture the possible fragmentation of the flux tube and its internal magnetic structure, it does preserve the frozen-in condition of the magnetic field. The thin flux tube model is also useful in that each simulation can be performed quickly on single processor desktops, as compared to the three-dimensional simulations, which require multi-million processor hours on massively parallel supercomputers. As such, the thin flux tube approach is a useful platform for a parameter space study, and provides a starting point for the investigation of the effects of convective flows on flux tubes of realistic magnetic field strengths with magnetic fluxes similar to those of active regions observed on the Sun.

This article builds upon the work as presented in Article 1. In Article 1, it was found that as the magnetic field strength increases from 15 kG to 100 kG, the dynamic evolution of the emerging flux loops changes from being convection dominated to magnetic buoyancy dominated. For these flux tubes, the convective flow is found to reduce the rise time and the latitude of emergence through the anchoring of flux loop footpoints by downdrafts. The addition of solar-like convection also promotes tilt angles that are consistent with the observed mean tilt of solar active regions in part because of the mean kinetic helicity of the upflows in the simulated convection. In this article, we expand the study by increasing the number of simulations performed by about 3.5 times per magnetic field strength to improve the statistics of the results, sampling different time spans of the convective flows. Solar active regions exhibit magnetic flux in the range of 10^{20} Mx to 10^{22} Mx, which include ephemeral regions and pores to the largest scale sunspots (Zwaan, 1987). To study the dependence of flux tube evolution on magnetic flux, we also include cases with flux of 10^{20} Mx and 10^{21} Mx in this study in addition to the 10^{22} Mx cases. This in combination with the increased number of simulations per magnetic field strength results in a total of 5940 flux tubes to analyze. As with Article 1, we consider flux tubes with initial latitudes of 1° to 40° in both the northern and southern hemispheres. Besides the observed active-region properties described by Joy's Law, in this article we also consider additional observational diagnostics, including the scatter of active-region tilts and sunspot rotation rates to further constrain the initial field strength of active-region flux tubes at the base of the solar convection zone. We also discuss the properties of the magnetic field at the apices of the emerging flux loops. In comparison to Article 1, our new study confirms that with convection all flux tubes with magnetic field strengths of 15 kG – 100 kG do exhibit a positive Joy's Law trend, and given the uncertainties that we obtain, we note no significant dependence of the flux tube tilt angle on magnetic flux.

A description of the thin flux tube model and how it is incorporated with the spherical shell of solar-like convection employed in this study is outlined in Section 2. We highlight the results obtained from our simulations in Section 3, discussing flux tube rise times, latitudes of emergence, tilt angles, magnetic field properties, and the rotation rate of the simulated emerging loops. A summary of the results is given in Section 4.

2. Model Description

The dynamics of a thin, isolated magnetic flux tube can be described by the thin flux tube approximation. These equations are derived from the ideal MHD equations, operating under the assumption that the flux tube radius is small compared to all other relevant scales of variation associated with the flux tube (*e.g.* Spruit, 1981; Cheng, 1992; Longcope and Klapper, 1997; Fan, 2009). As in Article 1, we solve the following equations, which describe the evolution of each Lagrangian element of the one-dimensional flux tube:

$$\rho\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -2\rho(\mathbf{\Omega}_0\times\mathbf{v}) - (\rho_\mathrm{e}-\rho)\big[\mathbf{g}-\mathbf{\Omega}_0\times(\mathbf{\Omega}_0\times\mathbf{r})\big] + \mathbf{l}\frac{\partial}{\partial s}\left(\frac{B^2}{8\pi}\right) + \frac{B^2}{4\pi}\mathbf{k} - C_\mathrm{d}\frac{\rho_\mathrm{e}|(\mathbf{v}-\mathbf{v}_\mathrm{e})_\perp|(\mathbf{v}-\mathbf{v}_\mathrm{e})_\perp}{(\pi\Phi/B)^{1/2}}, \tag{1}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{B}{\rho}\right) = \frac{B}{\rho}\left[\frac{\partial(\mathbf{v}\cdot\mathbf{l})}{\partial s} - \mathbf{v}\cdot\mathbf{k}\right], \tag{2}$$

$$\frac{1}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{1}{\gamma p}\frac{\mathrm{d}p}{\mathrm{d}t}, \tag{3}$$

$$p = \frac{\rho RT}{\mu}, \tag{4}$$

$$p + \frac{B^2}{8\pi} = p_\mathrm{e}, \tag{5}$$

where $\mathbf{r}$, $\mathbf{v}$, B, ρ, p, T, which are functions of time t and arc length s measured along the tube, denote, respectively, the position, velocity, magnetic field strength, gas density, pressure, and temperature of a Lagrangian tube segment with subscript e referring to the external plasma which are functions of depth only, $\mathbf{l}\equiv\partial\mathbf{r}/\partial s$ is the unit vector tangential to the flux tube, and $\mathbf{k}\equiv\partial^2\mathbf{r}/\partial s^2$ is the tube's curvature vector, subscript $\perp$ denotes the component perpendicular to the flux tube, Φ is the constant total flux of the tube, μ is the mean molecular weight of the surrounding external plasma, $\mathbf{g}$ is the gravitational acceleration and a function of depth, $\mathbf{\Omega}_0$ is the angular velocity of the reference frame co-rotating with the sun, with Ω_0 set to 2.7×10^{-6} rad s^{-1} in this calculation, C_d is the drag coefficient set to unity, γ is the ratio of specific heats, R is the ideal gas constant, and $\mathbf{v}_\mathrm{e}(\mathbf{r},t)$ is a time-dependent velocity field relative to the rotating frame of reference that impacts the dynamics of the thin flux tube through the drag force term. In the above equations, we do not introduce an explicit magnetic diffusion or kinematic viscosity term. However, we do introduce a drag force (last term in Equation (1)), which describes the interaction of the external fluid with the flux tube in a high Reynolds number regime (*e.g.* Batchelor, 1967). The flux tube evolves passively in the external fluid such that the tube imparts no back reaction on the fluid in which it is embedded. The magnetic field of the flux tube is untwisted such that it only has a component in the $\hat{l}$ direction, and the tube is discretized with 800 uniformly spaced grid points along its arc length s. A description of the numerical methods used to solve the flux tube evolution as determined by the above set of equations is discussed in detail by Fan, Fisher, and DeLuca (1993).

The thin flux tube approximation is no longer satisfied when the radius of the flux tube is on the order of the local pressure scale height. With this in mind, the thin flux tube approximation is satisfied for the bulk of the convection zone, where the flux tube spends the majority of its time during evolution. For large flux values of $\approx10^{22}$ Mx, and small magnetic field strength of ≈15 kG, the radius of the flux tube will be the greatest, and will expand significantly as it rises toward the top of the convection zone. Due to these effects, we must

stop our simulation before the flux tube can reach the photosphere. In this case, we stop our simulations once the flux tube has reached a depth of ≈ 21 Mm below the solar surface.

A reference one-dimensional solar structure model (Christensen-Dalsgaard *et al.*, 1996) provides the stratification parameters of the external field-free plasma in the convection zone, with an extension of a simple polytropic, sub-adiabatically stratified thin overshoot layer, as described by Fan and Gong (2000). The sub-adiabatic overshoot region extends from 4.8×10^{10} cm ($r = 0.69\ R_\odot$) to 5.026×10^{10} cm from Sun center, with the super-adiabatic convection zone extending from 5.026×10^{10} cm to 6.75×10^{10} cm ($r = 0.97\ R_\odot$), which is still ≈ 21 Mm below the solar surface. Profiles of T_e, ρ_e, p_e, and the adiabaticity δ, that we use can be found in Article 1.

What sets the thin flux tube simulations in Article 1 (and continued here) apart from previous simulations is the inclusion into the drag force term of a time-dependent convective velocity field $\mathbf{v}_e(\mathbf{r}, t)$ relative to the rotating frame of reference. This three-dimensional global convection simulation is computed separately from the thin flux tube simulation using the ASH (anelastic spherical harmonic) code, as described by Miesch, Brun, and Toomre (2006) (hereafter MBT06). In the anelastic approximation, the velocity of convection flows is taken to be much slower than the speed of sound in the fluid, and convective flows and thermal variations are treated as a linear perturbation to a background state taken from a one-dimensional solar structure model. The computed $\mathbf{v}_e(\mathbf{r}, t)$ captures giant-cell convection and the associated mean flows such as meridional circulation and differential rotation, in a rotating convective envelope spanning $r = 0.69\ R_\odot$ to $r = 0.97\ R_\odot$ (4.8×10^{10} cm to 6.75×10^{10} cm from Sun center). ASH is a pseudo-spectral code with the particular simulation used here resolved by a grid of 129 points in r, 256 points in θ, and 512 points in ϕ. To find more specific details regarding this ASH simulation, see Article 1.

Boundary conditions and simulation parameters for the ASH simulation are similar to Case AB3 in MBT06. Particularly, the lower thermal boundary condition is the same as in Case AB3, with a latitudinal entropy gradient imposed to implicitly capture thermal coupling to the tachocline. This feature helps promote a conical rotation profile (see MBT06). The radial entropy gradient imposed at the outer boundary is steeper in this simulation, and therefore more in line with solar structure models (*e.g.* Christensen-Dalsgaard *et al.*, 1996), as compared to Case AB3. In this simulation the values of thermal diffusivity [κ] and turbulent viscosity [ν] at the outer boundary ($r = 0.97 R_\odot$) are 4×10^{13} cm^2 s^{-1} and 2×10^{13} respectively, and each decreases with depth in proportion to the inverse square root of the background density $\hat{\rho}^{-1/2}$. Somewhat larger than the density contrast of Case AB3 due to a slightly deeper domain, the density contrast across this ASH simulation domain is 69, corresponding to 4.2 density scale heights. This yields a mid-convection zone Rayleigh number [R_a] of 5×10^6 and Reynolds number [R_e] of order 50. The value of R_e in this simulation is about a factor of two smaller than that used in the flux tube simulations of Jouve and Brun (2009), but the value of R_a is larger.

While this ASH simulation is more laminar than some others (Miesch *et al.*, 2008; Jouve and Brun, 2009), it still possesses all of the relevant features necessary to explore the fundamental interactions between thin flux tubes and the mean flows associated with global convection. These features include: asymmetric, rotationally aligned cells at low latitudes (density-stratified banana cells), rapidly evolving downflows in the upper convection zone at high latitudes dominated by helical plumes, and a strong, solar-like differential rotation. Even the highest resolution simulations exhibit similar basic features, therefore we would not expect the essential results to change significantly with more turbulent convection.

A typical giant-cell convection pattern at a depth of 25 Mm below the solar surface is shown in Figure 1. Broad upflow cells are surrounded by narrow downflow lanes, which can

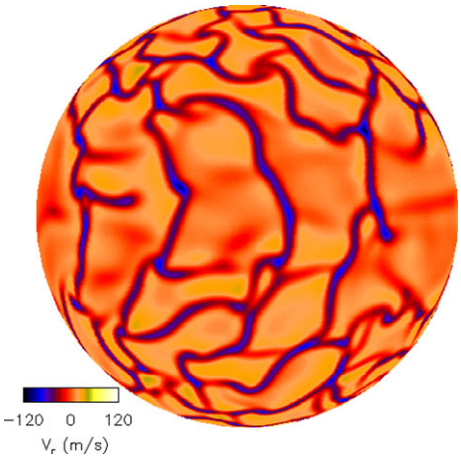


Figure 1 A snapshot of the convective radial velocity at a depth of 25 Mm below the solar surface. This simulation exhibits giant-cell convection.

reach maximum downflow speeds of nearly 600 $\mathrm{m\,s^{-1}}$ at a mid-convection zone depth of about 86 Mm below the surface. Throughout the bulk of the convection zone, the combined influence of density stratification and the Coriolis force induces anti-cyclonic vorticity in expanding upflows and cyclonic vorticity in contracting downflows. These effects yield a mean kinetic helicity density $[H_\mathrm{k}]$ that is negative in the northern hemisphere and positive in the southern hemisphere (see Article 1 for an image of the associated kinetic helicity). Such a helicity pattern is typical for rotating, compressible convection (*e.g.* Miesch and Toomre, 2009).

Columnar, elongated downflow lanes align preferentially with the rotation axis at low latitudes, reflecting the presence of so-called "banana cells" (see Figure 1). Such structures propagate in a prograde direction relative to the polar regions, due in part to differential rotation and an intrinsic phase drift similar to traveling Rossby waves (Miesch and Toomre, 2009). These features dominate the convective Reynolds stress, aiding in the maintenance of a strong differential rotation that is faster at the Equator than at the poles, comparable to that inferred from helioseismic inversions. The total angular velocity $[\Omega/2\pi]$ (with respect to the inertial frame) is solar-like, and decreases monotonically from ≈ 470 nHz at the Equator to ≈ 330 nHz at the poles, and exhibits nearly conical contours at mid-latitudes (see Article 1 for an image of the associated differential rotation), as observed in the solar convection zone (Thompson *et al.*, 2003).

As is in Article 1, our simulations start with isolated toroidal magnetic flux rings in mechanical equilibrium (neutral buoyancy), located at a radial distance to the center of the Sun of $r = r_0 = 5.05 \times 10^{10}$ cm, slightly above the base of the convection zone. To ensure initial neutral buoyancy, the internal temperature of the flux tube is reduced compared to the external temperature. The toroidal ring in neutral buoyancy is perturbed with small undular motions which consist of a superposition of Fourier modes with azimuthal order ranging from $m = 0$ through $m = 8$ with random phase relations. In this work and in Article 1, the external time-dependent three-dimensional convective flow described impacts the flux tube through the drag force associated with the flux tube. We consider a range of initial field strengths of 15 kG, 30 kG, 40 kG, 50 kG, 60 kG, and 100 kG, and initial latitudes ranging from 1° to 40° for the toroidal flux ring, with a constant magnetic flux of 10^{20} Mx to 10^{22} Mx. This range of magnetic flux values is typical of ephemeral regions and pores to the strongest sunspots. We also greatly improve statistics by performing seven simulations sampling different time ranges of the ASH convective flow for each magnetic field strength,

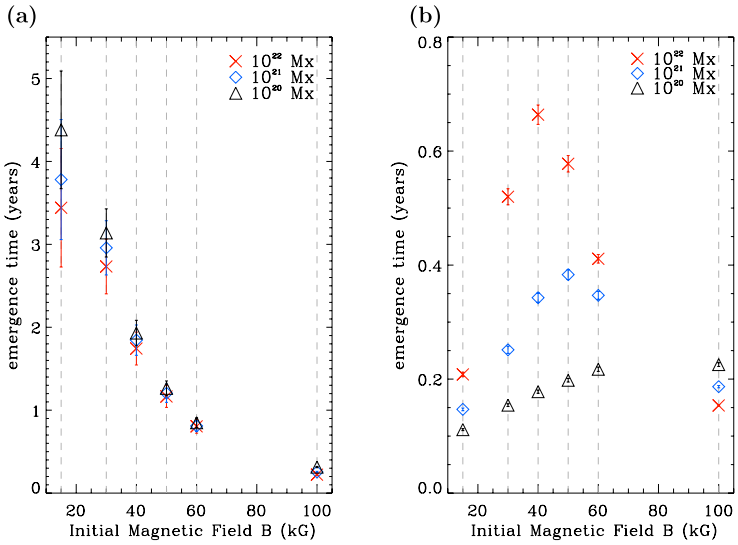


Figure 2 (a) Average rise times for flux tube simulations without the influence of convection. Red crosses are for flux tubes with magnetic flux of 10^{22} Mx, blue diamonds for 10^{21} Mx, and black triangles for 10^{20} Mx. Each symbol represents an average of nine flux tube simulations. This figure shows a decrease in emergence time with an increase in magnetic field strength and flux. (b) Average rise times for flux tube simulations with the influence of convection. Each symbol represents an average of ≈ 330 flux tube simulations. With the addition of convection, rise times of the flux tubes are shorter for tubes with smaller magnetic flux, except at 100 kG where the trend is reversed. Flux tubes of mid-field strength now take the longest time to emerge. Bars represent the standard deviation of the mean.

initial latitude, and flux, for a total number of flux tube simulations of 5940. The thin flux tube simulations sample a span of ≈ 2.7 years of consecutive ASH convective flows.

3. Results

3.1. Rise Time and Latitude of Emergence

The drag force term (the last term in Equation (1)) depends upon the magnetic field $[B]$ and the flux $[\Phi]$ through the ratio of $(\Phi/B)^{1/2}$, which is proportional to the cross-sectional radius of the thin flux tube. As this ratio appears in the denominator of the drag force term, thinner tubes will experience more drag. For any given magnetic field, the diameter of the flux tube can be reduced by decreasing the magnetic flux. In the absence of convection, the drag force acting on the rising tube reduces the velocity of the flux tube in all directions. Figure 2(a) shows a trend of increasing rise time of the flux tube simulations for both a decreasing magnetic field strength and a decreasing magnetic flux in the absence of convection, which has been found in previous thin flux tube simulations without the influence of convection (*e.g.* Moreno-Insertis, 1983; Choudhuri and Gilman, 1987; D'Silva and Choudhuri, 1993; Fan, Fisher, and DeLuca, 1993). Taking these effects into consideration, a tube with a reduced flux and a reduced magnetic field should take the longest to rise, as is shown in Figure 2(a), with 15 kG, 10^{20} Mx tubes taking the longest time to rise.

With the addition of convection, however, the rise time of the flux tubes is shorter for tubes with smaller magnetic flux except at 100 kG, as seen in Figure 2(b). In this case, the

drag force term affects how strongly the flux tube is coupled with the convective velocity field. Rising flux tubes with lower magnetic flux are advected strongly by convection. This aids the flux tube in emerging at the surface faster than it could at a larger magnetic flux, provided the drag force due to convection is significant compare to the buoyancy force. On the other hand, the magnetic buoyancy and magnetic tension of the flux tubes with an initial magnetic field strength of 100 kG dominate the drag force due to convection. As a result, flux tubes of 10^{20} Mx still take a longer time to emerge than those with a flux of 10^{22} Mx, as is the trend without convection. Similar to Article 1, flux tubes with mid-field strengths of 40 – 50 kG and magnetic flux of 10^{21} Mx and 10^{22} Mx take the longest time to emerge. For these magnetic flux and field strengths, the average convective downflows and magnetic buoyancy of the flux tube are of similar magnitudes (see Article 1). A tug-of-war exists between these two effects until one eventually dominates. Considering the ~ 11-year duration of the solar cycle, for flux tubes originating near the convection zone base, a maximum rise time of about eight months for flux tubes subject to convective flows (see Figure 2(b)) is much more realistic than a maximum rise time of about five years without convection (see Figure 2(a)).

Without the external flow field, the increased drag force on the flux tubes at lower magnetic flux aids in reducing the poleward deflection of the flux tube by the Coriolis force, such that the tube rises more radially. As a result, the flux tube emerges at lower latitudes for lower magnetic flux, as has been found in previous thin flux tube calculations in the absence of convection (*e.g.* Choudhuri and Gilman, 1987; D'Silva and Choudhuri, 1993; Fan, Fisher, and DeLuca, 1993). This is supported by Figure 3, which shows the latitudinal deflection (emergence latitude minus initial latitude) of the flux tube as a function of initial latitude in the absence of convection for magnetic flux of 10^{20} Mx (black triangles), 10^{21} Mx (blue diamonds), and 10^{22} Mx (red crosses). Also, as found in previous studies (Choudhuri and Gilman, 1987; Caligari, Moreno-Insertis, and Schüssler, 1995), a low-latitude zone with the absence of flux emergence due to poleward deflection is found for magnetic field strengths of 15 kG to 40 kG. As such, these flux tubes would not be able to reproduce active regions near the Equator.

With convection, the previous problem of poleward slippage for flux tubes of weak initial field strength is rectified. Convection produces a scatter in the emerging latitude compared to the cases without convection, which is both poleward and equatorward. Figure 4 shows the latitudinal deflection of the flux tube for fluxes of 10^{20} Mx, 10^{21} Mx, and 10^{22} Mx resulting from simulations with convection (plus signs), compared to the case without convection (diamonds). A negative latitudinal deflection indicates that the flux tube emerges close to the Equator than where it originally started. Such a phenomenon only occurs due to external convective flows. It is even possible for portions of the flux tube to cross into the opposite hemisphere. We note a systematic trend of reduced poleward deflection of the flux tube for all fluxes considered here. This is in agreement with Article 1, where we only studied 10^{22} Mx cases. It is also evident that there is a greater amount of scatter in latitudinal deflection at lower flux, which occurs because the drag force is greater and thus the tube will be advected more by convective flows. Also, as the magnetic field strength increases for a particular flux, the latitudinal deflection decreases, as is the case for tubes both with and without convection. This is a result of the buoyancy force overpowering the Coriolis force at large magnetic fields, forcing the tube to rise more radially.

3.2. Joy's Law

The phenomenon associated with solar active regions known as Joy's Law describes the tilting behavior of emerging flux regions toward the Equator (*e.g.* Hale *et al.*, 1919). A line

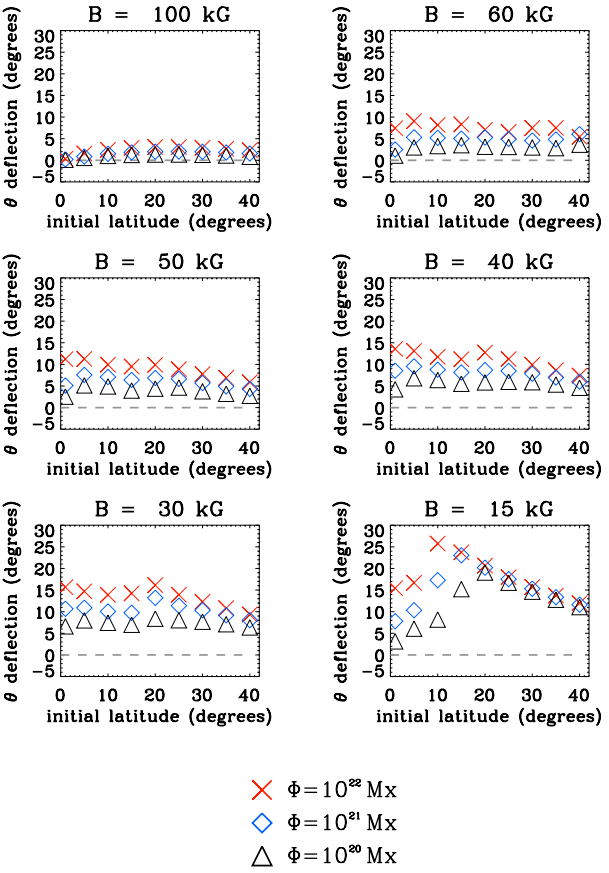


Figure 3 Latitudinal deflection (emergence latitude minus initial latitude) of the flux tube apex as a function of initial latitude for initial flux tube magnetic field strengths of 100 kG, 60 kG, 50 kG, 40 kG, 30 kG, and 15 kG in the absence of convection for magnetic flux of 10^{22} Mx (red crosses), 10^{21} Mx (blue diamonds), and 10^{20} Mx (black triangles). An increased drag force for tubes with a smaller flux reduces the poleward deflection of the tube.

drawn between the center of the two bipolar regions will be tilted with respect to the East–West direction such that the leading polarity (in the direction of solar rotation) is closer to the Equator than the following. On average, the tilt of this line with respect to the East–West direction, quantified by the tilt angle, increases with an active region's latitude of emergence. In this article, the tilt angle is computed as the angle between the tangent vector at the apex of the emerging loop (once it has reached the top of the simulation domain), and the local East–West direction. We define a positive sign of tilt as a clockwise (counter-clockwise) rotation of the tangent vector away from the East–West direction in the northern (southern) hemisphere, consistent with the direction of the observed mean tilt of active regions. If the magnitude of the tilt angle exceeds 90°, then the active region is of an anti-Hale arrangement, possessing the wrong leading polarity in the direction of solar rotation for that particular hemisphere. This is the same convention as used in Article 1.

Tilt angles of the flux tube as a function of emergence latitude for magnetic flux of 10^{20} Mx, 10^{21} Mx, and 10^{22} Mx are shown in Figure 5, for initial magnetic field strengths of 100 kG, 60 kG, 50 kG, 40 kG, 30 kG, and 15 kG both with (plus signs) and without convection (diamonds). Each of the panels in this figure includes about 3.5 times more flux tube simulations per magnetic field strength compared to the number of simulations conducted in Article 1: sampling different time ranges of the ASH convective velocity flow, therefore greatly improving statistics. As a result, each panel contains ≈ 330 points. Here

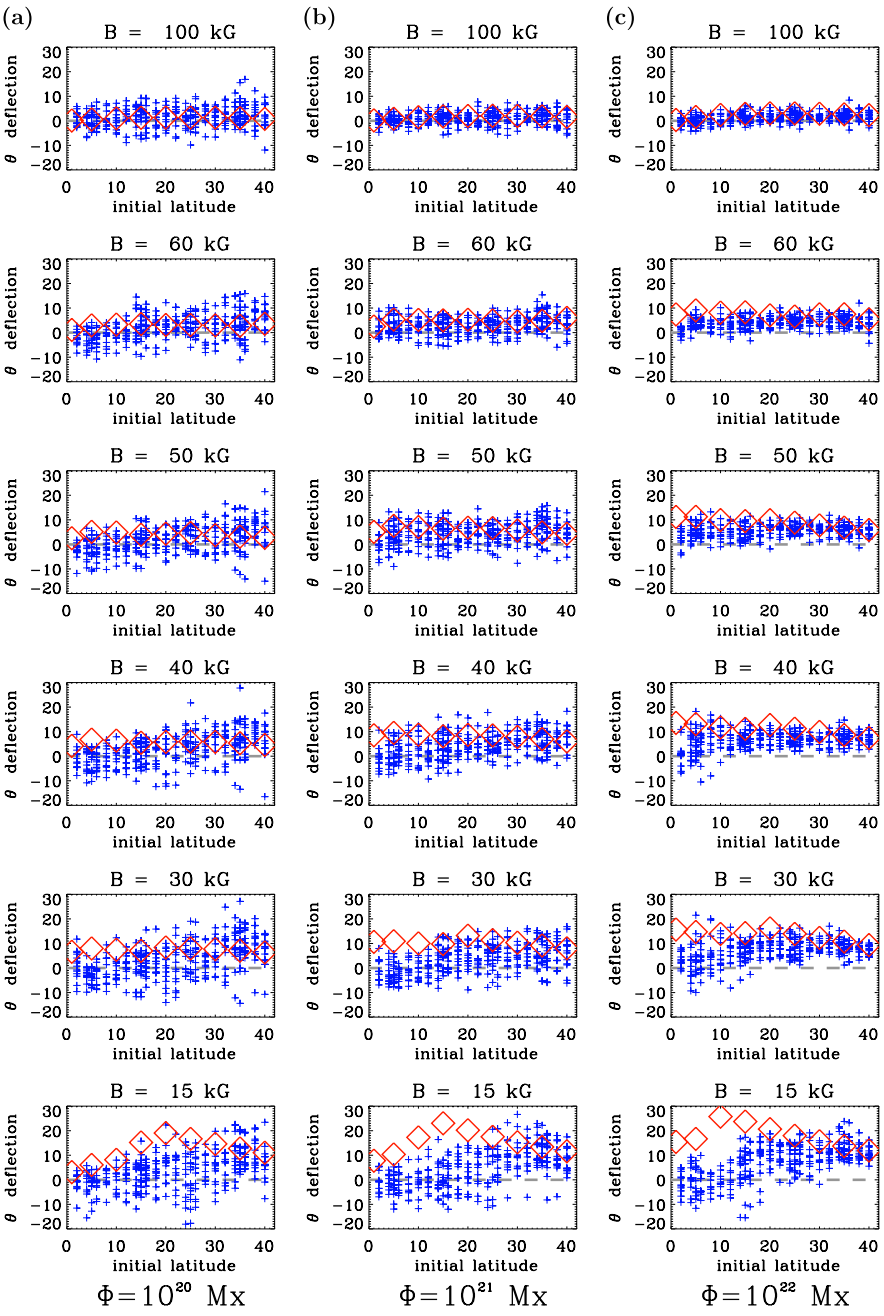


Figure 4 Latitudinal deflection (emergence latitude minus initial latitude) of the flux tube apex as a function of initial latitude for initial flux tube magnetic field strengths of 100 kG, 60 kG, 50 kG, 40 kG, 30 kG, and 15 kG with a magnetic flux of (a) 10^{20} Mx in the left column, (b) 10^{21} Mx in the middle column, (c) and 10^{22} Mx in the right column. Red diamond are for flux tubes without convection, and plus signs are for flux tubes with convective effects. Both axes are in units of degrees. Addition of a convective velocity field results in flux tubes that are able to emerge near the Equator at lower magnetic fields.

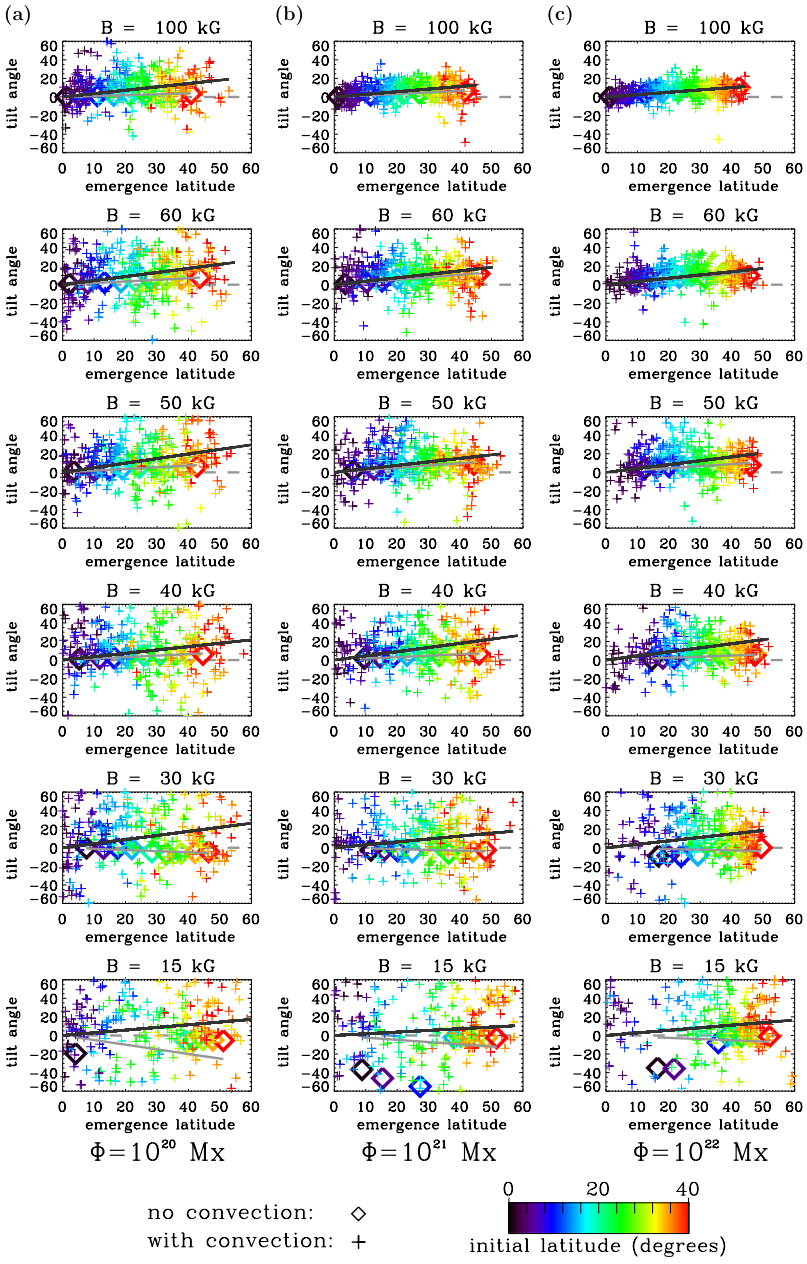


Figure 5 Tilt angles as a function of emergence latitude for initial magnetic field strengths of 100 kG, 60 kG, 50 kG, 40 kG, 30 kG, and 15 kG for cases with (plus signs) and without (diamond points) the influence of convection for magnetic flux of (a) 10^{20} Mx in the left column, (b) 10^{21} Mx in the middle column, and (c) 10^{22} Mx in the right column, with data sets sampling seven different convective velocity flow fields. The gray line is the linear best fit for flux tubes in the absence of convection, and the black line is the best-fit line for the cases subjected to convective flows. Slopes for all of these best-fit lines are shown in Table 1. A color bar indicates the original starting latitude of the flux tube. All axes are in units of degrees. Convection introduces a scatter of the tilt angle about the best-fit line, and also aids in increasing the slope of the best-fit line, especially at lower magnetic field strengths.

Table 1 Slopes m_A of the linear best-fit lines to the tilt angle as a function of emergence latitude for flux tubes without convective effects. At low magnetic field strengths, the slope of the best-fit line is negative, indicating a departure from the Joy's Law trend.

B without convection	10^{20} Mx	10^{21} Mx	10^{22} Mx
100 kG	0.10 ± 0.01	0.16 ± 0.01	0.25 ± 0.01
60 kG	0.18 ± 0.01	0.27 ± 0.01	0.30 ± 0.02
50 kG	0.18 ± 0.01	0.25 ± 0.01	0.22 ± 0.02
40 kG	0.16 ± 0.01	0.16 ± 0.01	0.12 ± 0.01
30 kG	-0.12 ± 0.01	-0.08 ± 0.03	-0.07 ± 0.05
15 kG	-0.50 ± 0.38	-0.24 ± 0.22	-0.13 ± 0.13

we also expand the investigation to include cases with flux of 10^{20} Mx and 10^{21} Mx. As in Article 1, we assume that the tilt angle increases monotonically with increasing emergence latitude, and so perform a linear least-squares fit of Joy's Law to the data using the equation $\alpha = m_A \lambda$, where α, m_A, and λ represent the tilt angle, slope, and latitude respectively. The fit is forced to go through zero because no tilt is expected for equatorial sunspot groups. Figure 5 shows in gray the best-fit line for the data without convection, and black for the flux tubes with convection. The slopes of these best-fit lines along with their uncertainties are reported in Table 1 for flux tubes without convective effects, and Table 2 for flux tubes with convective effects.

Without the influence of convection, untwisted flux tubes tilt in the appropriate direction for their respective hemispheres due to the Coriolis force acting on the limbs of the emerging loop (*e.g.* D'Silva and Choudhuri, 1993). However, for weaker initial field strengths of 15 kG and 30 kG, some emerging loops show negative tilt angles, opposite to the sign of the active-region mean tilts. This occurs because plasma flow along the flux tube near the apex changes from diverging to converging as it enters the upper convection zone (*e.g.* Caligari, Moreno-Insertis, and Schüssler, 1995; Fan and Fisher, 1996). The Coriolis force acting on the converging flow drives a tilt of the wrong sign, resulting in a negative best-fit line slope for tubes without convection at magnetic field strengths of 15 kG and 30 kG. This effect is especially severe for flux tubes of 15 kG and 10^{20} Mx. The tilt angle of these tubes is so large, and of the wrong sign for initial latitudes of 5°, 10°, and 15°, that they do not appear on the plot (Figure 5(a), lower left). As can be seen from Tables 1 and 2, we find that at all magnetic field strengths, the slopes of the best-fit lines for the flux tubes with convection are increased from the cases without convection. This occurs in part because the convective velocity field upflows have an associated kinetic helicity that helps to drive the tilt of the flux tube apex in the appropriate Joy's Law direction for its respective hemisphere (*i.e.* toward the Equator). Such a mean kinetic helicity corresponds to a vertical vorticity in upflows that is clockwise in the northern hemisphere, and counter-clockwise in the southern hemisphere.

With convection and for all magnetic fluxes, all magnetic field strengths of 15 kG – 100 kG show positive best-fit line slopes, agreeing with Joy's Law. This is an improvement upon Article 1, for which there were fewer simulations only for a magnetic flux of 10^{22} Mx and the best-fit line slopes for 15 kG and 30 kG of 0.12 ± 0.18 and 0.15 ± 0.21 respectively, had too large uncertainties to report a definitive Joy's Law trend. It is interesting to note that with convection, the mid-magnetic field strength flux tubes (40 kG – 50 kG) have the largest best-fit slopes. This is probably due to the fact that at these mid-field strengths the joint effects of magnetic buoyancy and the convective flows are such that the tubes have a longer

Table 2 Slopes m_A of the linear best-fit lines to the tilt angle as a function of emergence latitude for flux tubes with convective effects. Convection aids in boosting the slope of the best-fit line, and the slope also peaks at mid-field strengths of 40 kG and 50 kG. In comparison, using white-light sunspot group tilt-angle data, Dasi-Espuig *et al.* (2010) find $m_A = 0.26 \pm 0.05$ for Mount Wilson data and $m_A = 0.28 \pm 0.06$ for Kodaikanal data.

B with convection	10^{20} Mx	10^{21} Mx	10^{22} Mx
100 kG	0.36 ± 0.05	0.28 ± 0.02	0.25 ± 0.02
60 kG	0.44 ± 0.08	0.37 ± 0.04	0.35 ± 0.02
50 kG	0.50 ± 0.08	0.38 ± 0.04	0.42 ± 0.04
40 kG	0.36 ± 0.10	0.45 ± 0.08	0.44 ± 0.06
30 kG	0.44 ± 0.11	0.31 ± 0.11	0.37 ± 0.09
15 kG	0.29 ± 0.12	0.18 ± 0.14	0.27 ± 0.12

rise time compared to the other field strengths. Therefore, the systematic effects from the Coriolis force and kinetic helicity in convective upflows have a longer time to act on the flux tubes, and the flux tube will likely encounter multiple convection cells during its evolution. Since the mean kinetic helicity is obtained by averaging the kinetic helicity of many cells over time, a flux tube that takes a longer time to emerge will be influenced more by the mean kinetic helicity rather than stochastic fluctuations.

For solar observational data, the observed tilt angle of the active region is often plotted as a function of its emergence latitude for active regions of all sizes and magnetic flux. Figure 6 shows the tilt angles from the flux tube simulations together for selected values of magnetic field and flux. A linear least-squares fit is performed on these data for various flux and magnetic field combinations as shown in Figure 6, using the fit of $\alpha = m_A \lambda$. From white-light sunspot group tilt-angle data spanning Solar Cycles 15 – 21, Dasi-Espuig *et al.* (2010) find a linear best-fit line slope m_A of 0.26 ± 0.05 for Mount Wilson data, and 0.28 ± 0.06 for Kodaikanal data. For simulated flux tubes of all magnetic field strengths considered and fluxes of 10^{21} Mx and 10^{22} Mx (Figure 6(b)), as these are most likely the magnetic flux strengths required to produce sunspot groups that can be identified in white light, we find a slope of $m_A = 0.34 \pm 0.02$, which does fall within the range of tilt angles that Dasi-Espuig *et al.* (2010) find given the uncertainties on the slope.

Sometimes a fit is also performed using $\alpha = m_B \sin(\lambda)$, which is a good choice assuming the origin of the tilt angle is related to the Coriolis force, as this force varies with latitude as $\sin(\lambda)$. Here, m_B is the best-fit line slope. Stenflo and Kosovichev (2012) perform such a fit, which gives a slope of $m_B = 32.1^\circ \pm 0.7^\circ$, using 15 years of MDI full-disk magnetograms. Wang and Sheeley (1989) find a similar trend, using National Solar Observatory Kitt Peak magnetograms of bipolar active regions for Solar Cycle 21. Performing the same fit on the tilt angles of our simulated flux tubes for magnetic flux values of 10^{20} Mx, 10^{21} Mx, and 10^{22} Mx (Figure 6(c)), which encompasses the range of flux that these magnetograms consider, we find $m_B = 22^\circ \pm 1^\circ$. This value is much lower than what Wang and Sheeley (1989) and Stenflo and Kosovichev (2012) determine. However, using Mount Wilson white-light sunspot group data collected from 1917 – 1985, Fisher, Fan, and Howard (1995) find $m_B = 15.69^\circ \pm 0.66^\circ$. For our simulated flux tubes of 10^{21} Mx and 10^{22} Mx (Figure 6(b)), we find $m_B = 21^\circ \pm 1^\circ$, which is slightly larger than the value derived from Fisher, Fan, and Howard (1995). Recalling that mid-field strengths of 40 kG – 50 kG have the largest linear best-fit slopes m_A, we perform the fit $\alpha = m_B \sin(\lambda)$ for all fluxes we consider here only for field strengths of 40 kG – 50 kG (Figure 6(d)). The result is a significantly increased

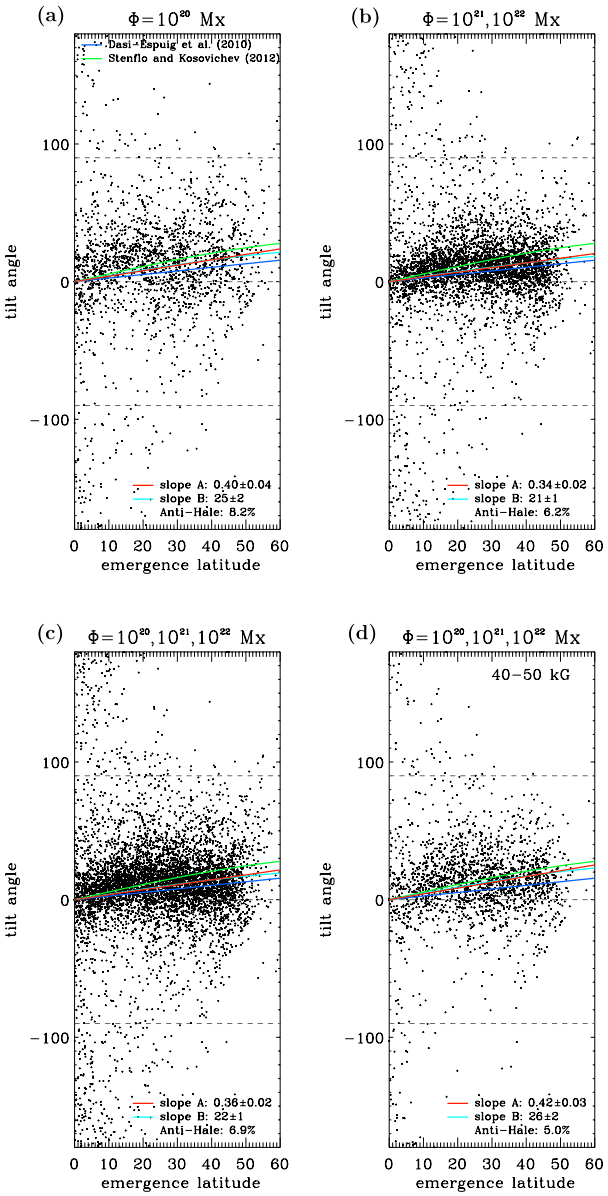


Figure 6 Tilt angles for various flux and magnetic fields plotted together. The dark blue line is the linear best-fit line for tilt angle as a function of latitude as found from white-light sunspot group observations (Dasi-Espuig *et al.*, 2010), and the green line is the best-fit line for tilt angle as a function of the sine of latitude as found from magnetograms (Stenflo and Kosovichev, 2012). In all graphs, we also plot the best-fit line of the simulations for tilt angle as a function of latitude (red, slope A), and for the tilt angle as a function of the sine of latitude (light blue, slope B) for (a) tilt angles of all magnetic field strengths and flux of 10^{20} Mx only, (b) tilt angles for all magnetic field strengths and fluxes of 10^{21} Mx and 10^{22} Mx, (c) tilt angles for all magnetic field strengths and fluxes of 10^{20} Mx, 10^{21} Mx, and 10^{22} Mx, (d) tilt angles for all fluxes for magnetic field strengths of 0 kG – 50 kG. The best-fit lines for the thin flux tube simulations are bounded by the observational best-fit lines. We find that 6.9 % of our simulated flux tubes are of an anti-Hale configuration, as compared to the ≈ 4 % as found by Stenflo and Kosovichev (2012).

value of $m_B = 26° \pm 2°$, although it still falls short of the $32.1° \pm 0.7°$ value found by Stenflo and Kosovichev (2012). For all values of magnetic flux and magnetic field that we consider for Figures 6(a) through 6(d), the best-fit lines we obtain for our simulated flux tubes are bounded by the tilt-angle best-fit lines as derived from observations. We note that the best-fit lines for the Joy's Law tilt-angle dependence derived from white-light sunspot group data (Fisher, Fan, and Howard, 1995; Dasi-Espuig *et al.*, 2010) are much lower than those derived from magnetograms (Wang and Sheeley, 1989; Stenflo and Kosovichev, 2012). This might occur because weaker active regions can be identified in magnetograms, which may not appear in white-light sunspot group images. However, it may also be a result of selection effects employed by various authors, as well as the fact that anti-Hale regions might not be identified from white-light data.

According to Fan, Fisher, and McClymont (1994), in the absence of the effects of convection, the tilt angle of the flux tube should increase with increasing flux as a result of the Coriolis force acting on the buoyantly rising flux tube. With convection, given the uncertainties that we obtain, we do not see a statistically significant dependence of the Joy's Law slope on magnetic flux (see Table 2), which agrees with the results of Stenflo and Kosovichev (2012). We have shown that the addition of convection increases the Joy's Law slope, regardless of the amount of time that the flux tubes stays in the bulk of the convection zone. This indicates that convective effects can have a significant contribution to the tilt angle of the flux tube throughout its evolution. Therefore, the results of Stenflo and Kosovichev (2012) do not rule out the paradigm that flux tubes obtain at least a portion of their tilt angle during their buoyant rise.

We also note that 6.9 % of our simulated flux tubes (see Figure 6(c)) emerge with polarities that violate Hale's Law, in comparison to the ≈ 4 % as found via observations of medium to large-sized active regions (Wang and Sheeley, 1989; Stenflo and Kosovichev, 2012). In our simulations, these violations arise as a result of flux tubes emerging in the opposite hemisphere from which they originated, or as a result of the flux tube becoming so distorted by convection that the legs of the emerging loop can become reversed. Both of these situations usually occur for weak magnetic field strength cases, with the majority of emerging anti-Hale flux tubes with moderate magnetic fields between 40 kG – 50 kG happening as a result of hemisphere crossing due to convective flows. Therefore, in light of our findings, observations of anti-Hale spots of various magnetic flux does not rule out the rising flux tube paradigm.

3.3. Tilt Angle Scatter

We attempt to further constrain the magnetic field strength at which the solar dynamo might be operating by observing the scatter of tilt angles our simulated flux tubes produce about the Joy's Law trend. In Figure 5, there is a clear scatter of the tilt angle about the best-fit line for the data with convection, which becomes greater at lower magnetic flux. Again, this trend is the result of stronger coupling between the flux tube and convective flows at a reduced magnetic flux. To quantify the scatter of the tilt angles around the best-fit line, we investigate the standard deviation of the tilt about its fitted value

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (\alpha_i - \alpha_{\text{fit}})^2}{N}}, \tag{6}$$

where α_i is the ith tilt angle, α_{fit} is the ith tilt angle as a result of the fit, and N is the number of points considered.

Table 3 Standard deviation of the tilt angle about its fitted value. The standard deviation tends to increase with decreasing flux.

	15 kG	30 kG	40 kG	50 kG	60 kG	100 kG
10^{20} Mx	63 ± 4	58 ± 3	54 ± 3	41 ± 2	40 ± 2	22 ± 1
10^{21} Mx	73 ± 4	55 ± 3	42 ± 2	30 ± 2	19 ± 1	10.4 ± 0.6
10^{22} Mx	71 ± 4	52 ± 3	35 ± 2	23 ± 1	12.5 ± 0.7	8.2 ± 0.4
10^{21}, 10^{22} Mx	74 ± 2	54 ± 2	39 ± 2	26 ± 1	16.3 ± 0.6	9.4 ± 0.4
10^{20}, 10^{21}, 10^{22} Mx	69 ± 2	55 ± 2	44 ± 1	32 ± 1	26.4 ± 0.8	14.9 ± 0.5

Using Mount Wilson white-light sunspot group data (1917 – 1985), Fisher, Fan, and Howard (1995) find that $\sigma < 40^\circ$ for all emergence latitudes. We evaluate σ for each field strength and flux value, and the results are given in Table 3. Comparing to the results of Fisher, Fan, and Howard (1995), this would exclude magnetic field strengths of 15 kG and 30 kG for the combination of fluxes of 10^{21} Mx and 10^{22} Mx, together (row 4 in Table 3) as the progenitors of solar active regions, which has a scatter that exceeds the observed value. We only consider these flux values because Fisher, Fan, and Howard (1995) use white-light sunspot group data. If we include all magnetic fields and these two fluxes together, we obtain a σ value of 41.4 ± 0.7, which is fairly close to what Fisher, Fan, and Howard (1995) obtain. However, if we exclude magnetic fields of 15 kG and 30 kG for fluxes of 10^{21} Mx and 10^{22} Mx, we obtain a σ value of 25.3 ± 0.5, well within the range that Fisher, Fan, and Howard (1995) suggest.

Active regions of flux values of 10^{20} Mx would appear on magnetograms, but would probably not be associated with white-light sunspot groups. Since we have simulations for 10^{20} Mx, we calculate σ for these values as well in Table 3. For this magnetic flux, σ is very large and exceeds the values suggested by Fisher, Fan, and Howard (1995) for all magnetic field strengths ≤ 40 kG. However, this result does not exclude these magnetic field values as progenitors of active regions for a flux of 10^{20} Mx, since these regions most likely would not appear in the Fisher, Fan, and Howard (1995) study. We do note that while the Joy's Law trend is not statistically dependent on flux (see Table 2), the scatter of the tilt angles about their best-fit line increases with decreasing flux (see Table 3). Wang and Sheeley (1989) and Stenflo and Kosovichev (2012) also find that the spread in the distribution of the tilt angles increases as one goes to smaller regions, or smaller magnetic flux.

3.4. Preferred Tilt Angle

Howard (1996) investigated the distribution of sunspot group tilt angles from a set of Mount Wilson white-light photographs (1917 – 1985) as well as sunspot groups derived from plage in Mount Wilson daily magnetograms (1967 – 1995). They bin the tilt angles in 2.5° increments and plot the number of occurrences of the tilt angles in these bins. In this way, they find that sunspot groups tend to form most often with tilt angles between $2.5^\circ - 5^\circ$ for white-light data, and between $7.5^\circ - 10^\circ$ for magnetogram data. The tilt-angle distribution that we obtain for flux tubes of all magnetic fields and magnetic flux studied here using tilt-angle bins of 2.5° is shown in Figure 7. We find that for all magnetic fluxes considered, the tilt-angle distribution peaks between bins of $10^\circ - 12.5^\circ$ (black line), where a non-linear least-squares Gaussian fit (red line) gives a center, or preferred tilt, of 10.6°. This is in good agreement with the magnetogram data of Howard (1996) and Stenflo and Kosovichev (2012), who find a preferred, or most common, tilt of $\approx 10^\circ$ for all bipolar regions emerging between a $15^\circ - 20^\circ$ latitude range including the largest active regions down to regions of $\approx 10^{20}$ Mx.

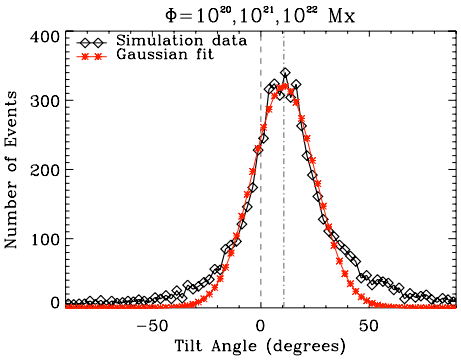


Figure 7 Distribution of tilt angles in 2.5° bins for all initial magnetic field strengths and all magnetic fluxes considered in this study. The peak of the distribution function is between 10° – 12.5° (black line), with a non-linear least-squares Gaussian fit (red line) of center 10.6°, as shown by the dash-dotted line.

Howard (1996) also finds an average tilt angle of 4.3° for sunspot groups and 6.3° for the magnetograms, whereas we find an average of $9.5° \pm 0.6°$ for all fluxes considered together, where the uncertainty is the standard deviation of the mean. The average tilt-angle value is significantly larger than what Howard (1996) finds. However, this could be due to the nature of our chosen convective velocity simulations, the thin flux tube idealization, or a result of the selection criteria that Howard (1996) uses to identify sunspot groups or plages. He counts an active region on every day that it is observed, so an active region may be represented a number of times for a single solar disk passage. As such, their data also include the evolution of the region's tilt angle as it ages.

3.5. Magnetic Fields at the Flux Tube Apex

Using thin flux tube simulations, Fan, Fisher, and DeLuca (1993) show that the preceding leg of an emerging flux loop has a stronger magnetic field than the following leg as a result of the Coriolis force, which induces a differential stretching of the rising loop. This provides an explanation for the observed coherent, less fragmented morphology for the leading polarity flux of an active region (*e.g.* Bray and Loughhead, 1979). Subsequent simulations using the more physical mechanical equilibrium state (*e.g.* Caligari, Moreno-Insertis, and Schüssler, 1995; Fan and Fisher, 1996), as opposed to temperature equilibrium (Fan, Fisher, and DeLuca, 1993), found that the leading leg of the emerging loop has a stronger magnetic field than the following only for flux tubes with an initial field strength below 60 kG. As is in Article 1, we investigate the magnetic field asymmetry by calculating $\mathrm{d}B/\mathrm{d}s$, the derivative of the magnetic field along the arc length s in the direction of solar rotation, at the apex of the emerging loop. If $\mathrm{d}B/\mathrm{d}s$ is greater (less) than zero, then the leading (following) leg has a stronger magnetic field. We find that for all fluxes, tubes with an initial magnetic field ≤ 50 kG tend to emerge with the appropriate magnetic field asymmetry of a stronger field strength in the leading side, consistent with the tendency of a more coherent leading polarity in an active region.

Recent three-dimensional radiation MHD simulations of magnetic flux emergence and sunspot formation in the top layer of the convection zone and the photosphere (M. Rempel, private communication, 2012) have suggested that the morphological asymmetry of sunspots may be caused by a flow of plasma along the emerging tube out of the leading portion of the active region into the following leg, and is less dependent on the asymmetry in field strength of the emerging loop. If this is indeed the case, then we cannot rule out flux tubes with stronger initial magnetic field strengths of > 50 kG at the base of the solar convection zone as the progenitors of solar active regions, simply due to their magnetic field asymmetries.

The nature of the magnetic field below the visible surface of the Sun is not well known. Simulations of flux emergence in the upper layers of the convection zone have been computed in domains that span from the photosphere to depths of 2 Mm to 20 Mm below the surface (*e.g.* Cheung, Schüssler, and Moreno-Insertis, 2007; Rempel, Schüssler, and Knölker, 2009; Stein *et al.*, 2011). The simulations assume a range of initial magnetic field strengths of the flux tube at the bottom of their simulation domain anywhere from ≈ 2 kG to ≈ 30 kG. We show the magnetic field strength at the apex of our thin flux tube simulations, once the tube has reached the top of the simulation domain at a depth of 25 Mm below the solar surface, for multiple initial magnetic field strengths and flux in Figure 8. Our results do support average magnetic field strength values of ≈ 500 G to ≈ 15 kG for flux tubes reaching a height of ≈ 21 Mm below the solar surface, assuming initial field strengths of 15 kG to 100 kG at the base of the convection zone for magnetic fluxes of 10^{20} Mx to 10^{22} Mx. The magnetic field strength of the flux tube at the apex is largest for larger initial fields, and the scatter of the magnetic field at the apex for a particular magnetic field strength increases with decreasing flux.

3.6. Emerging Loop Rotation Rate

Observations show that sunspots tend to rotate faster than the solar surface plasma (Howard and Harvey, 1970; Golub and Vaiana, 1978). The azimuthal velocity of leader, follower, and all sunspots (in the reference frame with a solid body rotation rate of $\Omega_0 = 2.7 \times 10^{-6}$ rad s^{-1}) as derived from Gilman and Howard (1985) are plotted in Figure 9, as well as the observed azimuthal rotation rate of the solar plasma at the surface (blue line) as determined from surface spectroscopic Doppler-velocity measurements (*e.g.* Thompson *et al.*, 2003) and the rotation rate at $r = 0.95\ R_\odot$ (red line) as found via inversions of helioseismic observations (Howe *et al.*, 2000). This image suggests that sunspots tend to rotate at nearly the same rate as the solar plasma at $0.95\ R_\odot$, and therefore rotate faster than solar surface plasma. This might also indicate that sunspots are anchored at a depth of $0.95\ R_\odot$, near the surface shear layer interface.

In order to quantify the rotation rate of emerging loops from our thin flux tube simulation, we calculate the apparent azimuthal speed of the center point between the leading and following intersections of the emerging loop with the constant-r surface of $0.95\ R_\odot$ during the last two time steps before the loop apex reaches the top of the simulation domain. This reflects the apparent azimuthal speed of an emerging active region, and it is not the actual azimuthal speed of the flux tube plasma at the tube apex. We do this only for flux tubes of 10^{21} Mx and 10^{22} Mx assuming tubes of flux order 10^{20} Mx do not produce large sunspots. Caligari, Moreno-Insertis, and Schüssler (1995) computed the azimuthal phase speed of the summit of an emerging flux loop, without convective effects. This phase motion is related to the wave character of the rising flux loop, which is similar to a transverse wave propagating along a string. They find that the azimuthal phase speed decreases with increasing height of the summit, and that it changes sign at about 50 Mm below the solar surface such that the angular velocity of the summit is smaller than that of the external plasma. However, they suggest that the resulting active region will still show a higher rotation velocity than the surrounding plasma because of the inclination difference of the leading and following legs with respect to the local vertical. This will occur when the following leg has a steeper slope than the leading leg, which is caused by the conservation of angular momentum as the tube rises (Moreno-Insertis, Caligari, and Schüssler, 1994; Caligari, Moreno-Insertis, and Schüssler, 1995; Caligari, Schüssler, and Moreno-Insertis, 1998). In Article 1, we find that, on average, all magnetic field strengths between 15 kG and 100 kG at a magnetic flux of 10^{22} Mx show

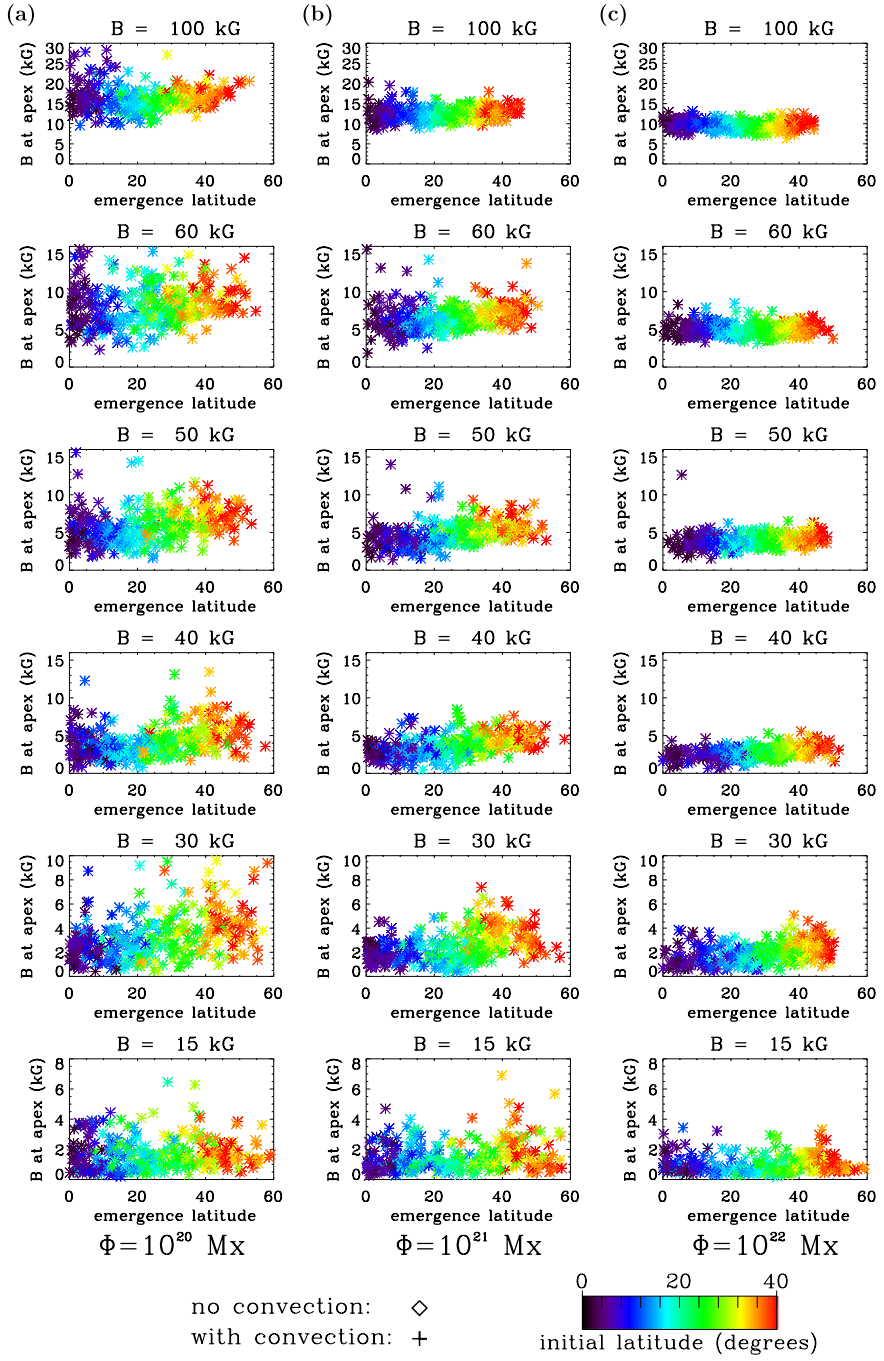


Figure 8 Magnetic field strength at the apex of the flux loop for (a) 10^{20} Mx in the left column, (b) 10^{21} Mx in the middle column, (c) and 10^{22} Mx in the right column. Our results support average magnetic field strengths of ≈ 500 G to ≈ 15 kG for flux tubes apices that have reached the top of the simulation domain, which is still 25 Mm below the solar surface.

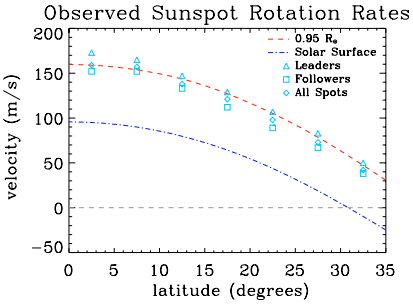


Figure 9 Rotation rate of leader, follower, and all sunspots, plotted with the observed azimuthal rotation rate of the solar plasma at the surface (blue dash-dotted line) and $r = 0.95\ R_\odot$ (red dashed line). All values are plotted with reference to the solid body rotation rate of $\Omega_0 = 2.7 \times 10^{-6}$ rad s^{-1}, so that the zero line is the solid body rotation of the Sun. This image suggests that sunspots rotate at nearly the same rate as the solar plasma at 0.95 $R_\odot$.

the appropriate inclination asymmetry of the emerging flux loop such that the leading and following polarities of a resulting active region will appear to drift apart from each other and rotate faster than the surrounding plasma. To make comparisons to actual sunspot rotation rates, we feel that the method presented in this section is a better alternative.

Figures 10(a) and 10(b) show the average of the apparent azimuthal speed in 5° bins, for cases with a magnetic flux of 10^{21} Mx and 10^{22} Mx respectively, where bars on the points are the standard deviation of the mean. For comparison, these plots also show the average azimuthal rotation rate of the ASH simulation at $r = 0.95\ R_\odot$ (red line), and the azimuthal flow speed that one would expect at the surface (blue line) assuming that the differential rotation in the top shear layer between the surface and 0.95 $R_\odot$ decreases by ≈ 10 nHz (what we will subsequently call the *inferred* surface rate) as found by helioseismology (*e.g.* Thompson *et al.*, 2003).

In all cases, flux tubes that emerge at high latitudes [above 40°] rotate at a rate that is close to, or faster than, the ASH rate at $r = 0.95\ R_\odot$ and hence faster than the inferred surface rate expected of our simulation. With initial magnetic field strengths of 15 kG and 30 kG, for both magnetic fluxes, the average rotation rates of the majority of flux tubes that emerge at $\leq 35°$ are less than the inferred surface rate. However, at mid-magnetic-field strengths of 40 kG and 50 kG, the rotation rates of the flux tubes roughly follow the inferred surface rate. Only for initial magnetic field strengths of ≥ 60 kG are the flux tubes capable of rotating at, or faster than, the inferred surface rate for all emergence latitudes, considering magnetic fluxes of 10^{21} Mx and 10^{22} Mx.

We recognize that this method of quantifying the rotation rate of the emerging flux tubes in our simulation has some limitations. Due to the nature of the ASH convection simulation and the fact that the thin flux tube simulation breaks down in the upper portion of the convection zone, it is not possible to allow the thin flux tube to emerge all of the way to the solar surface. Instead, we operate under the assumption that the rotation rate of the emerging loop at a constant-r surface 0.05 $R_\odot$ below the solar surface is a good representation for how an active region will behave at the solar surface, as is reflected in Figure 9 for solar observations. The discrepancy between the observed and simulated rotation rates could in part be attributed to the particular ASH simulation that we use, which does not precisely reproduce the solar Ω-profile. Our simulated flux tubes are effectively anchored at the convection zone base, whereas real flux tubes may decouple from the deep convection zone at some point during their evolution, and become anchored closer to the surface. This is why we choose to

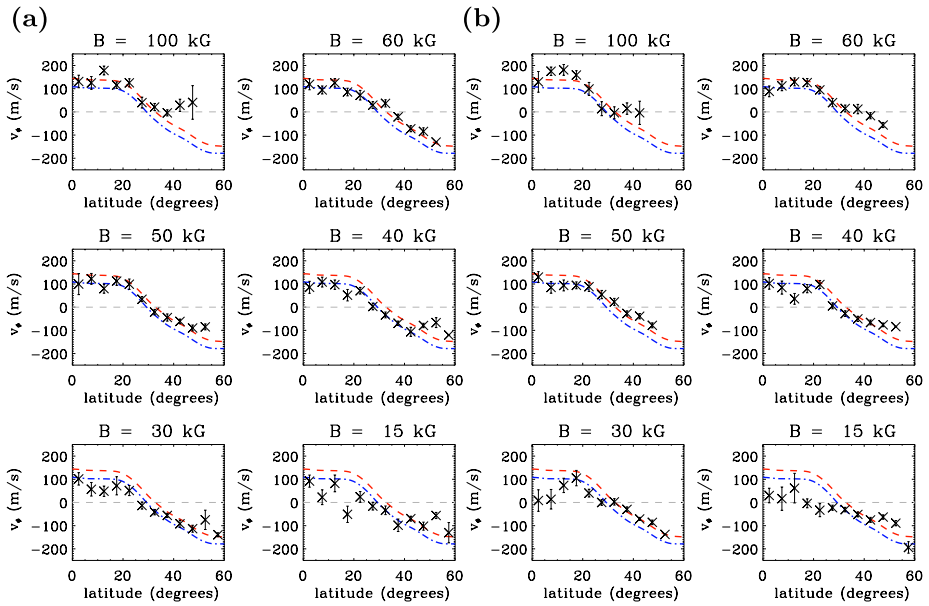


Figure 10 Average rotation rate of emerging flux loops subjected to the convective flow with magnetic flux of (a) 10^{21} Mx on the right (rightmost six panels) and (b) 10^{22} Mx on the left (leftmost six panels). Averages are taken in 5° bins, with bars representing the standard deviation of the mean. Red and blue lines correspond to the ASH rotation rate at $r = 0.95\ R_\odot$ and inferred solar surface-rotation rate respectively, assuming a difference of 10 nHz between the two. All values are plotted with reference to the solid body rotation rate of $\Omega_0 = 2.7 \times 10^{-6}$ rad s^{-1}, so that the zero line is the solid body rotation of the Sun. Strong flux tubes with initial field strengths of ≥ 60 kG rotate faster than, or nearly equal to, the inferred surface rate.

investigate the rotation rate closer to the surface. As our thin flux tube is a one-dimensional string of mass elements, we cannot address the issue that the flux tube could lose its coherency and become fragmented (Longcope, Fisher, and Arendt, 1996), which may result in a stronger coupling between the tube and convective fluid motions, and could be a significant contributing factor to the rotation rate of active regions.

In an attempt to gain a stronger coupling between the flux tube and convective fluid motions, we perform some simulations where the drag coefficient [C_d] in the last term of Equation (1) is increased from unity to the constant values of 1.5 and 2. Also, we perform simulations where we adjust the drag coefficient such that it exponentially varies from 1 at the base of the convection zone to 1.5 or 2 at the top of the simulation domain. These tests were performed for flux tubes with initial field strengths of 40 kG and magnetic flux of 10^{22} Mx. These efforts to alter the drag coefficient to produce stronger coupling to convection did not result in faster rotation rates compared to the surface rate for all latitudes less than 25°. This indicates that 40 kG flux tubes or less will not be able to reproduce sunspot rotation rates, utilizing the thin flux tube model as it currently stands.

4. Summary

By embedding the thin flux tube model in a three-dimensional, turbulent, convective velocity field representing the solar convective envelope, we study how convection can influence

the properties of emerging active-region flux tubes. In comparing these properties to those obtained from solar active-region observations, we attempt to constrain the magnetic field strengths of dynamo-generated magnetic fields at the base of the solar convection zone. The thin flux tube approximation, although idealized, allows us to investigate active-region scale flux tubes at weak to strong magnetic field strengths of 15 kG – 100 kG under perfect frozen-in flux conditions. We find that subjecting the thin flux tube to turbulent convective flows does indeed alter flux tube dynamics, and that it can have a significant impact on the properties of the emerging flux loop in comparison to flux tube simulations performed in the absence of a convective velocity field. Also, the addition of convection aids the flux tube in more closely reproducing some observed properties of solar active regions. This article extends the previous work of Article 1 by including flux tube simulations with additional magnetic fluxes of 10^{20} Mx, 10^{21} Mx, as well as 10^{22} Mx, and increasing the number of simulations performed per magnetic field strength by a factor of 3.5 in order to improve statistics such that uncertainties are reduced. We also include more observational diagnostics, such as tilt-angle scatter and sunspot rotation rate, to put further constrains on the field strength of the dynamo-generated magnetic field as the progenitors of solar active regions.

Decreasing the magnetic flux of the tube results in an increase of the drag force acting on the rising flux loop. With convection, flux tube rise times decrease with decreasing flux for initial magnetic field strengths of 15 kG – 60 kG because the increased drag force causes the flux tube to be more closely coupled with convection, and so they are advected by turbulent flows more strongly than the 10^{22} Mx case. For all magnetic fluxes that we consider here, flux tubes are able to emerge near the Equator with the aid of convective flows, which solves the previous problem of poleward slip for flux tubes of low magnetic field strengths without convective effects.

With the increased number of thin flux tube simulations, and hence improved statistics in this study, we are able to confidently say that all magnetic field strengths of 15 kG to 100 kG, and all magnetic fluxes studied here, produce emerging loops with tilt angles that follow the Joy's Law trend. This is an improvement upon Article 1 where the slope of the linear Joy's Law fit of the tilt angles of the emerging loops as a function of emerging latitude had too large an uncertainty for magnetic field strengths of 15 kG and 30 kG to report a definitive Joy's Law trend. Of particular note is the fact that helical convective upflows help to drive the tilt angle of the flux tube in the appropriate Joy's Law direction for both hemispheres. Including all tilt angles together for all magnetic field strengths and magnetic flux of 10^{21} Mx and 10^{22} Mx, we calculate the linear best-fit line for the tilt angle as a function of emergence latitude and find a slope of $m_A = 0.34 \pm 0.02$, which overlaps with the range of 0.26 ± 0.05 and 0.28 ± 0.06 as suggested by white-light sunspot group image from Mount Wilson and Kodaikanal, respectively (Dasi-Espuig *et al.*, 2010). Performing a fit for the tilt angle as a function of the sine of the latitude for 10^{20} Mx, 10^{21} Mx, and 10^{22} Mx, we find a best-fit line slope of $m_B = 22^\circ \pm 1^\circ$, which is greater than the value $15.69^\circ \pm 0.66^\circ$ obtained from white-light sunspot group data by Fisher, Fan, and Howard (1995), but less than the value of $32.1^\circ \pm 0.7^\circ$ found by Stenflo and Kosovichev (2012) using MDI magnetograms. When we exclude all fields except for 40 kG – 50 kG for all of the fluxes that we consider here, m_B increases significantly to $26^\circ \pm 2^\circ$, which is closer to the value derived from magnetograms.

On the other hand, the scatter of the tilt angles around their linear Joy's Law fit line is shown to be too large for initial magnetic field strengths of 15 kG and 30 kG with fluxes of 10^{21} Mx and 10^{22} Mx, as compared to the observed value of Fisher, Fan, and Howard (1995) for white-light sunspot group images. While the scatter of the tilt angle increases with decreasing flux (Table 3), we find no statistically significant dependence of the Joy's Law trend on flux (Table 2), consistent with the results of Stenflo and Kosovichev (2012). We

also find that the most common tilt angle produced by our study is 10.6° for tubes with a flux of 10^{20} Mx, 10^{21} Mx, and 10^{22} Mx, which agrees with Howard (1996) who finds that most tilt angles fall within the range of 7.5° – 10° as obtained from Mount Wilson magnetogram data, although our average tilt angles are higher.

Similar to previous studies (Caligari, Moreno-Insertis, and Schüssler, 1995; Fan and Fisher, 1996; Weber, Fan, and Miesch, 2011), we find that for magnetic field strengths ≤ 50 kG the leading leg of the emerging loop tends to have a larger magnetic field than the following, which may provide an explanation for the observed better cohesion of the leading polarity of an emerging active region as compared to the following polarity. This trend of asymmetry in field strength reverses for tubes with an initial magnetic field of ≥ 60 kG. However, it may be the case that the morphological asymmetry of sunspot regions is less dependent on magnetic field asymmetry, and is rather a result of the retrograde plasma flow inside the flux tube from the leading leg into the following leg as suggested by recent simulations of sunspot formation (M. Rempel, private communication, 2012). If this is indeed the case, then we cannot exclude magnetic field strengths of greater than 60 kG from the dynamo-generated magnetic field regime. A study of the magnetic field of the flux tube at the top of the simulation domain suggests typical values of 500 G to 15 kG for tubes that reach ≈ 21 Mm below the photosphere.

Observations show that sunspot groups tend to rotate faster than the surrounding solar surface plasma (Howard and Harvey, 1970; Golub and Vaiana, 1978). We use the apparent azimuthal motion of the center of the intersections of the emerging loop with a constant-r surface near the top of the simulation domain as a measure of the rotation rate of the emerging region and compare them to the average azimuthal rotation rate of the ASH convection simulation at $r = 0.95 R_\odot$, and the surface rotation rate that we would expect assuming the surface shear layer as inferred from helioseismology. For tubes with a flux of 10^{21} Mx and 10^{22} Mx, we find that at high emergence latitudes, the average rotation rate of the emerging loops tends to be greater than the inferred surface rotation rate for all field strengths considered. At lower latitudes, below about 35°, loops with initial field strength ≥ 60 kG tend to rotate faster than the inferred surface rate, consistent with the observed sunspot rotation rate, while loops with initial fields of about 40 kG – 50 kG tend to rotate at a similar rate as the surface rate. However, for initial magnetic fields below 40 kG, the rotation rate at low latitudes tends to be slower than the surface rate, contrary to observations. Thus comparison with the observed sunspot rotation rate seem to favor stronger fields as the progenitor of solar active regions. However, because of the limitations of our model, we recognize that there remain large uncertainties in our results of the azimuthal motion of emerging loops and how they relate to the observed sunspot rotations. These limitations arise because our simulations stop before the tube enters the solar surface shear layer, and because we do not address the deformation and fragmentation of the flux tubes, which may result in a stronger and more complex coupling of the magnetic fields with the fluid motion, especially in the upper layers of the convection zone.

Overall, the results in this study suggest that the initial field strength of active-region progenitor flux tubes needs to be sufficiently large, probably ≥ 40 kG, in order for them to satisfy the Joy's Law trend for mean tilt angles as well as the observed amount of scatter of the tilt angles about the mean Joy's Law behavior. Weaker magnetic fields tend to produce too large a scatter to be consistent with the observed results. Mid-field strengths of 40 kG – 50 kG, which take the longest to rise, do the best job at matching magnetogram observations of Joy's Law dependence. Although only 50 kG or greater field strengths can rotate at or faster than the solar surface rate. So, according to our thin flux tube approach, magnetic field values need to be of moderate to large field strengths for tubes with fluxes

of 10^{21} Mx and 10^{22} Mx to produce sunspot rotation behavior. The addition of multiple magnetic flux values to this study allows us to perform a more comprehensive study of the dependence of flux tube evolution with regard to magnetic flux. We find that with convective effects, the tube rise time substantially decreases with decreasing flux. There also is no statistically significant dependence on the Joy's Law trend in relation to the value of the magnetic flux, although the spread of the tilt angles about the Joy's Law trend does increase as flux decreases. Convective effects are important to flux tube development at these field strengths and magnetic flux values, and should be incorporated into future studies.

With this new study, all magnetic field strengths now show a positive Joy's Law trend, and we are now learning that the magnetic field asymmetry of sunspots could be due to a retrograde flow of plasma along the emerging flux tube. As a result, the estimate of the required ≥ 40 kG magnetic field strength is based here mainly on the tilt-angle scatter and rotation rates. However, identifying the rotation rate of the emerging flux tube, in particular, is an aspect where the thin flux tube approach may be lacking in its ability to adequately capture all of the physical processes involved. We suggest that this aspect of flux tube evolution would be a useful topic for further high-resolution three-dimensional MHD simulations of global flux emergence. In addition, more of these simulations are needed using realistic active-region flux tube parameters to further constrain the magnetic field strength generated by the solar dynamo based on observational diagnostics.

Acknowledgements This work is supported by NASA SHP grant NNX10AB81G to the National Center for Atmospheric Research (NCAR). NCAR is sponsored by the National Science Foundation. We would like to thank Nick Featherstone for reading our manuscript as our internal reviewer, and for offering helpful comments. Also, we would like to thank our referee for offering a critical review of our manuscript, which contributed to the production of a more substantial article.

References

Batchelor, G.K.: 1967, *An Introduction to Fluid Dynamics*, Cambridge University Press, New York.
Bray, R.J., Loughhead, R.E.: 1979, *Sunspots*, Dover, New York.
Brown, B.P., Browning, M.K., Brun, A.S., Miesch, M.S., Toomre, J.: 2010, *Astrophys. J.* **711**, 424. ADS: 2010ApJ...711..424B, doi:10.1088/0004-637X/711/1/424.
Caligari, P., Moreno-Insertis, F., Schüssler, M.: 1995, *Astrophys. J.* **441**, 886. ADS: 1995ApJ...441..886C.
Caligari, P., Schüssler, M., Moreno-Insertis, F.: 1998, *Astrophys. J.* **502**, 481. ADS: 1998ApJ...502..481C, doi:10.1086/305875.
Charbonneau, P.: 2010, *Living Rev. Solar Phys.* **7**, 3. http://www.livingreviews.org/lrsp-2010-3.
Cheng, J.: 1992, *Astron. Astrophys.* **264**, 243. ADS: 1992A&A...264..243C.
Cheung, M.C.M., Schüssler, M., Moreno-Insertis, F.: 2007, *Astron. Astrophys.* **467**, 703. ADS: 2007A&A...467..703C.
Choudhuri, A.R., Gilman, P.A.: 1987, *Astrophys. J.* **316**, 788. ADS: 1987ApJ...316..788C.
Christensen-Dalsgaard, J., Däppen, W., Ajukov, S.V., Anderson, E.R., Anita, H.M., Basu, S., *et al.*: 1996, *Science* **272**, 1286. ADS: 1996Sci...272.1286C, doi:10.1126/science.272.5266.1286.
Dasi-Espuig, M., Solanki, S.K., Krivova, N.A., Cameron, R., Penuela, T.: 2010, *Astron. Astrophys.* **518**, A7. ADS: 2010A&A...518A...7D.
D'Silva, S., Choudhuri, A.R.: 1993, *Astron. Astrophys.* **272**, 621. ADS: 1993A&A...272..621D.
Fan, Y.: 2009, *Living Rev. Solar Phys.* **6**, 4. http://www.livingreviews.org/lrsp-2009-4.
Fan, Y., Abbett, W.P., Fisher, G.H.: 2003, *Astrophys. J.* **582**, 1206. ADS: 2003ApJ...582.1206F, doi:10.1086/344798.
Fan, Y., Fisher, G.H.: 1996, *Solar Phys.* **166**, 17. ADS: 1996SoPh..166...17F, doi:10.1007/BF00179354.
Fan, Y., Fisher, G.H., DeLuca, E.E.: 1993, *Astrophys. J.* **405**, 390. ADS: 1993ApJ...405..390F.

Fan, Y., Fisher, G.H., McClymont, A.N.: 1994, *Astrophys. J.* **436**, 907. ADS: 1994ApJ...436..907F.
Fan, Y., Gong, D.: 2000, *Solar Phys.* **192**, 141. ADS: 2000SoPh..192..141F, doi:10.1023/A:1005260207672.
Ferriz-Mas, A., Schüssler, M.: 1993, *Geophys. Astrophys. Fluid Dyn.* **72**, 209. ADS: 1993GApFD..72..209F, doi:10.1080/03091929308203613.
Fisher, G.H., Fan, Y., Howard, R.F.: 1995, *Astrophys. J.* **438**, 463. ADS: 1995ApJ...438..463F.
Gilman, P.A.: 2000, *Solar Phys.* **192**, 27. ADS: 2000SoPh..192...27G, doi:10.1023/A:1005280502744.
Gilman, P.A., Howard, R.: 1985, *Astrophys. J.* **295**, 233. ADS: 1985ApJ...295..233G.
Golub, L., Vaiana, G.S.: 1978, *Astrophys. J. Lett.* **219**, L55. ADS: 1978ApJ...219L..55G.
Hale, G.E., Ellerman, F., Nicholson, S.B., Joy, A.H.: 1919, *Astrophys. J.* **49**, 153. ADS: 1919ApJ....49..153H.
Howard, R.F.: 1996, *Solar Phys.* **169**, 293, ADS. 1996SoPh..169..293H, doi:10.1007/BF00190606.
Howard, R., Harvey, J.: 1970, *Solar Phys.* **12**, 23. ADS: 1970SoPh...12...23H, doi:10.1007/BF02276562.
Howe, R., Christensen-Dalsgaard, J., Hill, F., Komm, R.W., Larsen, R.M., Schou, J., Thompson, M.J., Toomre, J.: 2000, *Science* **287**, 2456. ADS: 2000Sci...287.2456H, doi:10.1126/science.287.5462.2456.
Jouve, L., Brun, J.: 2009, *Astrophys. J.* **701**, 1300. ADS: 2009ApJ...701.1300J, doi:10.1088/0004-637X/701/2/1300.
Longcope, D.W., Fisher, G.H., Arendt, S.: 1996, *Astrophys. J.* **464**, 999. ADS: 1996ApJ...464..999L.
Longcope, D.W., Klapper, I.: 1997, *Astrophys. J.* **488**, 443. ADS: 1997ApJ...488..443L, doi:10.1086/304680.
Miesch, M.S., Brun, A.S., Toomre, J.: 2006, *Astrophys. J.* **641**, 618. ADS: 2006ApJ...641..618M, doi:10.1086/499621.
Miesch, M.S., Toomre, J.: 2009, *Annu. Rev. Fluid Mech.* **41**, 317. ADS: 2009AnRFM..41..317M, doi:10.1146/annurev.fluid.010908.165215.
Miesch, M.S., Brun, A.S., DeRosa, M.L., Toomre, J.: 2008, *Astrophys. J.* **673**, 557. ADS: 2008ApJ...673..557M, doi:10.1086/523838.
Moreno-Insertis, F.: 1983, *Astron. Astrophys.* **122**, 241. ADS: 1983A&A...122..241M.
Moreno-Insertis, F.: 1986, *Astrophys. J.* **166**, 291. ADS: 1986A&A...166..291M.
Moreno-Insertis, F., Caligari, P., Schüssler, M.: 1994, *Solar Phys.* **153**, 449. ADS: 1994SoPh..153..449M, doi:10.1007/BF00712518.
Moreno-Insertis, F., Schüssler, M., Ferriz-Mas, A.: 1992, *Astron. Astrophys.* **264**, 686. ADS: 1992A&A...264..686M.
Nelson, N.J., Brown, B.P., Brun, A.S., Miesch, M.S., Toomre, J.: 2011, *Astrophys. J. Lett.* **739**, L38. ADS: 2011ApJ...739L..38N, doi:10.1088/2041-8205/739/2/L38.
Rempel, M.: 2006a, *Astrophys. J.* **637**, 1135. ADS: 2006ApJ...637.1135R, doi:10.1086/498440.
Rempel, M.: 2006b, *Astrophys. J.* **647**, 662. ADS: 2006ApJ...647..662R, doi:10.1086/505170.
Rempel, M., Schüssler, M., Knölker, M.: 2009, *Astrophys. J.* **691**, 640. ADS: 2009ApJ...691..640R, doi:10.1088/0004-637X/691/1/640.
Schüssler, M., Caligari, P., Ferriz-Mas, A., Moreno-Insertis, F.: 1994, *Astron. Astrophys.* **281**, L69. ADS: 1994A&A...281L..69S.
Spiegel, E.A., Weiss, N.O.: 1980, *Nature* **287**, 616. ADS: 1980Natur.287..616S, doi:10.1038/287616a0.
Spruit, H.C.: 1981, *Astron. Astrophys.* **98**, 155. ADS: 1981A&A....98..155S.
Stein, R.F., Lagerfjärd, A., Nordlund, A., Georgobiani, D.: 2011, *Solar Phys.* **268**, 271. ADS: 2011SoPh..268..271S, doi:10.1007/s11207-010-9510-y.
Stenflo, J.O., Kosovichev, A.G.: 2012, *Astrophys. J.* **745**, 129. ADS: 2012ApJ...745..129S, doi:10.1088/0004-637X/745/2/129.
Thompson, M.J., Christensen-Dalsgaard, J., Miesch, M.S., Toomre, J.: 2003, *Annu. Rev. Astron. Astrophys.* **41**, 599. ADS: 2003ARA&A..41..599T, doi:10.1146/annurev.astro.41.011802.094848.
van Ballegooijen, A.A.: 1982, *Astron. Astrophys.* **113**, 99. ADS: 1982A&A...113...99V.
van Driel-Gesztelyi, L., Petrovay, K.: 1990, *Solar Phys.* **126**, 285. doi:10.1007/BF00153051.
Wang, Y.M., Sheeley, N.R. Jr.: 1989, *Solar Phys.* **124**, 81. ADS: 1989SoPh..124...81W, doi:10.1007/BF00146521.
Weber, M.A., Fan, Y., Miesch, M.S.: 2011, *Astrophys. J.* **741**, 11 (Article 1). ADS: 2011ApJ...741...11W, doi:10.1088/0004-637X/741/1/11.
Zwaan, C.: 1987, *Annu. Rev. Astron. Astrophys.* **25**, 83. ADS: 1987ARA&A..25...83Z, doi:10.1146/annurev.aa.25.090187.000503.

DOI 10.1007/978-1-4899-8005-2_17
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0148-9

The Sub-surface Structure of a Large Sample of Active Regions

C.S. Baldner · R.S. Bogart · S. Basu

Received: 1 May 2012 / Accepted: 24 September 2012 / Published online: 2 November 2012

Abstract We employ ring-diagram analysis to study the sub-surface thermal structure of active regions. We present results using a large number of active regions over the course of Solar Cycle 23. We present both traditional inversions of ring-diagram frequency differences, with a total sample size of 264, and a statistical study using Principal Component Analysis. We confirm earlier results on smaller samples that sound speed and adiabatic index are changed below regions of strong magnetic field. We find that sound speed is decreased in the region between approximately $r = 0.99\,R_\odot$ and $r = 0.995\,R_\odot$ (depths of 3 Mm to 7 Mm) and increased in the region between $r = 0.97\,R_\odot$ and $r = 0.985\,R_\odot$ (depths of 11 Mm to 21 Mm). The adiabatic index $[\Gamma_1]$ is enhanced in the same deeper layers where sound-speed enhancement is seen. A weak decrease in adiabatic index is seen in the shallower layers in many active regions. We find that the magnitudes of these perturbations depend on the strength of the surface magnetic field, but we find a great deal of scatter in this relation, implying that other factors may be relevant.

Keywords Helioseismology · Observations

1. Introduction

Measurements of the thermal structure beneath active regions using the techniques of local helioseismology are of substantial interest in solar physics. An increasingly wide variety of sophisticated models of the structure of sunspots and active regions are becoming available, and helioseismology can be used to test the validity of these models. In this work, we use ring diagrams to study a large number of active regions from Solar Cycle 23.

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

C.S. Baldner (✉) · R.S. Bogart
Hansen Experimental Physics Laboratory, Stanford University, Stanford, CA, 94305-4085, USA
e-mail: baldner@stanford.edu

S. Basu
Department of Astronomy, Yale University, P.O. Box 208101, New Haven, CT 06520-8101, USA

The use of ring diagrams (Hill, 1988) to study the near-surface acoustic properties, dynamics, and thermal structure of the Sun is by now well established (*e.g.* review by Gizon and Birch, 2005). In previous works, inversions of ring-diagram frequencies for structure have been performed on small numbers of active regions (Basu, Antia, and Bogart, 2004; Bogart *et al.*, 2008) to determine the changes in sound speed and adiabatic index. In these works, sound speed and adiabatic index were found to be enhanced in the layers between approximately 0.975 $R_\odot$ and 0.985 $R_\odot$ and depressed in the shallower layers between 0.99 $R_\odot$ and 0.998 $R_\odot$.

In this work, we study the sub-surface structure of a much larger sample of active regions than has previously been examined. We proceed in two ways: i) we simply extend the work of Basu, Antia, and Bogart (2004) and Bogart *et al.* (2008), performing structure inversions on ring-diagram data in the same manner as these earlier studies, and ii) in order to better take advantage of the large data set that we have, we perform a Principal Component Analysis (PCA) and study the resulting principal components. In Section 2 we introduce the ring-diagram data sets that we use in this work and discuss the differences between the sample that we use for individual inversions and the sample that we use in our PCA. In Section 3, we describe the use of Principal Component Analysis to study the ring-diagram mode parameters. In Section 4, we present the inversions of both a large number of individual rings and of the principal components from Section 3. We summarize our findings in Section 5.

2. Data

2.1. Ring Diagrams

Ring diagrams are three-dimensional power spectra of localized regions on the Sun. In this work, the data are resolved line-of-sight Doppler measurements ("Dopplergrams") taken with the *Michelson Doppler Imager* (MDI) instrument on the *Solar and Heliospheric Observatory* (SOHO). In order to achieve sufficiently high spatial frequency and sufficient signal-to-noise, we use only full-disk Dopplergrams, and require a duty cycle greater than 80 %. Because full-disk data are an MDI high-rate data product, sufficient data coverage is only available during yearly Dynamics campaigns, which are typically two or three months long. The data have a one-minute cadence and are tracked for 8192 minutes. The midpoint of the tracking interval coincides with the region of interest crossing the central meridian of the solar disk. The data are projected onto a rectangular grid using Postel's projection and corrected for the distortion in the MDI optics. This process is described in more detail by Patron *et al.* (1997) and Basu, Antia, and Tripathy (1999).

The spectra are fit in the way described by Basu, Antia, and Bogart (2004). A model of the power spectrum with 13 free parameters is fit at constant frequency. This model accounts for advection in the zonal and meridional directions as well as asymmetry in the spatial-frequency direction and azimuthally around the ring. The parameter of interest in this work is the spatial frequency $[k]$, which is returned as a function of the temporal frequency $[\nu]$ and the radial order $[n]$. In the plane-wave approximation, the spatial frequency is related to the spherical harmonic degree $[\ell]$ by $\ell = k\, R_\odot$. We interpolate the ring fits to integer values of ℓ and interpret them in the same way as global-mode parameters. We will refer to these interpolated ring-diagram fits as "mode-parameter sets"; all results in this work are based on these data sets.

The level of activity in a ring is characterized by a Magnetic Activity Index (MAI: Basu, Antia, and Bogart, 2004). This is a measure of the strong unsigned magnetic flux within the ring-diagram aperture, averaged over the tracking period.

2.2. Ring-Diagram Selection

In this work, we use two different data sets to study the structure beneath active regions. The first set consists of a set of frequency differences between active regions from the NOAA catalog and nearby regions of the quiet Sun, and it is essentially equivalent to the data sets used in earlier works (*e.g.* Basu, Antia, and Tripathy, 1999; Rajaguru, Basu, and Antia, 2001; Basu, Antia, and Bogart, 2004), albeit with a much larger sample. The second set consists of mode frequencies for a set of largely non-overlapping rings, and it is used for the statistical study described in Section 3.

The first sample is the same as used by Baldner, Bogart, and Basu (2011a). Regions are selected from the NOAA active-region catalog between June 1996 and April 2008. We include only NOAA active regions that are identified at disk center – in other words, we require that the active region be present through the center of the tracking interval. We require, too, sufficient data coverage as noted above. As in Basu, Antia, and Bogart (2004), we perform inversions relative to quiet-Sun ring-diagram mode parameters. The comparison ring diagrams are tracked at the same latitude as the active region that they are compared with. They are tracked across the disk in exactly the same way as the active regions, and they are located within 60° longitude of their active region. This is done to minimize the systematic errors in mode-parameter estimation arising from projection effects and secular changes in the instrument, as well as to minimize the effects of difficulties in modeling the near-surface layers of the solar interior. A preliminary version of this analysis was presented by Baldner, Bogart, and Basu (2011b). Although the parent active-region sample is the same in both works, the actual ring-diagram data used in the inversions have changed somewhat – in particular, the mode sets used and the comparison regions chosen. The most significant changes to the data set are improved treatment of the errors in the mode fitting and inversion, and more careful comparison region selection.

The second data set in this work is used in the PCA described in the next section. We use mode frequencies rather than frequency differences for the PCA itself. The PCA requires a filled matrix of mode parameters $[\nu_{n,\ell}]$ for a set of modes common to all of the rings in the sample. A common set of $[n, \ell]$ modes are chosen, and the ring-diagram fits are interpolated to these targets.

For the PCA procedure, we wish to minimize overlapping data, which could lead to some spurious statistical results. We choose only ring diagrams that were constructed from tracked data patches that do not overlap. The ring diagrams in this sample are the same size and tracking length as in the data set used for our individual inversions but are distributed somewhat differently so as to minimize spatial overlap rather than guaranteeing coverage of all NOAA active regions.

3. Principal Component Analysis

The use of Principal Component Analysis (PCA) in the analysis of helioseismic data was described by Baldner and Basu (2008) in the context of global-mode analysis. In brief, PCA finds an efficient representation of a set of data vectors as a linear combination of orthogonal principal components. It is efficient in the sense that, when properly ordered, each principal component is responsible for less variance in the data than the one that precedes it. In many applications, the data can be adequately represented using a relatively small number of principal components. An added benefit is that errors in measurements can be greatly reduced. In this work, we are primarily interested in using PCA to separate spurious signals in the

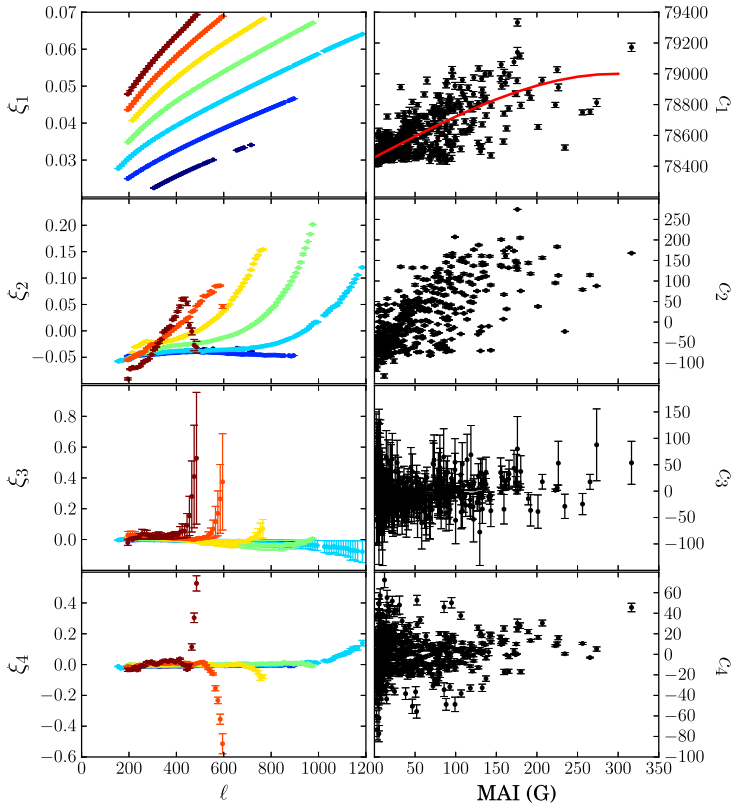


Figure 1 The first four principal components and the associated scaling coefficients. The first four principal components [ξ_1 through ξ_4] are shown in the left-hand panels as a function of degree [ℓ]. Each order [n] is shown in a different color. The right-hand panels show the scaling coefficients for each ring as a function of the ring's MAI. The red line in the top-right panel is a smoothed fit to the scaling coefficients.

mode parameters from those actually associated with solar activity, and to parametrize the changes in solar structure around active regions in a simple manner.

We use the following notation: if the set of all m frequency measurements [ν_j] for the ith ring diagram form the vector $\mathbf{D}_i = \{\nu_0 \ \ldots \ \nu_m\}$, each observation $\mathbf{D}_i$ can be reproduced completely with the m principal components [ξ_k] (each a vector of length m) and the m scaling coefficients [$c_{k,i}$] associated with that measurement:

$$\mathbf{D}_i = \sum_{k=0}^{m} c_{k,i}\xi_k. \tag{1}$$

The principal components [ξ_k] are normalized to unity. Errors in each component of the principal components and in the scaling coefficients are computed using a Monte Carlo simulation.

The results of the PCA are shown in Figure 1. The first component may be interpreted as the basic ℓ–ν diagram for a ring-diagram fit. The scaling coefficients [c_1] give the appropriate scaling for each observation, and we can see a general trend with increasing magnetic activity, although the scatter is large. The second component also has a significant dependence on magnetic activity. The following two principal components are shown in Figure 1

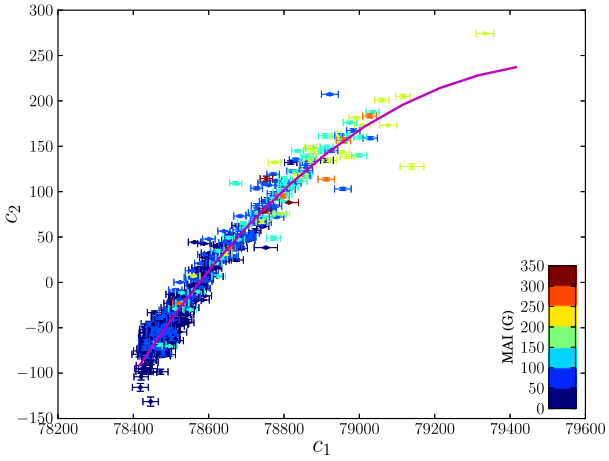


Figure 2 The second principal component scaling coefficients as a function of the first. The points are color-coded by MAI. The magenta line is a smoothed B-spline fit to the points.

and are dominated by tails at the end of each n ridge. These tails arise in the fits as the signal-to-noise decreases at the high-ℓ end of a ridge or the ridges begin to overlap significantly at low ℓ. In either case, the fits become less reliable. In inversions on real data, we truncate the fits where these tails become significant.

Higher-order components are not shown – they are in general dominated either by tails or by small numbers of outliers. Most of these components are significant only in one or at most a small number of rings. Some components show dependence on latitude (for example, ξ_5 and ξ_{12}) or on epoch (ξ_7). By choosing to neglect these components when reconstructing the data sets, we remove many small systematic effects and errors from the frequency measurements.

Given the apparent dependence of the scaling coefficients c_1 and c_2 on magnetic activity, it is desirable to reduce these coefficients to simple functions of MAI. In Figure 1, we show a smoothed fit using B-splines to the c_1 coefficients as a function of MAI. These may be used to approximate the dependence of these coefficients on magnetic activity. Since c_1 and c_2 both depend on MAI, we examine in Figure 2 their dependence on each other. It is clear that they are tightly correlated. Thus, for a given c_1, which we can relate approximately to MAI, we can determine c_2 fairly accurately. The scaling coefficients c_3 for the third principal component do not correlate well with c_1 or c_2, or with any other quantity that we examined.

It is important to know how well we can reproduce the actual ring-diagram mode parameters with a small number of vectors. In Figure 3 we plot the residuals between the reconstruction using three principal components of three randomly selected rings, and the actual ring data. We find that, especially for low MAIs, the agreement is fairly good. At the edges of ridges there are discrepancies, but, as noted above, the spectrum fits tend to become unreliable at either end of the power ridges.

4. Inversions for Structure

To determine the structure of the solar layers beneath active regions (or, in general, regions of high surface magnetic field), we invert frequency differences for sound speed or for the first adiabatic index [Γ_1]. The treatment of this problem as a linear inversion is by now well established both in global and in local seismology. In this work, we use Subtractive

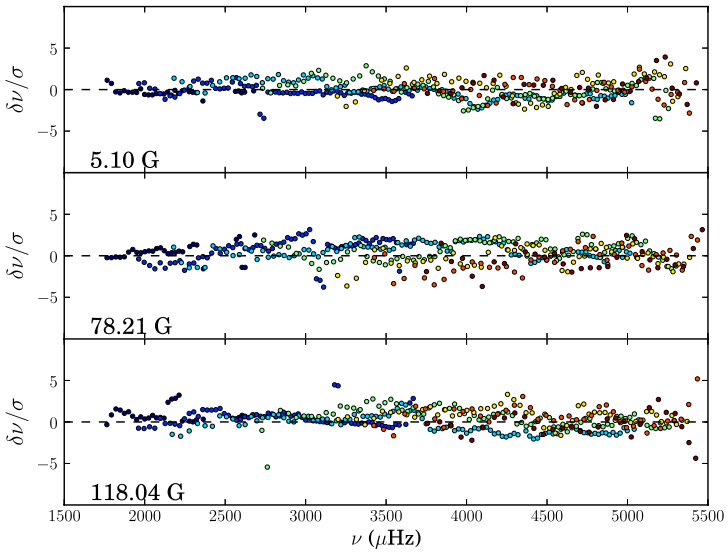


Figure 3 The difference between frequencies obtained from three ring diagrams and their representations using the first two principal components. The residuals are divided by their errors. The MAIs of the three rings are shown in the lower right. As in Figure 1, different orders n are plotted in different colors.

Optimally Localized Averages (SOLA: Pijpers and Thompson, 1992, 1994) to perform the inversions.

In our SOLA inversion, there are four free parameters: the error suppression parameter [μ], the cross-term trade-off parameter [β], the width of the target kernel [Δ_A], and the number of B-splines used to remove the surface term [Λ]. The selection of these parameters has been described in earlier works, *e.g.* Rabello-Soares, Basu, and Christensen-Dalsgaard (1999).

4.1. Inversions of Individual Rings

The principal difficulty in performing an inversion for structure from helioseismic data is in choosing the appropriate inversion parameters, which has limited earlier works. Baldner, Bogart, and Basu (2011b) found that a few combinations of inversion parameters worked well on a small subset of the active regions in our sample that was studied in detail. In this work, we again performed a thorough exploration of the parameter space on a small subset of the rings, including comparisons to rings published in earlier works. We find that the inversion results are not strongly dependent on the choice of surface term, so long as a surface term is used. When the surface term is not included in the inversions, very large perturbations are returned by the inversions, which are almost certainly spurious. We restrict ourselves to $\Lambda = 4$. In exploring the effects of the target kernel width, we find that smaller values of $\Delta(r_0)$ tend to cause oscillatory solutions to many inversions. Since we have not found any particularly sharp features in our inversions, we use a fairly large value of $\Delta(r_0) = 0.055$, which suppresses some oscillatory behavior in certain inversions, and does not overly smooth actual structure in better inversions.

Finally, for the choices of the error suppression term and the trade-off parameter, we have found that the values of these parameters for acceptable inversions in the rings we have

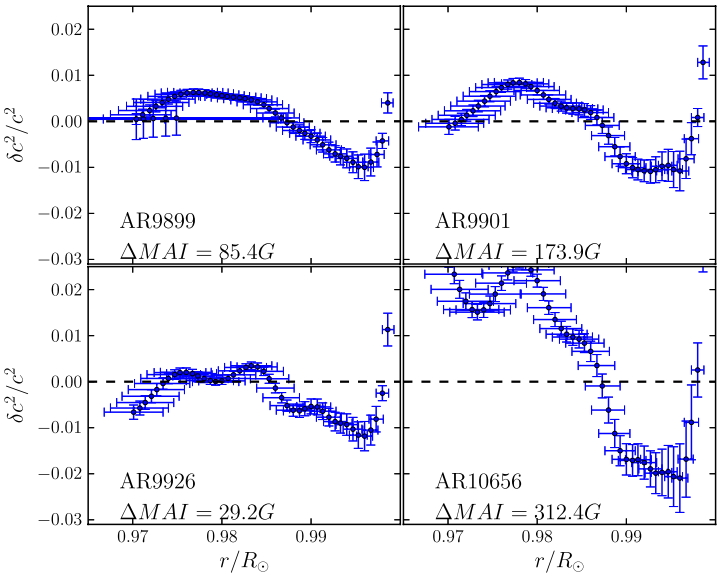


Figure 4 Examples of inversions for adiabatic squared sound speed $[c^2]$ for four active regions. The sense of the inversions is active minus quiet. The plotted horizontal error bars are the distance between the first and third quartile points of the inversion averaging kernels, and they represent the resolution of the inversion. The vertical error bars are the formal errors in the inversions.

examined carefully fall in a fairly narrow range. For the full sample, then, we run a batch of inversions for each region over the sample. Examining the inversions for each region can then be done fairly quickly and the best set of inversion parameters selected. We reject regions with unstable inversions – that is, regions whose inversion results are very strongly dependent on the choice of inversion parameters. What remains comprises our sample of sound speed and adiabatic index inversions.

Inversions for the difference in squared sound speed $[c^2]$ and adiabatic index $[\Gamma_1]$ were performed for all regions in the sample with ΔMAI > 40 G. Figure 4 shows example sound-speed inversions for four different rings with a range of active-region strengths as a function of depth. Figure 5 shows inversions for adiabatic index for the same regions.

Figure 6 shows averages of the inverted sound speed for all regions in our sample for different depth ranges. For sound speed averaged between $0.975\,R_\odot$ and $0.985\,R_\odot$ sound speeds are generally enhanced in the presence of magnetic fields, while in the region from $0.99\,R_\odot$ to $0.998\,R_\odot$ sound speeds decrease. In both regions, the magnitude of the change tends to increase with magnetic-field strength, although the relationship seems to be more of an envelope than a linear relation. Further, there seems to be some saturation of the effect at very high magnetic-field strengths.

In the shallowest layers ($r > 0.998\,R_\odot$), the sound-speed inversion results show a sharp positive change. This feature is present in every inversion in our sample. Caution should be used in interpreting this, however, as the averaging kernels become strongly asymmetric near the surface and the cross-term contributions begin to become significant. Examples of this sharp feature can be seen in the inversions in Figure 4.

We have also inverted for the first adiabatic index. In Figure 7, we show the same averages as Figure 6, but for the Γ_1 inversions. In general, Γ_1 is an easier quantity to invert for (in

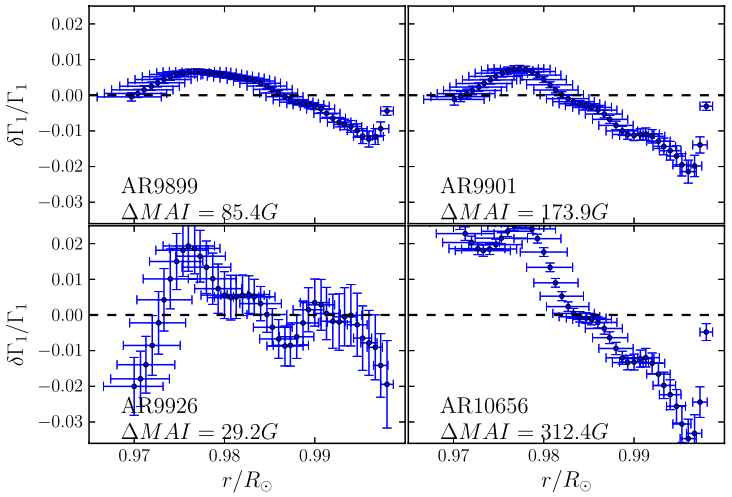


Figure 5 Examples of inversions for adiabatic index [Γ_1] for four active regions. The regions shown are the same as those in Figure 4.

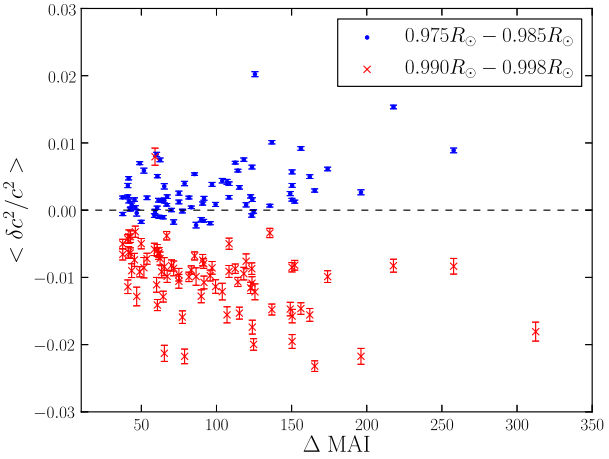


Figure 6 Averages of inversions for c^2 over two depth ranges are shown, plotted as a function of ΔMAI. Blue points are averages of inverted sound speed between 0.975 $R_\odot$ and 0.985 $R_\odot$; red crosses are averages of inverted sound-speed difference between 0.99 $R_\odot$ and 0.998 $R_\odot$.

the sense that the inversions tend to be less sensitive to the inversion parameters), and so we include a larger number of inversions for Γ_1 than we could for c^2.

We find that, as for the c^2 results, there is a depression in Γ_1 in the shallower layers that we invert (above $r = 0.99\,R_\odot$), and that, in many cases, there is a corresponding enhancement below approximately $r = 0.98\,R_\odot$, as was found in earlier works. We find that the deeper enhancement is, for most rings, much less pronounced than for the c^2. For some regions, in fact, we do not see any positive perturbation at all, and in general we find only a weak correlation with magnetic activity.

The depth ranges of the negative and positive perturbations are found to be relatively constant. In Figure 8, we plot the depths at which our inversion results change from negative to positive. We find that, at lower activity levels, there is greater variance in the depths of the perturbations, but this can be explained by lower signal-to-noise in the inversions. Beyond that, we do not find any significant change in the depth ranges of the perturbations.

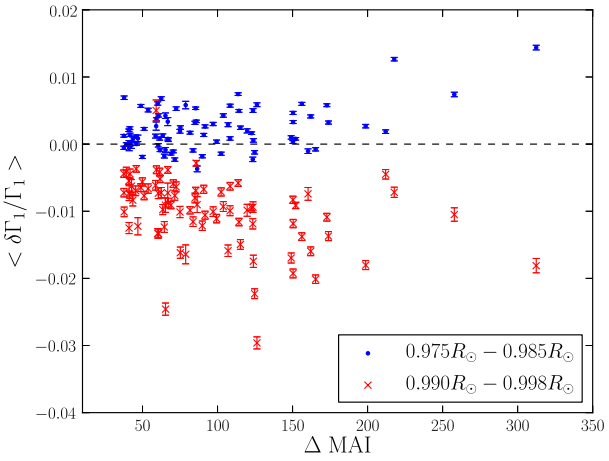


Figure 7 Averages of inversions for differences in Γ_1 over two depth ranges are shown, plotted as a function of ΔMAI. Blue points are averages of inverted sound speed between 0.975 $R_\odot$ and 0.985 $R_\odot$; red crosses are averages of inverted sound speed between 0.99 $R_\odot$ and 0.998 $R_\odot$.

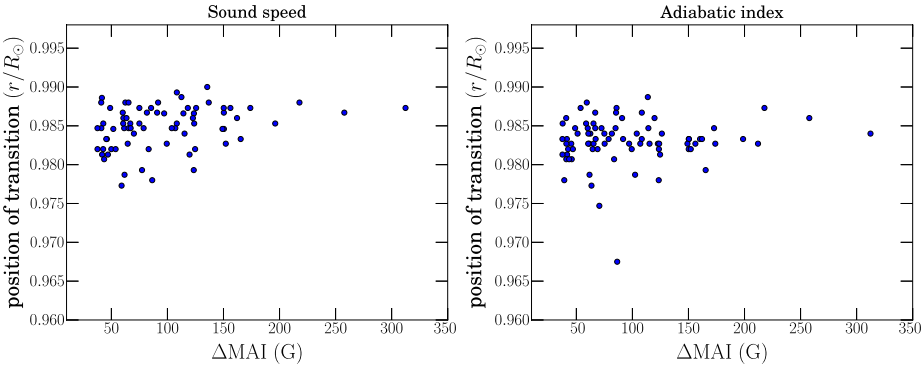


Figure 8 Position of transition, in fractional radius, between the negative (shallower) inversion results and the deeper (positive) perturbations in the inversion results, plotted as a function of ΔMAI. At left, the sound speed results are shown, at right are the results for adiabatic index. These transition points are found by treating the inversion points as a continuous curve and finding the zero-crossing point.

4.2. Inversions of PCA Data

In this section we use the PCA reconstruction of ring-diagram mode parameters described in Section 3 to perform inversions for structure. As discussed earlier, the advantages of using PCA on this data set are reduced errors in the reconstructed data set and removal of certain systematic effects in parameter estimation due to secular changes in the MDI instrument over time and projection effects. In addition, because we may reduce the data to a number of linearly independent vectors, we substantially reduce the number of inversions that need to be done.

Most of the variation in the ring-diagram mode-parameter sets is captured by the first two principal components (see Figure 3). Furthermore, as the coefficients c_1 are relatively well correlated with MAI, and c_2 is tightly correlated with c_1, we can reconstruct ring-diagram frequencies as a function of MAI. We then compute frequencies $[\nu_{n,\ell}]$, the difference in frequency $[\delta\nu_{n,\ell}]$ between the target MAI and zero MAI, and the error in $\delta\nu_{n,\ell}$. With these

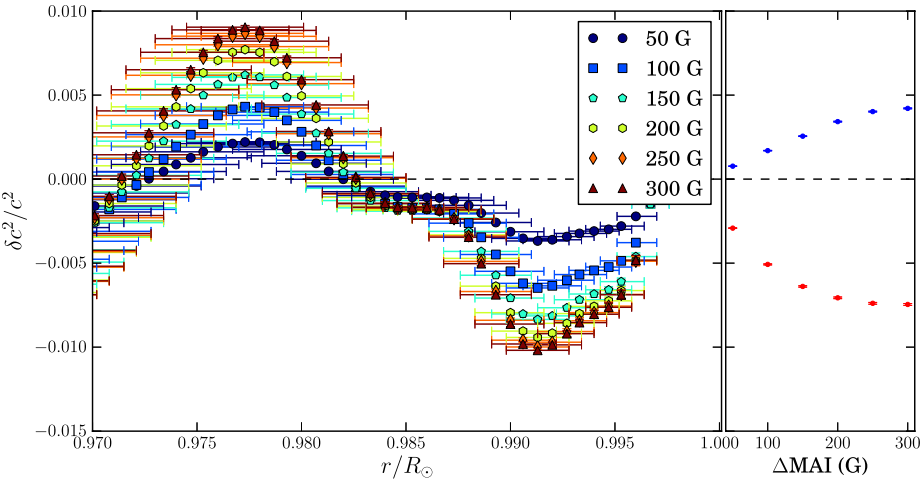


Figure 9 Inversion of the first two principal components at different MAIs for c^2. The scaling coefficients are chosen as discussed in the text. The left-hand panel shows individual inversions as a function of radius. As in Figure 4, the horizontal error bars are the difference between the first and third quartile points of the averaging kernels, and represent the resolution of the inversion. The vertical error bars are the formal errors in the inverted quantities. The latter are generally smaller than the points. The right-hand panel shows the average of the inversion results over the same depth ranges as Figure 6 as a function of MAI.

quantities, we may perform a SOLA inversion in the same manner as in the previous section.

The inversion results for six target MAIs between 50 G and 300 G are shown in Figure 9. As in the inversions of individual rings, the inversions of the PCA reconstructions show a distinctive two-layer structure with a shallow negative perturbation and a deeper positive perturbation. In addition, the magnitudes of these perturbations scale with MAI, and the saturation effect hinted at in the results in Figure 6 is clearly present here.

While most of the variation is spanned by ξ_1 and ξ_2, there remains significant signal in subsequent components. We are interested to know, then, what effects these higher components have on inversions for structure, and so we invert components separately. It must be noted, however, that inversion results from different components, each with different errors, cannot be simply combined together linearly. With that in mind, we proceed to invert various principal components separately.

In Figure 10, we show inversions for sound speed for principal components ξ_2 through ξ_5. Because it is not clear how to parametrize the scaling coefficients for any of the components except ξ_2, we simply scale the component by the largest coefficient for that component, and perform the inversion. We invert the components themselves, not explicit differences – thus the inverted sound speeds are the difference between a ring with a large coefficient c_i and a ring without any contribution from that component at all. We see that ξ_3 and ξ_4 do return changes in structure, but ξ_5 is consistent with no change.

We also invert for adiabatic index. Figure 11 shows inversions for the first two principal components at various MAIs between 50 G and 300 G. We find general consistency with the individual ring-diagram inversions: a shallow negative and deeper positive perturbation. The correlation with MAI in the shallow layer is very weak, and in some rings there is no negative perturbation at all. We find that the boundary between the negative and positive perturbations is somewhat deeper than in the individual inversion for Γ_1, occurring between 0.975 $R_\odot$ and

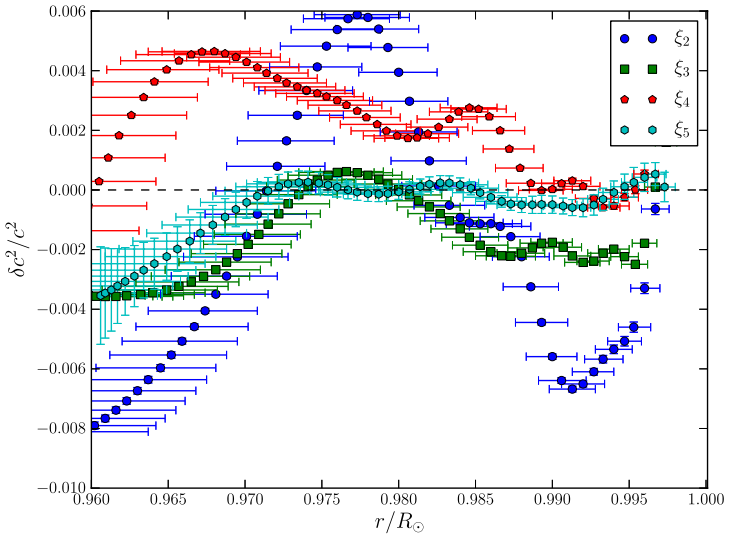


Figure 10 Inversion of the ξ_2 through ξ_5 principal components individually. Each component ξ_i is scaled by the largest associated coefficient c_i in our data set. The base frequencies are the same as the base frequencies used in Figure 9.

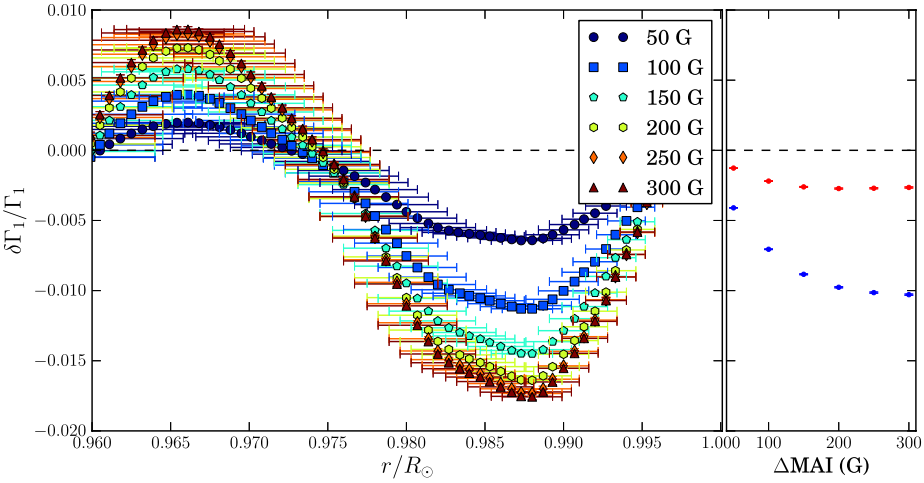


Figure 11 Inversion of the first two principal components at different MAIs for adiabatic index $[\Gamma_1]$. The data are the same as those inverted in Figure 9. The left-hand panel shows the individual inversions, the right-hand panel shows averages over the same depth ranges as Figure 7 as a function of MAI.

0.98 $R_\odot$. This discrepancy does not appear to arise from the choice of modes used in the inversion, nor does it depend on the choice of inversion parameters. One possibility is that the contributions from other principal components may shift the boundary upward, which would in turn imply a dependence of Γ_1 on something other than MAI. This is in fact what we find. On examination, the ξ_4 principal component returns a positive perturbation in the region we expect, and little signal elsewhere. The magnitude of the ξ_4 component (the c_4

coefficients from Figure 1) show a large scatter at low MAI values, but a pronounced bump above approximately 150 G. This is broadly consistent with the individual inversion results shown in Figure 7, where many of the regions with $\Delta\mathrm{MAI} < 150$ G do not show positive perturbations.

We have explored the first ten principal components in some detail and not found evidence of this, however.

5. Conclusions

In this work, we have presented inversions for structure of a large number of active regions, and used Principal Component Analysis to parametrize more accurately these inversion results. We confirm earlier results using ring-diagram analysis (Basu, Antia, and Bogart, 2004; Bogart *et al.*, 2008), as well as high-degree global-mode analysis (Rabello-Soares, 2012) that finds a two-layer thermal structure beneath surface magnetic activity. We find a shallow negative sound-speed perturbation and a somewhat deeper positive perturbation, with a similar structure in the adiabatic index. The magnitudes of these perturbations generally increase with the magnitude of the surface magnetic field.

Bogart *et al.* (2008) found a linear correlation between magnitude of the sound-speed change and strength of the active region. We find a similar relation, but with substantial scatter in our inversions of individual rings. It is perhaps more correct to say that we find an "envelope" related to MAI than a real correlation. Further, we find that the correlation appears to saturate at high field strengths. We find that the positive perturbation in Γ_1 is much weaker than in c^2, and that the correlation with magnetic activity is much less significant. We find consistent values for the strength of the perturbations in the shallower layers with Basu, Antia, and Bogart (2004) and with Bogart *et al.* (2008), but we find consistently smaller values for the deeper, positive perturbations than those found in the previous works. The choices of inversion parameters and mode sets can have some effect on the magnitudes of the inversions. In Basu, Antia, and Bogart (2004), RLS inversions tended to return slightly larger perturbations than the SOLA inversions, which we employ in this work. The most significant effect on our results appears to be our decision to neglect the f-mode in our inversions, which, as noted above, made our inversion results more stable. It also appears to have decreased the magnitude of the sound-speed change in the deeper layers.

Bogart *et al.* (2008) also reported that the depths of the positive perturbations in Γ_1 were deeper than those in the c^2 inversions. We find this unambiguously in the PCA inversion results (Figure 11), and it is also consistent with the inversions of individual rings.

Interpreting these results is not straightforward. The inversions that we perform in this work are, strictly speaking, only valid in a spherically symmetric, non-magnetized star. Clearly, neither of these conditions is satisfied in active regions. In the presence of magnetic fields, wave propagation is affected both by the direct effect of the Lorentz force, and by the effects of magnetic fields on the thermal structure. At some point – at some minimum field strength or field configuration – the entire linearized inversion becomes meaningless, as the contributions from magnetic fields to the wave propagation (which is, among other problems, no longer insensitive to direction) become significant.

In some studies, the choice is made to interpret the inferred change in sound speed as a change in the local wave speed, as was done in Gizon *et al.* (2009) and Moradi *et al.* (2010), rather than interpreting it as the effects of a thermal perturbation as we do here. In these works, significant disagreement was found when ring-diagram inversions were compared to time–distance inversions for the same solar data. The resolution of this discrepancy remains

unclear. Lin, Basu, and Li (2009), making the assumption that we measure a wave-speed perturbation in these inversions, claimed that the magnetic effect can be disentangled from the thermal effect. In the case of time–distance analysis, Braun *et al.* (2012) have shown that travel-time shifts determined from a model sunspot match neither what the thermal perturbation should give nor what they would expect from the magnetoacoustic fast mode speed. The implication of this work is that, in the case of time–distance analysis, at least, it is not correct to interpret inversion results as either changes in the thermal sound speed or as local wave-speed perturbations. Equivalent work has not yet been done for ring-diagram analysis, but must be done to determine the extent to which the assumptions we have made are valid.

In studying our sample of rings using PCA, we have decomposed the ring-diagram frequency measurements into a set of linearly independent components. The dependence of the first two of these on magnetic activity allows us to parametrize most of the frequency variance across our sample of rings as a function of MAI, and to determine what changes in sound speed and adiabatic index could give rise to these changes. We find that both sound speed and adiabatic index have a dependence on MAI that is consistent with what has been found in individual rings.

There is further variance in the ring-diagram frequency measurements, however, and MAI alone does not adequately parametrize the ring-diagram frequencies, as can be seen in Figure 1. Many of these changes are due to errors in the ring fitting and systematic effects due to projection effects and secular changes in the MDI instrument, but some of the more significant principal components may represent changes that are solar in origin and that may be associated with thermal changes below the solar surface, as shown in Figure 10.

We have demonstrated that PCA can be a useful tool in ring-diagram analysis and structure inversion. The large volume of data being returned from the *Helioseismic and Magnetic Imager* (HMI) on the *Solar Dynamics Observatory* (SDO) spacecraft represents a significant data analysis challenge. It is possible that this technique might prove a feasible alternative to attempting inversions on every ring diagram produced by HMI.

Acknowledgements This work was partially supported by a NASA Earth and Space Sciences fellowship NNX08AY41H to CSB. CSB and RSB are currently supported by NASA grant NAS5-02139 to Stanford University. SB acknowledges support from NASA grant NNX10AE60G. This work utilizes data from the *Solar Oscillations Investigation/Michelson Doppler Imager* (SOI/MDI) on the *Solar and Heliospheric Observatory* (SOHO). SOHO is a project of international cooperation between ESA and NASA. MDI is supported by NASA grant NNX09AI90G to Stanford University.

References

Baldner, C.S., Basu, S.: 2008, Solar cycle related changes at the base of the convection zone. *Astrophys. J.* **686**, 1349 – 1361. doi:10.1086/591514.

Baldner, C.S., Bogart, R.S., Basu, S.: 2011a, Evidence for solar frequency dependence on sunspot type. *Astrophys. J. Lett.* **733**, L5. doi:10.1088/2041-8205/733/1/L5.

Baldner, C.S., Bogart, R.S., Basu, S.: 2011b, The thermal structure of sunspots from ring diagram analysis. *J. Phys.* **CS-271**(1), 012006. doi:10.1088/1742-6596/271/1/012006.

Basu, S., Antia, H.M., Bogart, R.S.: 2004, Ring-diagram analysis of the structure of solar active regions. *Astrophys. J.* **610**, 1157 – 1168. doi:10.1086/421843.

Basu, S., Antia, H.M., Tripathy, S.C.: 1999, Ring diagram analysis of near-surface flows in the Sun. *Astrophys. J.* **512**, 458 – 470. doi:10.1086/306765.

Bogart, R.S., Basu, S., Rabello-Soares, M.C., Antia, H.M.: 2008, Probing the subsurface structures of active regions with ring-diagram analysis. *Solar Phys.* **251**, 439 – 451. doi:10.1007/s11207-008-9213-9.

Braun, D.C., Birch, A.C., Rempel, M., Duvall, T.L.: 2012, Helioseismology of a realistic magnetoconvective sunspot simulation. *Astrophys. J.* **744**, 77. doi:10.1088/0004-637X/744/1/77.

Gizon, L., Birch, A.C.: 2005, Local helioseismology. *Living Rev. Solar Phys.* **2**(1). http://www.livingreviews.org/lrsp-2005-6.

Gizon, L., Schunker, H., Baldner, C.S., Basu, S., Birch, A.C., Bogart, R.S., Braun, D.C., Cameron, R., Duvall, T.L., Hanasoge, S.M., Jackiewicz, J., Roth, M., Stahn, T., Thompson, M.J., Zharkov, S.: 2009, Helioseismology of sunspots: a case study of NOAA region 9787. *Space Sci. Rev.* **144**, 249 – 273. doi:10.1007/s11214-008-9466-5.

Hill, F.: 1988, Rings and trumpets – Three-dimensional power spectra of solar oscillations. *Astrophys. J.* **333**, 996 – 1013. doi:10.1086/166807.

Lin, C.-H., Basu, S., Li, L.: 2009, Interpreting helioseismic structure inversion results of solar active regions. *Solar Phys.* **257**, 37 – 60. doi:10.1007/s11207-009-9332-y.

Moradi, H., Baldner, C., Birch, A.C., Braun, D.C., Cameron, R.H., Duvall, T.L., Gizon, L., Haber, D., Hanasoge, S.M., Hindman, B.W., Jackiewicz, J., Khomenko, E., Komm, R., Rajaguru, P., Rempel, M., Roth, M., Schlichenmaier, R., Schunker, H., Spruit, H.C., Strassmeier, K.G., Thompson, M.J., Zharkov, S.: 2010, Modeling the subsurface structure of sunspots. *Solar Phys.* **267**, 1 – 62. doi:10.1007/s11207-010-9630-4.

Patron, J., Gonzalez Hernandez, I., Chou, D.-Y., Sun, M.-T., Mu, T.-M., Loudagh, S., Bala, B., Chou, Y.-P., Lin, C.-H., Huang, I.-J., Jimenez, A., Rabello-Soares, M.C., Ai, G., Wang, G.-P., Zirin, H., Marquette, W., Nenow, J., Ehgamberdiev, S., Khalikov, S. (TON Team): 1997, Comparison of two fitting methods for ring diagram analysis of very high L solar oscillations. *Astrophys. J.* **485**, 869. doi:10.1086/304469.

Pijpers, F.P., Thompson, M.J.: 1992, Faster formulations of the optimally localized averages method for helioseismic inversions. *Astron. Astrophys.* **262**, L33 – L36.

Pijpers, F.P., Thompson, M.J.: 1994, The SOLA method for helioseismic inversion. *Astron. Astrophys.* **281**, 231 – 240.

Rabello-Soares, M.C.: 2012, Solar-cycle variation of sound speed near the solar surface. *Astrophys. J.* **745**, 184. doi:10.1088/0004-637X/745/2/184.

Rabello-Soares, M.C., Basu, S., Christensen-Dalsgaard, J.: 1999, On the choice of parameters in solar-structure inversion. *Mon. Not. Roy. Astron. Soc.* **309**, 35 – 47. doi:10.1046/j.1365-8711.1999.02785.x.

Rajaguru, S.P., Basu, S., Antia, H.M.: 2001, Ring diagram analysis of the characteristics of solar oscillation modes in active regions. *Astrophys. J.* **563**, 410 – 418. doi:10.1086/323780.

DOI 10.1007/978-1-4899-8005-2_18
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0089-3

Horizontal Flows in the Photosphere and Subphotosphere of Two Active Regions

Yang Liu · Junwei Zhao · P.W. Schuck

Received: 17 April 2012 / Accepted: 25 July 2012 / Published online: 21 August 2012

Abstract We compare horizontal flow fields in the photosphere and in the subphotosphere (a layer 0.5 Mm below the photosphere) in two solar active regions: AR 11084 and AR 11158. AR 11084 is a mature, simple active region without significant flaring activity, and AR 11158 is a multipolar, complex active region with magnetic flux emerging during the period studied. Flows in the photosphere are derived by applying the Differential Affine Velocity Estimator for Vector Magnetograms (DAVE4VM) on HMI-observed vector magnetic fields, and the subphotospheric flows are inferred by time–distance helioseismology using HMI-observed Dopplergrams. Similar flow patterns are found for both layers for AR 11084: inward flows in the sunspot umbra and outward flows surrounding the sunspot. The boundary between the inward and outward flows, which is slightly different in the photosphere and the subphotosphere, is within the sunspot penumbra. The area having inward flows in the subphotosphere is larger than that in the photosphere. For AR 11158, flows in these two layers show great similarities in some areas and significant differences in other areas. Both layers exhibit consistent outward flows in the areas surrounding sunspots. On the other hand, most well-documented flux-emergence-related flow features seen in the photosphere do not have counterparts in the subphotosphere. This implies that the horizontal flows caused by flux emergence do not extend deeply into the subsurface.

Keywords Solar active regions, photosphere and subphotosphere · Solar active regions, flows

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

Y. Liu (✉) · J. Zhao
W.W. Hansen Experimental Physics Laboratory, Stanford University, Stanford, CA 94305-4085, USA
e-mail: yliu@sun.stanford.edu

P.W. Schuck
Solar Physics Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

1. Introduction

Plasma flows in the solar photosphere are of great importance in understanding the dynamics in the solar photosphere and in studying connections of photospheric flow fields and solar flaring activities. Different methods have been developed to derive flow fields in the photosphere. Local correlation tracking, initially developed by November and Simon (1988), has been widely used to study velocities inside supergranules and active regions. Different types of feature-tracking methods (see, *e.g.*, Hagenaar *et al.*, 1999) were also developed to study various aspects of the solar dynamics, including magnetic-flux fragmentation and merging, determination of statistical parameters of solar dynamos, and others. Comparisons among different feature-tracking methods were also carried out (DeForest *et al.*, 2007), and the results from these different methods were in reasonable agreement. Many authors also realized that the magnetic feature-tracking method might not be valid when magnetic-flux emergence or other complicated activities were occurring inside the studied active regions. Therefore, some new methods, with the magnetic induction equation taken into account, were developed to enhance the reliability of the inferred flow fields in such regions (Kusano *et al.*, 2002; Longcope, 2004; Welsch *et al.*, 2004; Georgoulis and LaBonte, 2006; Schuck, 2008).

Meanwhile, with the rapid development of helioseismology, mapping subsurface plasma flows has become possible for ring-diagram analysis (see, *e.g.*, Komm *et al.*, 2005) and for time–distance helioseismology (see, *e.g.*, Zhao *et al.*, 2012). In particular, time–distance helioseismology provides a high spatial resolution inference of subsurface flow fields, which has allowed studies of supergranules and active regions. In fact, the subsurface flow fields obtained from time–distance measurements have also been compared with the results derived from local correlation tracking for the quiet Sun (De Rosa, Duvall, and Toomre, 2000; Švanda, Zhao, and Kosovichev, 2007), and reasonable agreements were reported. However, a systematic comparison of photospheric and subphotospheric flows derived for active regions has not yet been carried out.

In this article, we compare photospheric flows and subphotospheric flows inside two selected active regions, and try to understand the similarities and discrepancies of these results. The photospheric flow field is obtained from the Differential Affine Velocity Estimator for Vector Magnetograms (DAVE4VM: the Schuck, 2008) using the data from the *Helioseismic and Magnetic Imager* (HMI)-observed vector magnetograms. The subphotospheric flow field is from the HMI time–distance analysis pipeline (Zhao *et al.*, 2012) using the HMI Doppler velocity, but only results from the shallowest depth, *i.e.* a depth range of $0-1$ Mm (hereafter, -0.5 Mm for simplicity), are used. It is very interesting to see how the results from these two very different analysis techniques using different data inputs agree or differ. Moreover, the differences in the flow fields between different depths can help us understand the rapidly developing complex active regions and flaring activities inside these regions. We introduce the procedure of data reduction in Section 2, and present our results in Section 3, together with a brief discussion. Our conclusions are given in Section 4.

2. Data Reduction

We use Doppler velocity and vector magnetic-field data taken by HMI (Scherrer *et al.*, 2012; Schou *et al.*, 2012) to derive the subsurface flow field and photospheric flow field, respectively. The HMI instrument is a filtergraph with a full-disk coverage of 4096×4096 pixels. The spatial resolution is about $1''$ with a $0.5''$ pixel size. The spectral line is the Fe I 6173 Å

absorption line formed in the photosphere (Norton *et al.*, 2006). There are two CCD cameras in the instrument, the "front camera" and the "side camera." The front camera acquires filtergrams at six wavelengths across the Fe I 6173 Å line in two polarization states with 3.75 seconds between the images. It takes 45 seconds to acquire a set of 12 filtergrams. This set of data is used to derive Dopplergrams and line-of-sight magnetograms (Couvidat *et al.*, 2012). The side camera is dedicated to measuring the vector magnetic field. It takes 135 seconds to obtain the filtergrams in six polarization states at six wavelength positions. The Stokes parameters [*I, Q, U, V*] are then computed from those measurements. In order to suppress the p-mode oscillations and increase the signal-to-noise ratio, the Stokes parameters are usually derived from the filtergrams averaged over a certain period of time. Currently an average is computed over 720 seconds. This average requires extra filtergrams before and after the nominal 720-second temporal window because the filtergrams must be interpolated onto a regular and finer grid before averaging. The averaging uses a cosine-apodized boxcar with a full width half maximum (FWHM) of 720 seconds. The tapered temporal window used is actually 1215 seconds long. The Stokes parameters are then inverted to produce the vector magnetic-field using the Very Fast Inversion of the Stokes Vector (VFISV) inversion algorithm, a Milne–Eddington (ME)-based approach (Borrero *et al.*, 2011). The 180° ambiguity of the azimuth is solved based on the "minimum energy" algorithm (Metcalf, 1994; Metcalf *et al.*, 2006; Leka *et al.*, 2009). Finally, the disambiguated vector magnetic-field data of active regions are deprojected into the heliographic coordinates.

DAVE4VM (Schuck, 2008) is used to derive the vector velocity field in the photosphere from HMI time series deprojected, registered vector magnetic-field data. The cadence of the data is 720 seconds. The window size used in DAVE4VM is 19 pixels, which is determined by examining the slope, Pearson linear correlation coefficient, and Spearman rank order between $\bigtriangledown_h \cdot (v_z \mathbf{B}_h - \mathbf{v}_h B_z)$ and $\delta B_z / \delta t$, where v_z and v_h are vertical and horizontal velocities, and B_z and $\mathbf{B}_h$ are vertical and horizontal magnetic fields, as suggested in Schuck (2008).

The subphotospheric flow fields are obtained from the HMI time–distance data analysis pipeline, and all analysis details of the technique and an example of subphotospheric flow field of one active region can be found in Zhao *et al.* (2012). The data input for such a pipeline analysis is HMI-observed Dopplergrams with a sampling rate of 0.12° pixel^{-1}. Although the pipeline provides subsurface flow fields down to 20 Mm in depth, here we only use the shallowest depth, *i.e.* -0.5 Mm, for a direct comparison with the photospheric flow fields.

We select two active regions to carry out this comparison. One is a mature yet simple active region, AR 11084; the other is an emerging, multipolar, and complicated active region, AR 11158. The time–distance helioseismology method requires eight-hour long Doppler velocities to derive the horizontal velocity field in the subsurface; thus, the velocity is an average during that eight-hour time interval. For AR 11084, we use data during the period of 2 July 2010 00:00 UT to 08:00 UT, and for AR 11158 the data used cover the period from 14 February 2011 06:00 UT to 14:00 UT. For the velocity in the photosphere, DAVE4VM can provide the velocity field at a cadence of 720 seconds, the same as that for HMI vector magnetic-field data. For comparison, we average eight derived photospheric velocity fields during that eight-hour time interval, one per hour, to obtain an eight-hour averaged photospheric velocity field. This velocity field is used for comparison with the velocity obtained for the -0.5 Mm layer.

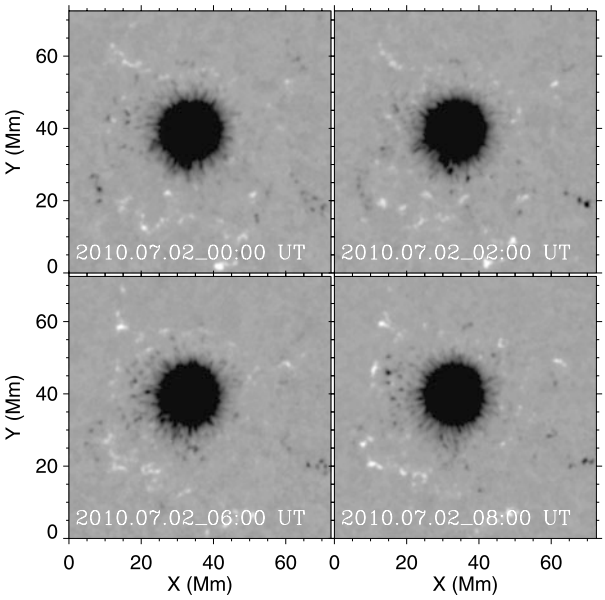


Figure 1 Vertical magnetic field in active region AR 11084 at 00:00 UT, 02:00 UT, 06:00 UT, and 08:00 UT of 2 July 2010. Black and white refer to negative and positive fields, respectively.

3. Results and Discussion

3.1. AR 11084

AR 11084 is a mature simple active region with a stable and relatively round sunspot located inside it. Figure 1 shows the evolution of the magnetic field in AR 11084 in the eight hours during which the study is carried out; no significant evolution is seen during this eight-hour interval for the sunspot.

Figure 2 shows horizontal flows in this active region plotted on a continuum intensity image. The arrows represent the flows in the photosphere (green) and in the -0.5 Mm layer (red). Only velocities where the horizontal speed in the photosphere is greater than 0.07 km s^{-1} are plotted. Note that the scales are different in the two maps in order to better display the results. In general, both maps show very similar flow patterns: inward flows in the sunspot umbra and outward flows in the areas surrounding the sunspot. The inward- and outward-flow pattern in mature sunspots in the photosphere was previously reported by, *e.g.*, Schröter (1962), Muller (1973), and Wang and Zirin (1992). Separation of the inward and outward flows takes place in the sunspot penumbra. But the boundaries between the two opposite-flow directions are different in these two layers. The area having inward flows at the -0.5 Mm layer is larger than that in the photosphere. In other words, the flows in the penumbra that have already become outflows in the photosphere are still inward flows at the -0.5 Mm layer. This inward flow may play a role in containing the sunspot (Zhao, Kosovichev, and Sekii, 2010).

To better illustrate the difference of the flows in these two layers, we mark with asterisks in Figure 3 the locations where the angle between the two flows is greater than 90°. This shows that the major differences between these two sets of results are inside the sunspot penumbra and other areas far away from the sunspot, where the method DAVE4VM becomes insensitive due to weak magnetic field strength.

While there are a few areas that show different flows in the two layers, most of the areas actually show very similar flow patterns in both layers. Figure 4 shows scatter plots of the

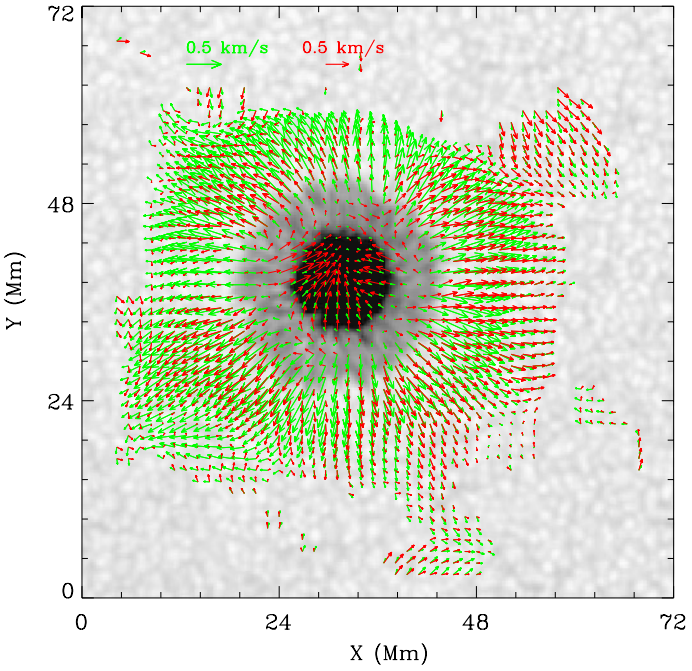


Figure 2 Horizontal flows in AR 11084 in the photosphere (green arrows) and in the −0.5 Mm layer (red arrows) overplotted on a continuum intensity image taken at 2 July 2010 04:00 UT. The photospheric flow is an average of eight velocity maps at a cadence of one hour from 00:00 UT to 08:00 UT derived by DAVE4VM. The flow in the −0.5 Mm layer is derived by a time–distance helioseismology method applied to eight-hour Dopplergrams observed from 2 July 2010 00:00 UT to 08:00 UT. Only velocities at locations where the horizontal velocity in the photosphere is greater than 0.07 km s^{-1} are plotted.

velocities in the two layers. In the top panels are scatter plots of v_x (right) and v_y (left). The magnitude ($|v| = \sqrt{v_x^2 + v_y^2}$) and azimuthal angle [$\theta$] of the horizontal velocity are plotted in the bottom panels. The azimuthal angle ranges from −180° to 180°. The horizontal axis represents velocity in the −0.5 Mm layer from the time–distance helioseismology technique, and the vertical axis velocity in the photosphere from the DAVE4VM method. The Pearson correlation coefficient is 0.89 for v_x, 0.76 for v_y, 0.48 for $|v|$, and 0.64 for θ. v_x has a better correlation than v_y, and this was also reported by Švanda, Zhao, and Kosovichev (2007) for the quiet-Sun study. Poor correlation in $|v|$ implies a gradient between the two layers. We also calculate the vector correlation coefficient and the Cauchy–Schwarz inequality (Schrijver *et al.*, 2006). The vector correlation coefficient [$C_{\rm vec}$] is defined as

$$C_{\rm vec} = \frac{\sum_i \mathbf{v}_i \cdot \mathbf{u}_i}{(\sum_i \mathbf{v}_i^2 \sum_i \mathbf{u}_i^2)^{1/2}}, \tag{1}$$

where $\mathbf{v}_i$ and $\mathbf{u}_i$ are velocities in two layers at pixel i. The Cauchy–Schwarz inequality [$C_{\rm cs}$] is defined as

$$C_{\rm cs} = \frac{1}{M} \sum_i \frac{\mathbf{v}_i \cdot \mathbf{u}_i}{|\mathbf{v}_i||\mathbf{u}_i|}, \tag{2}$$

where M is the total number of pixels in the region studied. Here we use only two components of the vector velocity field, *i.e.* the horizontal velocity, to compute the coefficients.

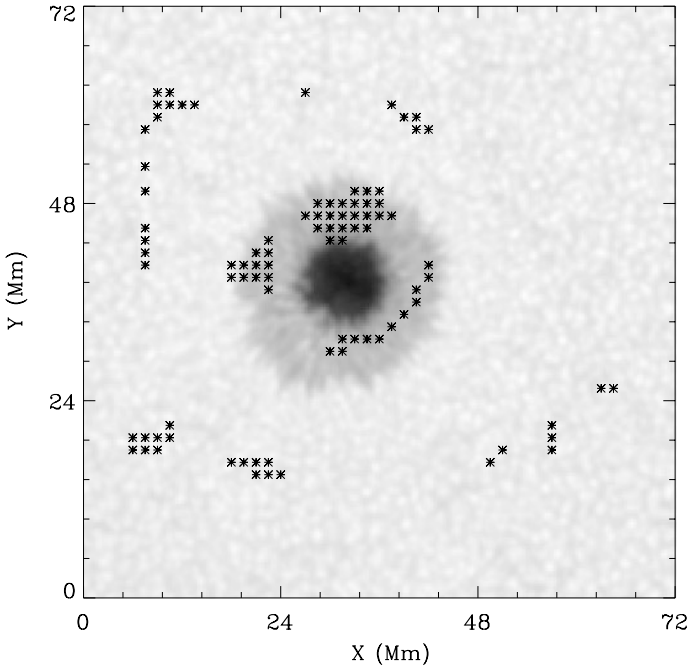


Figure 3 Background image is the 2 July 2010 04:00 UT continuum intensity data. Asterisks correspond to the areas where the angle between the two flows in the photosphere and the −0.5 Mm layer is greater than 90°.

The vector correlation coefficient is 0.80, and the Cauchy–Schwarz inequality is 0.76. This indicates that the flows in the two layers are very similar.

3.2. AR 11158

AR 11158 is a multipolar, complex active region, producing several major flares during its disk passage (see, *e.g.*, Sun *et al.*, 2012). The flows studied here in the photosphere are an average of eight one-hour cadence flow maps derived by DAVE4VM for the period of 14 February 2011 06:00 UT to 14:00 UT, during which magnetic flux was emerging in this region. This is demonstrated by the temporal profiles of magnetic-flux evolution in the region (see Figure 5). The two vertical dotted lines mark the time interval during which the flows are calculated. Note also that the positive and negative fluxes are well balanced. Figure 6 shows the vertical magnetic field in AR 11158 in this eight-hour time interval. The negative field patches *N1* and *N2* (see the bottom right panel) rotate counterclockwise, while the positive field patch *P1* (leading polarity) undergoes separation, moving toward the Northwest. Shear motion along the polarity inversion line between *N1* and *P2* is also clearly seen. These flow patterns are confirmed by the photospheric flow map by DAVE4VM (see Figure 7).

The flows in AR 11158 in the photosphere and at the −0.5 Mm layer are compared in Figure 8. Outward flows are clearly seen in the areas surrounding the sunspots in both layers, as also seen in AR 11084. However, the well-documented flux-emergence-related surface flows in the photosphere (see, *e.g.*, Brown *et al.*, 2003, Zhang, Li, and Song, 2007, Schmieder *et al.*, 1994, Deng *et al.*, 2006) – the rotation of negative field patches *N1* and

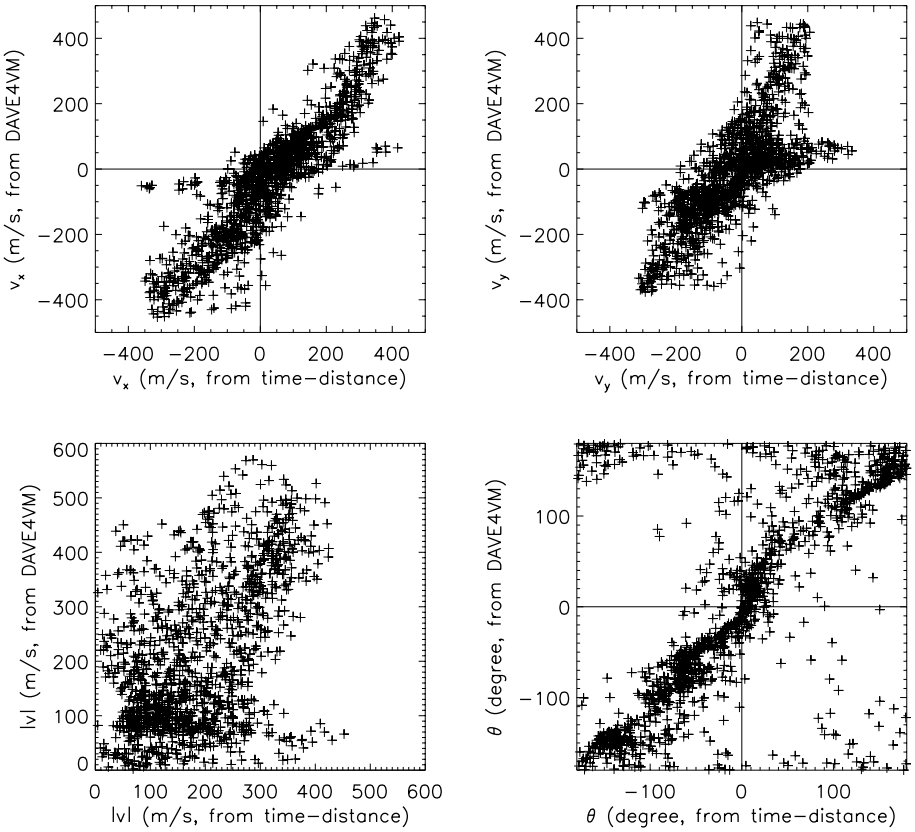


Figure 4 Scatter plots between the velocities of two layers for v_x (top left), v_y (top right), their magnitudes ($|v|$; bottom left), and azimuthal angles (θ; bottom right). The horizontal axis represents velocity in the -0.5 Mm layer from the time–distance helioseismology method, the vertical axis velocity in the photosphere from the DAVE4VM. The Pearson correlation coefficient is 0.89 for v_x, 0.76 for v_y, 0.48 for $|v|$, and 0.64 for θ.

N2, the separation motion of positive field patch *P1*, and the shear motion along the polarity inversion line between *P2* and *N1* – do not have counterparts in the -0.5 Mm layer. Instead, significant inward flows in the sunspots are seen (*e.g.* in *N1*, *N2*, and *P1*). Figure 9 shows the flow fields in these patches in detail. Figure 10 is another way to show the difference of the flows in the two layers: the asterisks mark areas where the angle between the two flows is greater than 90°. A significant difference occurs in the rotational sunspots, *N1* and *N2*, and in the separation-motion sunspot *P1*, indicating that the surface flows in the photosphere that are related to flux emergence do not extend very deeply into the subsurface. Those flows are probably caused by the magnetic field, which is emerging and expanding into the solar corona. Gradients in magnetic field due to its rapid emergence and expansion drive motions in the photosphere, as demonstrated in magnetohydrodynamic (MHD) simulations (see, *e.g.*, Magara and Longcope, 2003; Manchester *et al.*, 2004; Fan, 2009). This has less effect in the subphotosphere.

The scatter plots of velocities in two layers in AR 11158 are presented in Figure 11. As expected, the correlation coefficient is lower than that for AR 11084. Once again, the v_x has a better correlation than v_y.

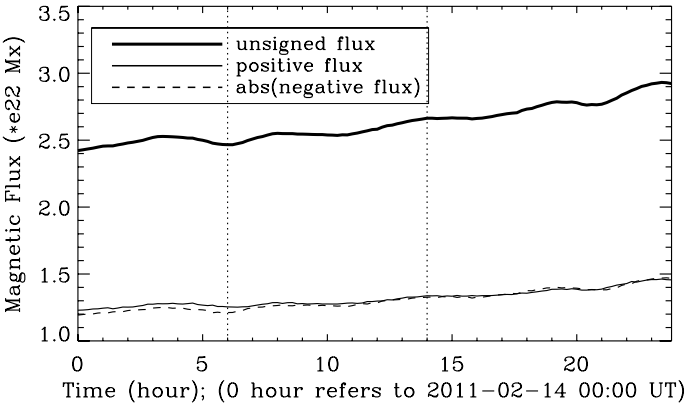


Figure 5 Temporal profiles of magnetic flux in AR 11158 on 14 February 2011. The thick solid line represents unsigned flux, while the solid and dashed lines are positive flux and absolute value of negative flux. Vertical magnetic fields greater than 100 Gauss, the noise level of the magnetic field strength, are used to measure the fluxes. Two vertical dotted lines mark the time interval from 06:00 UT to 14:00 UT in which the flows are analyzed here.

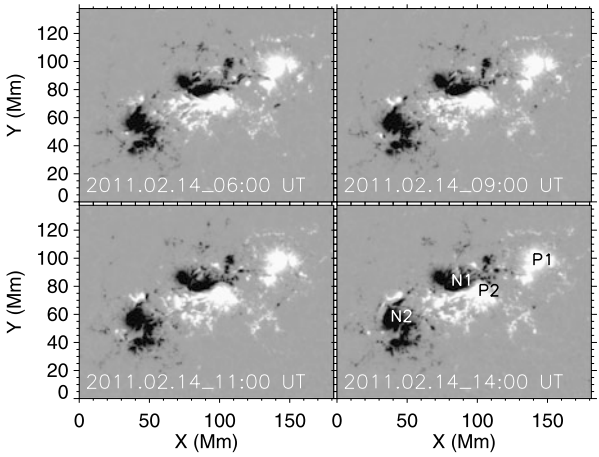


Figure 6 Evolution of magnetic field in AR 11158 during 06:00 UT – 14:00 UT 14 February 2011. White and black represent positive and negative vertical magnetic fields, respectively. *P1* and *P2* in the bottom right panel denote positive polarity patches; *N1* and *N2* denote negative field patches.

3.3. Discussion

For both active regions, we also extend the subsurface flows, obtained from the time–distance data analysis pipeline, down to −6 Mm in depth to compare with the −0.5 Mm flow fields. For depths down to −3 Mm, the flow patterns look quite similar to that at −0.5 Mm, indicating that similar flow structures extend to a depth of −3 Mm or so. Below −3 Mm, the outflow structure beyond the sunspot penumbra remains largely unchanged, but the inflow structure inside the penumbra has changed to outflows. The depth variation of subsurface flow fields obtained from time–distance analysis has been discussed by various authors (*e.g.* Zhao, Kosovichev, and Sekii, 2010). These large-scale outflows in sunspots have been reproduced by MHD numerical simulations. For example, Rempel (2011) showed that the large-scale outflows, extending from the photosphere to the deep interior, were associated with the sunspot penumbra. His simulation also showed converging flows in a sunspot without

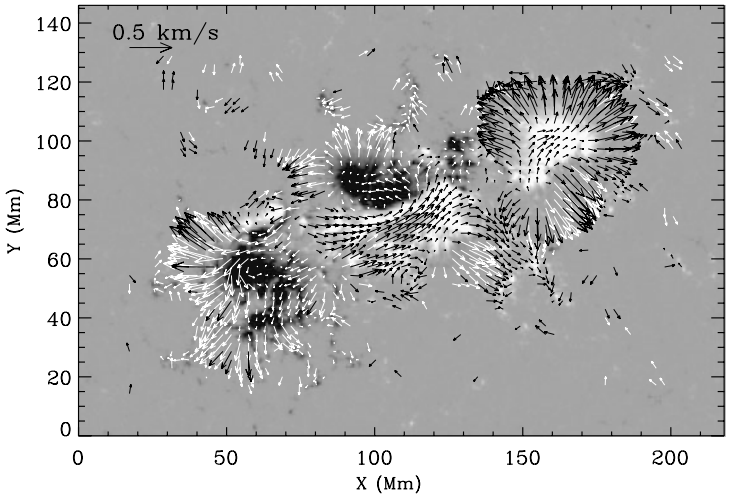


Figure 7 Horizontal velocity in the photosphere in AR 11158 derived by DAVE4VM. The image shows the vertical magnetic field with the positive field in white and negative field in black. The arrows represent horizontal velocity. The image is an average of eight velocity maps at a cadence of one hour from 06:00 UT – 14:00 UT 14 February 2011. Only velocities greater than 0.07 km s^{-1} are plotted. Black (white) arrows indicate that the vertical magnetic fields in the pixels are positive (negative).

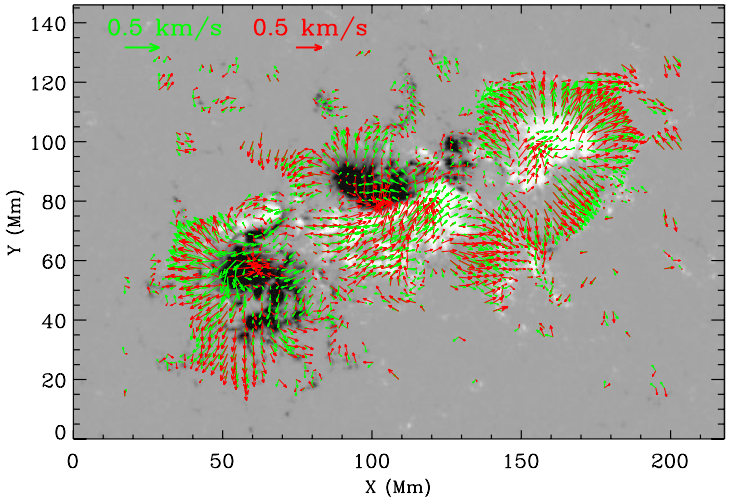


Figure 8 Vertical magnetic field in AR 11158 (image) overplotted by the horizontal velocities in the photosphere (green arrows) and in the −0.5 Mm layer (red arrows). Black and white in the image refer to negative and positive fields, respectively. The photospheric velocity is the same as in Figure 7. The flow in the −0.5 Mm layer is derived by a time–distance helioseismology method applied to eight-hour Dopplergrams observed from 06:00 UT – 14:00 UT 14 February 2011. Only velocities at locations where the horizontal velocity in the photosphere is greater than 0.07 km s^{-1} are plotted.

a penumbra. However, for a sunspot with a well-developed penumbra, no obvious converging flows were found in that simulation, which is not consistent with our results here, *i.e.*

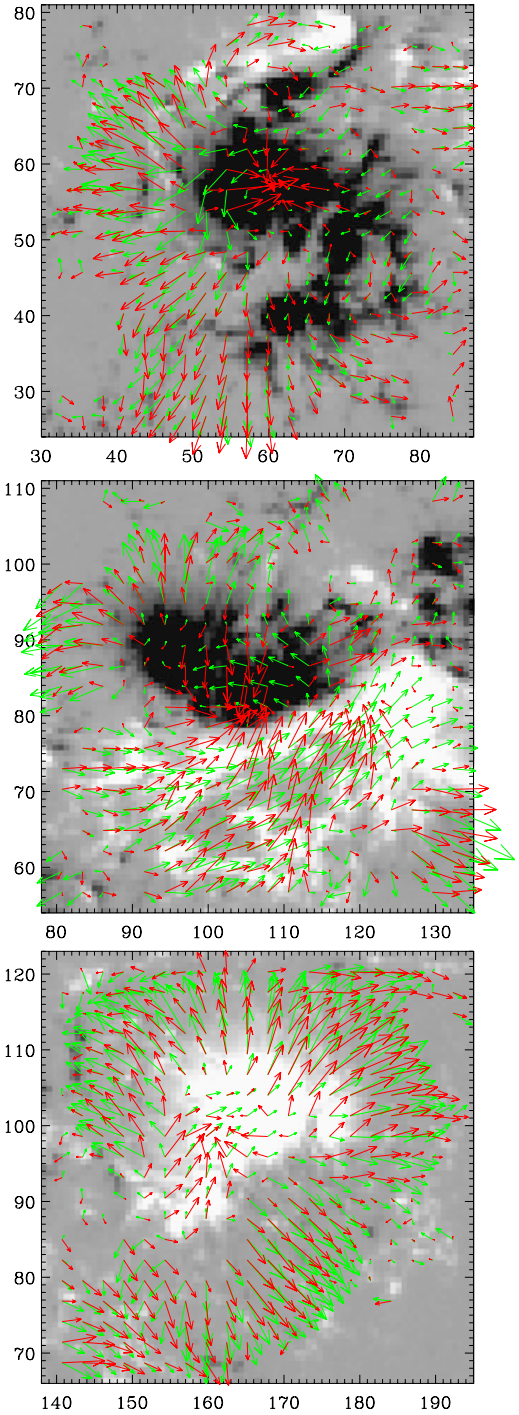


Figure 9 Enlarged sub-areas of Figure 8 to show details in the three interesting regions that include patch *N2* (top), patches *N1* and *P2* (middle), and patch *P1* (bottom) marked in Figure 6. Both vertical and horizontal axes are in megameters.

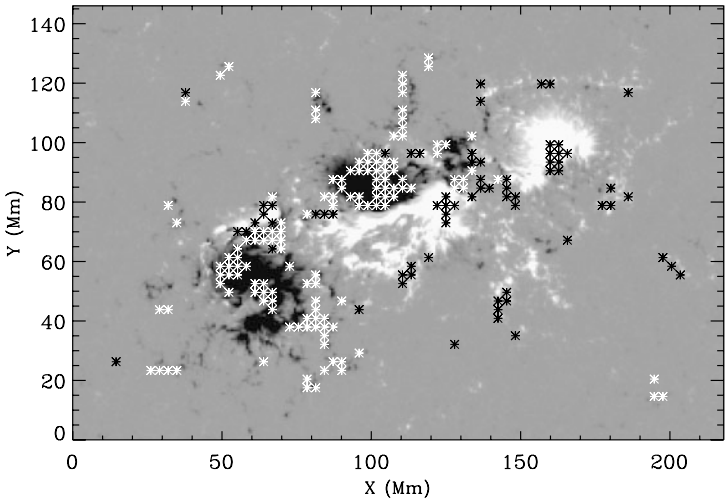


Figure 10 Background image is the vertical magnetic field in AR 11158. Asterisks represent the areas where the angle between the horizontal velocities in the photosphere and the −0.5 Mm layer is greater than 90°. White (black) asterisks correspond to the areas where the vertical field is negative (positive).

that the converging flows are in the umbra and part of the penumbra for both layers. This discrepancy deserves more study.

We also compared the vertical flows derived from both techniques, but found poor correlation between the two. The photospheric vertical flows in sunspots derived from DAVE4VM are dominated by upward flows, consistent with the strong outflows in the penumbra known as the Evershed effect. However, the subsurface vertical flows derived from time–distance analysis are dominated by downward flows, as discussed by Hindman, Haber, and Toomre (2009) and Zhao, Kosovichev, and Sekii (2010).

4. Conclusions

We compare flows in the photosphere and in the −0.5 Mm layer in two active regions. One is a mature simple active region; the other is an emerging, complex active region. We found that for the mature simple active region the flows in both layers show very similar patterns: inward flow in the sunspot umbra and outward flow in the areas surrounding the sunspot. The boundary separating these two types of flows occurs in the sunspot penumbra, but the location of separation is slightly different in the two layers. The inward-flow area in the sunspot is larger in the −0.5 Mm layer than that in the photosphere. In other words, the flows in these areas where there are inward flows in the −0.5 Mm layer have become outward flow in the photosphere. Inward flow in the sunspots in the subsurface is suggested to play a role in containing the sunspots (Zhao, Kosovichev, and Sekii, 2010).

For the emerging, complex active region, the flows in the two layers show both great similarities and significant differences. Although both layers show similar outward flows in the areas surrounding the sunspots, the well-documented flux-emergence-related surface flows seen in the photosphere, such as separation motion of leading and following polarity patches, fast rotation of sunspots, and apparent shear motion along the polarity inversion lines, do not have counterparts in the −0.5 Mm layer. This implies that the cause of these

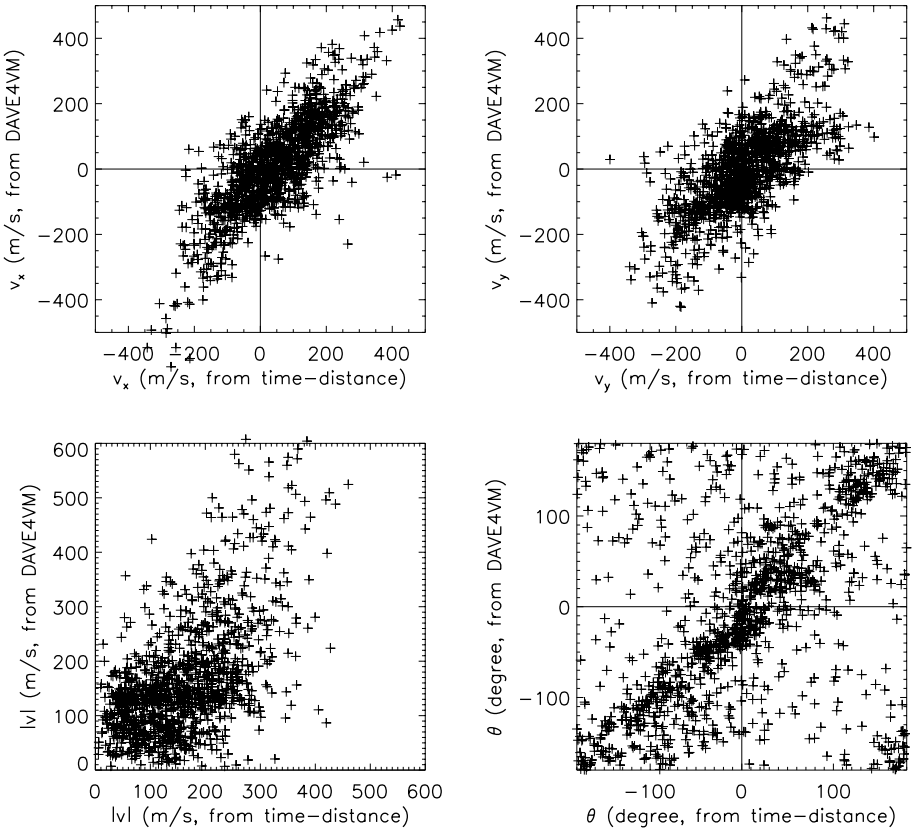


Figure 11 Scatter plots between the velocities of two layers for v_x (top left), v_y (top right), their magnitudes ($|v|$; bottom left), and azimuthal angles (θ; bottom right), for AR 11158. The horizontal axis represents velocity in the -0.5 Mm layer from the time–distance helioseismology method, while the vertical axis represents velocity in the photosphere from the DAVE4VM. The Pearson correlation coefficient is 0.78 for v_x, 0.70 for v_y, 0.56 for $|v|$, and 0.43 for θ. The vector correlation coefficient for the two components (horizontal velocity) is 0.73, and the Cauchy–Schwarz inequality is 0.59.

flows, which is probably the gradient of the magnetic field due to its rapid emergence and expansion into the corona, has less effect in the subsurface.

Acknowledgements The authors wish to thank the anonymous referee for the valuable comments and suggestions. This work was supported by NASA Contract NAS5-02139 (HMI) to Stanford University. The data are used courtesy of NASA/SDO and the HMI science team. SOHO is a project of international cooperation between ESA and NASA.

References

Borrero, J.M., Tomczyk, S., Kubo, M., Socas-Navarro, H., Schou, J., Couvidat, S., Bogart, R.: 2011, *Solar Phys.* **273**, 267. ADS: 2011SoPh..273..267B, doi:10.1007/s11207-010-9515-6.

Brown, D.S., Nightingale, R.W., Alexander, D., Schrijver, C.J., Metcalf, T.R., Shine, R.A., Title, A.M., Wolfson, C.J.: 2003, *Solar Phys.* **216**, 79. ADS: 2003SoPh..216...79B, doi:10.1023/A:1026138413791.

Couvidat, S., Rajaguru, S.P., Wachter, R., Sankarasubramanian, K., Schou, J., Scherrer, P.H.: 2012, *Solar Phys.* **278**, 217. ADS: 2012SoPh..278..217C, doi:10.1007/s11207-011-9927-y.

De Rosa, M., Duvall, T.L. Jr., Toomre, J.: 2000, *Solar Phys.* **192**, 351. ADS: 2000SoPh..192..351R, doi:10.1023/A:1005269001739.
DeForest, C.E., Hagenaar, H.J., Lamb, D.A., Parnell, C.E., Welsch, B.T.: 2007, *Astrophys. J.* **666**, 576.
Deng, N., Xu, Y., Yang, G., Cao, W., Liu, C., Rimmele, T.R., Wang, H., Denker, C.: 2006, *Astrophys. J.* **644**, 1278.
Fan, Y.: 2009, *Astrophys. J.* **697**, 1529.
Georgoulis, M.K., LaBonte, B.J.: 2006, *Astrophys. J.* **636**, 475.
Hagenaar, H.J., Schrijver, C.J., Title, A.M., Shine, R.A.: 1999, *Astrophys. J.* **511**, 932.
Hindman, B.W., Haber, D.A., Toomre, J.: 2009, *Astrophys. J.* **698**, 1749.
Komm, R., Howe, R., Hill, F., González Hernández, I., Toner, C., Corbard, T.: 2005, *Astrophys. J.* **631**, 636.
Kusano, K., Maeshiro, T., Yokoyama, T., Sakurai, T.: 2002, *Astrophys. J.* **577**, 501.
Leka, K.D., Barnes, G., Crouch, A.D., Metcalf, T.R., Gary, G.A., Jing, J., Liu, Y.: 2009, *Solar Phys.* **260**, 83. ADS: 2009SoPh..260...83L, doi:10.1007/s11207-009-9440-8.
Longcope, D.W.: 2004, *Astrophys. J.* **612**, 1181.
Magara, T., Longcope, D.W.: 2003, *Astrophys. J.* **586**, 630.
Manchester, W. IV, Gombosi, T., DeZeeuw, D., Fan, Y.: 2004, *Astrophys. J.* **610**, 588.
Metcalf, T.R.: 1994, *Solar Phys.* **155**, 235. ADS: 1994SoPh..155..235M, doi:10.1007/BF00680593.
Metcalf, T.R., Leka, K.D., Barnes, G., Lites, B.W., Georgoulis, M.K., Pevtsov, A.A., Balasubramaniam, K.S., Gary, G.A., Jing, J., Li, J., Liu, Y., Wang, H.N., Abramenko, V., Yurchyshyn, V., Moon, Y.-J.: 2006, *Solar Phys.* **237**, 267. ADS: 2006SoPh..237..267M, doi:10.1007/s11207-006-0170-x.
Muller, R.: 1973, *Solar Phys.* **29**, 55. ADS: 1973SoPh...29...55M, doi:10.1007/BF00153440.
Norton, A.A., Graham, J.P., Ulrich, R.K., Schou, J., Tomczyk, S., Liu, Y., Lites, B.W., López Ariste, A., Bush, R.I., Socas-Navarro, H., Scherrer, P.H.: 2006, *Solar Phys.* **239**, 69. ADS: 2006SoPh..239...69N, doi:10.1007/s11207-006-0279-y.
November, L.J., Simon, G.W.: 1988, *Astrophys. J.* **333**, 427.
Rempel, M.: 2011, *Astrophys. J.* **740**, 15.
Schuck, P.W.: 2008, *Astrophys. J.* **683**, 1134.
Scherrer, P.H., Schou, J., Bush, R.I., Kosovichev, A.G., Bogart, R.S., Hoeksema, J.T., Liu, Y., Duvall, T.L. Jr., Zhao, J., Title, A.M., Schrijver, C.J., Tarbell, T.D., Tomczyk, S.: 2012, *Solar Phys.* **275**, 207. ADS: 2012SoPh..275..207S, doi:10.1007/s11207-011-9834-2.
Schmieder, B., Hagyard, M.J., Ai, G., Zhang, H., Kalman, B., Gyori, L., Rompolt, B., Demoulin, P., Machado, M.E.: 1994, *Solar Phys.* **150**, 199. ADS: 1994SoPh..150..199S, doi:10.1007/BF00712886.
Schou, J., Scherrer, P.H., Bush, R.I., Wachter, R., Couvidat, S., Rabello-Soares, M.C., Bogart, R.S., Hoeksema, J.T., Liu, Y., Duvall, T.L.Jr., Akin, D.J., Allard, B.A., Miles, J.W., Rairden, R., Shine, R.A., Tarbell, T.D., Title, A.M., Wolfson, C.J., Elmore, D.F., Norton, A.A., Tomczyk, S.: 2012, *Solar Phys.* **275**, 229. ADS: 2012SoPh..275..229S, doi:10.1007/s11207-011-9842-2.
Schrijver, C.J., De Rosa, M.L., Metcalf, T.R., Liu, Y., McTiernan, J., Régnier, S., Valori, G., Wheatland, M.S., Wiegelmann, T.: 2006, *Solar Phys.* **235**, 161. ADS: 2006SoPh..235..161S, doi:10.1007/s11207-006-0068-7.
Schröter, E.H.: 1962, *Z. Astrophys.* **56**, 183.
Sun, X., Hoeksema, J.T., Liu, Y., Wiegelmann, T., Hayashi, K., Chen, Q., Thalmann, J.: 2012, *Astrophys. J.* **748**, 77.
Švanda, M., Zhao, J., Kosovichev, A.G.: 2007, *Solar Phys.* **241**, 27. ADS: 2007SoPh..241...27, doi:10.1007/s11207-007-0333-4.
Welsch, B.T., Fisher, G.H., Abbett, W.P., Regnier, S.: 2004, *Astrophys. J.* **610**, 1148.
Wang, H., Zirin, H.: 1992, *Solar Phys.* **140**, 41. ADS: 1992SoPh..140...41W, doi:10.1007/BF00148428.
Zhang, J., Li, L., Song, Q.: 2007, *Astrophys. J. Lett.* **662**, L35.
Zhao, J., Kosovichev, A.G., Sekii, T.: 2010, *Astrophys. J.* **708**, 304.
Zhao, J., Couvidat, S., Bogart, R.S., Parchevsky, K.V., Birch, A.C., Duvall, T.L. Jr., Beck, J.G., Kosovichev, A.G., Scherrer, P.H.: 2012, *Solar Phys.* **275**, 375. ADS: 2012SoPh..275..375Z, doi:10.1007/s11207-011-9757-y.

DOI 10.1007/978-1-4899-8005-2_19
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0031-8

Active Region Formation through the Negative Effective Magnetic Pressure Instability

Koen Kemel · Axel Brandenburg · Nathan Kleeorin · Dhrubaditya Mitra · Igor Rogachevskii

Received: 2 March 2012 / Accepted: 9 May 2012 / Published online: 10 July 2012

Abstract The negative effective magnetic-pressure instability operates on scales encompassing many turbulent eddies, which correspond to convection cells in the Sun. This instability is discussed here in connection with the formation of active regions near the surface layers of the Sun. This instability is related to the negative contribution of turbulence to the mean magnetic pressure that causes the formation of large-scale magnetic structures. For an isothermal layer, direct numerical simulations and mean-field simulations of this phenomenon are shown to agree in many details, for example the onset of the instability occurs at the same depth. This depth increases with increasing field strength, such that the growth rate of this instability is independent of the field strength, provided the magnetic structures are fully contained within the domain. A linear stability analysis is shown to support this finding. The instability also leads to a redistribution of turbulent intensity and gas pressure that could provide direct observational signatures.

Keywords Magnetohydrodynamics (MHD) · Sun: dynamo · Sunspots · Turbulence

1. Introduction

Active-region formation in the Sun is traditionally thought to be a deeply rooted phenomenon, because their size ($\approx 100\,\mathrm{Mm}$) is much larger than the naturally occurring scales

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

K. Kemel (✉) · A. Brandenburg · N. Kleeorin · D. Mitra · I. Rogachevskii
Nordita, Royal Institute of Technology and Stockholm University, Roslagstullsbacken 23, 10691 Stockholm, Sweden
e-mail: brandenb@nordita.org

K. Kemel · A. Brandenburg
Department of Astronomy, Stockholm University, 10691 Stockholm, Sweden

N. Kleeorin · I. Rogachevskii
Department of Mechanical Engineering, Ben-Gurion University of the Negev, POB 653, Beer-Sheva 84105, Israel

in the surface layers of the convection zone ($\approx 1-10$ Mm); see Golub *et al.* (1981). They are also long-lived (several months), which seems unnaturally long if associated with the near-surface layers (40 Mm depth), where typical time scales are about a day. This is particularly true of what are called *complexes of activity*, which can live for $\approx 1/2$ year (Golub and Vaiana, 1980). At the bottom of the convection zone, or even beneath it in the tachocline, the typical timescales are long ($\approx$ one month) and it might then, at these depths, be easier to envisage mechanisms for producing recurrent eruptions including sunspots over timescales exceeding $1/2$ year. A detailed report of recurrent sunspot activity between 23 August 1903 and 14 August 1904 was published by Epstein (1904), who referred to that region as a *Fleckenherd*, which can be translated as sunspot hearth. Later, Sanford (1941) referred to this as an *active region* and gave details about an event lasting 620 days from 20 August 1929 to 2 May 1931. In Epstein's early article, which was written in German,[1] he wrote that *a Fleckenherd is a region on the solar surface where sunspots* (*both larger and smaller ones, of shorter and longer duration*) *have occurred in a confined area for at least eight rotation periods.* Such regions on the Sun would nowadays be referred to as complexes of activity, while those discussed by Sanford (1941) would perhaps be called superactive regions or active zones (Bai, 1987, 1988), and might also be related to active longitudes (Vitinskij, 1969; Bogart, 1982).

All of these phenomena, from active regions to active longitudes, are probably caused by magnetic-flux enhancements of some sort. For the purpose of this article, we simply refer to them as active regions. Instabilities in the tachocline are generally held responsible for their formation such as the clamshell and tipping instabilities (Cally, Dikpati, and Gilman, 2003), which are global instabilities of a magnetic belt around the Sun. Based on calculations using the thin-fluxtube approximation, the field strength of such a magnetic belt is expected to be around 10^5 G (Choudhuri and Gilman, 1987; D'Silva and Choudhuri, 1993; Schüssler *et al.*, 1994); see Fan (2009) and Charbonneau (2010) for recent reviews on the subject. However, such strong fields may be unstable (Tayler, 1973); see Arlt, Sule, and Rüdiger (2007), who found 100 G to be the limit. A completely different idea is to invoke a coupling to non-axisymmetric dynamo modes (Ruzmaikin, 1998; Bigazzi and Ruzmaikin, 2004), but the field is still thought to reside near the bottom of the convection zone.

A general problem with deeply rooted active regions as a source of sunspots is the difficulty of keeping a buoyant flux tube intact, so that it can pierce through the surface to form a well-confined bipolar region. Another difficulty concerns the angular velocity of active regions. At the beginning of the solar cycle, at 30° latitude, they have a rotation rate of ≈ 446 nHz, matching the local angular velocity at a radius of $\approx 0.95 R_\odot$, but exceeding the value at the bottom of the convection zone by ≈ 10 nHz. Likewise, at the end of the cycle, at 4° latitude, they have a rotation rate of ≈ 462 nHz, matching the local angular velocity near the surface at a radius of $\approx 0.97 R_\odot$, again exceeding the value at the bottom of the convection zone by ≈ 10 nHz; see Figure 4 of Benevolenskaya *et al.* (1999) and Figure 2 of Brandenburg (2005). Other difficulties concern the high field strength in the tachocline required by thin-fluxtube calculations to explain the observed tilt angles (D'Silva and Choudhuri, 1993). On the other hand, Kosovichev and Stenflo (2008) pointed out that there is no

[1] We gave here a rather free translation of his original and somewhat antique formulation. For historical reasons, and for the benefit of those with restricted access to the original issue of *Astronomische Nachrichten*, we reproduce here his original definition: *Ein Fleckenherd auf der Sonne. So kann man füglich eine Gegend auf der Sonnenoberfläche bezeichnen, wo in einem verhältnismäßig beschränkten Bezirke im Verlaufe eine Jahres in mindestens acht Rotationen Flecke, teils größere, teils kleinere, teils von kürzerem, teils von längerem Bestande, aufgetreten sind. . . .* [*Th. Epstein, Frankfurt am Main, Oktober 1904*]

evidence for a dependence of the tilt angle on magnetic-field strength, as expected based on thin-fluxtube calculations. Furthermore, Stenflo and Kosovichev (2012) find no indication of a dependence of the tilt angle on the size of the region, which one may expect if the tilts were produced by the Coriolis force during the buoyant rise of flux loops from the tachocline. Several of the arguments discussed above have led to the consideration of solar activity as a shallow phenomenon; see Brandenburg (2005) for details. Even sunspots themselves suggest their being rather shallow (Zhao, Kosovichev, and Duvall, 2001; Kosovichev, 2002). This raises the question whether there is then a mechanism that could be responsible for accumulating magnetic flux near the surface.

The idea of magnetic-field clustering has been discussed in the contexts of both deeply rooted (Ruzmaikin, 1998) and shallow (Schatten, 2007) dynamo scenarios that are now becoming more fashionable (Pipin and Kosovichev, 2011). Here we discuss the negative effective magnetic-pressure instability (Kleeorin, Rogachevskii, and Ruzmaikin, 1989, 1990; Kleeorin and Rogachevskii, 1994; Kleeorin, Mond, and Rogachevskii, 1996; Rogachevskii and Kleeorin, 2007; Brandenburg, Kleeorin, and Rogachevskii, 2010; Brandenburg *et al.*, 2011) as a possible mechanism for producing magnetic-flux concentrations of the form of active regions. Of course a magnetic field $[\boldsymbol{B}]$ always gives rise to a positive magnetic pressure $[\boldsymbol{B}^2/2\mu_0]$, where μ_0 is the vacuum permeability. In a turbulent medium, however, magnetic fields also suppress the turbulence and thus decrease the turbulent pressure $[\rho \boldsymbol{u}^2/3]$, and modify the pressure caused by magnetic fluctuations $[\boldsymbol{b}^2/6\mu_0]$. Here, $\boldsymbol{u}$ and $\boldsymbol{b}$ are velocity and magnetic fluctuations, ρ is the density, μ_0 is the vacuum permeability, and the coefficients in the turbulent fluid and magnetic pressure are given for isotropic turbulence. Magnetic fluctuations can be due to both small-scale dynamo action as well as tangling of a large-scale field $[\overline{\boldsymbol{B}}]$. The total field is thus $\boldsymbol{B} = \overline{\boldsymbol{B}} + \boldsymbol{b}$. The sum of both effects $[p_{\rm turb} = \overline{\rho \boldsymbol{u}^2}/3 + \overline{\boldsymbol{b}^2}/6\mu_0]$ is positive definite, but it depends on $\overline{\boldsymbol{B}}$, and $p_{\rm turb}$ tends to be reduced as $\overline{B} \equiv |\overline{\boldsymbol{B}}|$ increases. Indeed, $p_{\rm turb} = 2E_T/3 - \overline{\boldsymbol{b}^2}/6\mu_0$, where the total turbulent energy $E_T = \overline{\rho \boldsymbol{u}^2}/2 + \overline{\boldsymbol{b}^2}/2\mu_0 \approx$ constant, so that the change of the turbulent pressure is negative ($\delta p_{\rm turb} < 0$) when the magnetic fluctuations are generated by tangling of the mean magnetic field by the velocity fluctuations at the expense of turbulent kinetic energy (Kleeorin, Rogachevskii, and Ruzmaikin, 1990; Brandenburg *et al.*, 2011). Thus, we write

$$p_{\rm turb}(\overline{B}) = p_{\rm turb}(0) - q_{\rm p}(\overline{B})\,\overline{B}^2/2\mu_0, \tag{1}$$

where $p_{\rm turb}(0)$ is the turbulent pressure at zero mean field and $q_{\rm p}(\overline{B})$ is a positive function of $\overline{B}$ such that $p_{\rm turb}$ is reduced in the presence of $\overline{B}$. The pressure $p_{\rm turb}(0)$ only includes those contributions from $\boldsymbol{b}^2$ that are associated with small-scale dynamo action, but not the magnetic fluctuations resulting from the tangling of the mean magnetic field. The relevant magnetic pressure in the evolution equation for the mean flow $[\overline{\boldsymbol{U}}]$ is then not just $\overline{B}^2/2\mu_0$, but it is affected by the $\overline{B}$-dependence of $p_{\rm turb}$, *i.e.* it depends on

$$p_{\rm turb}(\overline{B}) + \overline{B}^2/2\mu_0 = p_{\rm turb}(0) + \left[1 - q_{\rm p}(\overline{B})\right]\overline{B}^2/2\mu_0, \tag{2}$$

which is also still positive. But $1 - q_{\rm p}(\overline{B})$ may well become negative, which leads to what we call a *negative effective magnetic pressure*. The first term, $p_{\rm turb}(0)$, on the right-hand side of Equation (2) is independent of $\overline{\boldsymbol{B}}$, so it ignores the dynamics of the mean field and only affects the density scale height. Therefore, the expression

$$p_{\rm eff} = (1 - q_{\rm p})\overline{B}^2/2\mu_0 \tag{3}$$

is referred to as the *effective* magnetic pressure. We emphasize that the effective magnetic pressure is an averaged quantity describing the dependence on the mean magnetic field. One

could therefore also talk about a magnetic mean-field pressure. In addition, there is also the gas pressure [$p_{\rm gas}$]. Once the effective magnetic pressure drives a mean flow, the gas density changes, and as a consequence the gas pressure, so as to re-establish approximate total pressure balance. Therefore, $p_{\rm gas}$ and ρ will also depend on $\overline{B}$.

In the presence of gravity, the properties of magnetic buoyancy are drastically altered by a negative effective magnetic pressure. In the following we illustrate how this can lead to an instability. Since the flow velocities are highly subsonic, we can make the anelastic approximation, *i.e.* $\boldsymbol{\nabla}\cdot\overline{\rho}\overline{\boldsymbol{U}}=0$. This leads to $\boldsymbol{\nabla}\cdot\overline{\boldsymbol{U}}+\overline{\boldsymbol{U}}\cdot\boldsymbol{\nabla}\ln\overline{\rho}=0$, or

$$\boldsymbol{\nabla}\cdot\overline{\boldsymbol{U}}=\frac{\overline{U}_z}{H_\rho}, \tag{4}$$

where we have used the density scale height [H_ρ], so that $\boldsymbol{\nabla}\ln\overline{\rho}=(0,0,-1/H_\rho)$. This equation shows that a downward motion, $\overline{U}_z<0$, leads to compression: $\boldsymbol{\nabla}\cdot\overline{\boldsymbol{U}}<0$. This enhances an applied field locally. We consider an applied equilibrium magnetic field of the form $(0,B_0,0)$ and the mean field has only a y-component, *i.e.* $\overline{\boldsymbol{B}}=(0,\overline{B}_y(x,z),0)$, so we have

$$\frac{{\rm D}\overline{B}_y}{{\rm D}t}=-\overline{B}_y\boldsymbol{\nabla}\cdot\overline{\boldsymbol{U}}, \tag{5}$$

where ${\rm D}/{\rm D}t=\partial/\partial+\overline{\boldsymbol{U}}\cdot\boldsymbol{\nabla}$ is the advective derivative. Note that for a magnetic field with only a y-component, but $\partial/\partial y=0$, there is no stretching term, so there is no term of the form $\overline{\boldsymbol{B}}\cdot\boldsymbol{\nabla}\overline{\boldsymbol{U}}$. Using Equation (4), and linearizing Equation (5) around $\overline{\boldsymbol{U}}=\boldsymbol{0}$ and $\overline{\boldsymbol{B}}=\boldsymbol{B}_0$, we have

$$\frac{\partial\overline{B}_{1y}}{\partial t}=-B_0\frac{\overline{U}_{1z}}{H_\rho}, \tag{6}$$

where subscripts 1 denote linearized quantities. The vertical-velocity perturbation [$\overline{U}_{1z}$] is caused by magnetic buoyancy. Assuming total pressure equilibrium, $p_{\rm gas}+p_{\rm eff}=$ constant, we see that an increase in the effective magnetic pressure causes a *decrease* is the gas pressure, *i.e.* $\delta p_{\rm gas}=-\delta p_{\rm eff}$, just as in the regular magnetic-buoyancy instability. Therefore, the Archimedian buoyancy force is

$$-\frac{\delta\rho}{\rho}g=-\frac{\delta p_{\rm gas}}{p_{\rm gas}}g=\frac{\delta p_{\rm eff}}{\rho c_{\rm s}^2}g=\frac{{\rm d}p_{\rm eff}}{{\rm d}\overline{B}^2}\frac{\delta\overline{B}^2}{\rho c_{\rm s}^2}g, \tag{7}$$

where we have used $p_{\rm gas}=\rho c_{\rm s}^2$ for an isothermal gas. In the regular magnetic-buoyancy instability (Parker, 1966, 1979), without turbulence effects, we have $2\mu_0\,{\rm d}p_{\rm eff}/{\rm d}\overline{B}^2=1$. In the domain where the negative effective magnetic-pressure effect causes ${\rm d}p_{\rm eff}/{\rm d}\overline{B}^2$ to be negative, a magnetic-field enhancement leads to a further reduction of the local pressure, which is compensated by horizontal inflows, increasing density (and field strength), making this fluid parcel heavier, causing it to sink. Conversely, a local field reduction causes outflows and rises until it reaches the region where this feedback reverses. Thus, the instability loop is closed by considering the momentum equation in its linearized form:

$$\frac{\partial\overline{U}_{1z}}{\partial t}=\frac{{\rm d}p_{\rm eff}}{{\rm d}\overline{B}^2}\frac{2B_0\overline{B}_{1y}}{\rho c_{\rm s}^2}g. \tag{8}$$

Using $c_{\rm s}^2/g=H_\rho$ for an isothermal atmosphere, we then find the dispersion relation for the growth rate [λ] of the resulting instability

$$\lambda=\frac{v_{\rm A}}{H_\rho}\sqrt{-2\mu_0\,p'_{\rm eff}}-\eta_{\rm t}k^2, \tag{9}$$

where $v_A = B_0/\sqrt{\mu_0 \rho}$ is the Alfvén speed and

$$2\mu_0 \, p'_{\rm eff} = 2\mu_0 \, {\rm d}p_{\rm eff}/{\rm d}\overline{B}^2 = 1 - q_{\rm p} - {\rm d}q_{\rm p}/{\rm d}\ln \overline{B}^2 \tag{10}$$

is twice the derivative of the effective magnetic pressure. We have also included here the effects of turbulent magnetic diffusivity [$\eta_{\rm t}$] and turbulent magnetic viscosity [$\nu_{\rm t}$], assuming $\nu_{\rm t}/\eta_{\rm t} = 1$. Here, k is the effective wavenumber. A proper derivation of the growth rate of the instability, but again without including turbulent magnetic diffusivity and turbulent magnetic viscosity, is given in Appendix A. This analysis shows that the first term in Equation (9) is to be multiplied by a factor k_x/k, as is also familiar for gravity modes (Stein and Leibacher, 1974). This implies that $\lambda = 0$ for $k_x = 0$, and that $\lambda(k_x)$ has a maximum for intermediate values of k_x; see Appendix A.

The negative contribution of turbulence to the mean magnetic pressure and the resulting large-scale instability has been predicted long ago (Kleeorin, Rogachevskii, and Ruzmaikin, 1990; Kleeorin, Mond, and Rogachevskii, 1996; Kleeorin and Rogachevskii, 1994). However, this instability has been detected in direct numerical simulations (DNS) only recently (Brandenburg *et al.*, 2011; Kemel *et al.*, 2012a). This large-scale instability is called the negative effective magnetic pressure instability (NEMPI).

Equation (9) demonstrates that stronger stratification and thus a smaller scale height leads to an increased growth rate of the instability. This was qualitatively confirmed by Kemel *et al.* (2012b). Using numerical solutions of the full mean-field equations, they found furthermore that the growth rate of the instability is actually independent of v_A. This seems to be at odds with Equation (9). To understand this, we use the following approximation for $q_{\rm p}$, based on fits to the DNS results (Brandenburg *et al.*, 2011, 2012; Kemel *et al.*, 2012b):

$$q_{\rm p}(\beta) = \frac{\beta_\star^2}{\beta_{\rm p}^2 + \beta^2}, \tag{11}$$

where $\beta_\star$ and $\beta_{\rm p}$ are constants, $\beta = \overline{B}/B_{\rm eq}$ is the modulus of the normalized mean magnetic field, and $B_{\rm eq} = \sqrt{\mu_0 \rho} u_{\rm rms}$ is the equipartition field strength. Thus, for $\beta_\star \gg \beta \gg \beta_{\rm p}$, we have

$$\lambda \approx \beta_\star \frac{u_{\rm rms}}{H_\rho} - \eta_{\rm t} k^2, \tag{12}$$

so the growth rate is indeed independent of the imposed field strength.

In a mean-field model, $u_{\rm rms}$ is normally expressed in terms of $\eta_{\rm t} = u_{\rm rms}/3k_{\rm f}$, where $k_{\rm f}$ is the wavenumber of the energy-carrying eddies, so Equation (12) becomes

$$\frac{\lambda}{\eta_{\rm t} k^2} \approx 3\beta_\star \frac{k_{\rm f}/k}{k H_\rho} - 1, \tag{13}$$

which illustrates immediately the importance of large enough scale separation, *i.e.* large enough values of $k_{\rm f}/k$.

The purpose of this article is to show that NEMPI can work over a range of different field strengths. Such a result was recently predicted using the mean-field simulations (MFS) by Kemel *et al.* (2012b). We shall also investigate the close connection between MFS and DNS results, which allows us to determine the resulting effective magnetic pressure as a function of the mean magnetic field in the plane perpendicular to the mean field. Here we focus on a series of simulations with different field strengths, but for a fixed value of the magnetic Reynolds number and fixed value of the scale separation ratio. For a numerical study of the dependence on magnetic Reynolds number and on scale separation ratio, but fixed field strength, we refer to the recent work of Kemel *et al.* (2012a). In the following, we discuss first DNS of NEMPI and turn then to MFS. We begin with a simplistic illustration of the nature of NEMPI.

2. Vertical Profile of Effective Magnetic Pressure

The first successful DNS of NEMPI has been possible under the assumption of an isothermally stratified layer with an isothermal equation of state (Brandenburg *et al.*, 2011). Much of the same physics is also possible in adiabatically stratified layers, but NEMPI was found in this case only in mean-field models (Brandenburg, Kleeorin, and Rogachevskii, 2010; Käpylä *et al.*, 2012). The isothermal case has conceptual advantages that help us understand better the underlying physics of this instability. We make use of this advantage in the present article, too.

In most of the isothermal setups studied so far, the rms velocity is only weakly dependent on height, so the z-variation of $B_{\rm eq}$ was only caused by that of $\rho = \rho_0 \exp(-z/H_\rho)$. Before we perform DNS and MFS, let us determine the conditions for which NEMPI is most effective. To this end we plot the effective magnetic pressure, which is normalized by the local equipartition field strength $[B_{\rm eq}]$,

$$\mathcal{P}_{\rm eff}(\beta) = \frac{1}{2}\big[1 - q_{\rm p}(\beta)\big]\beta^2, \tag{14}$$

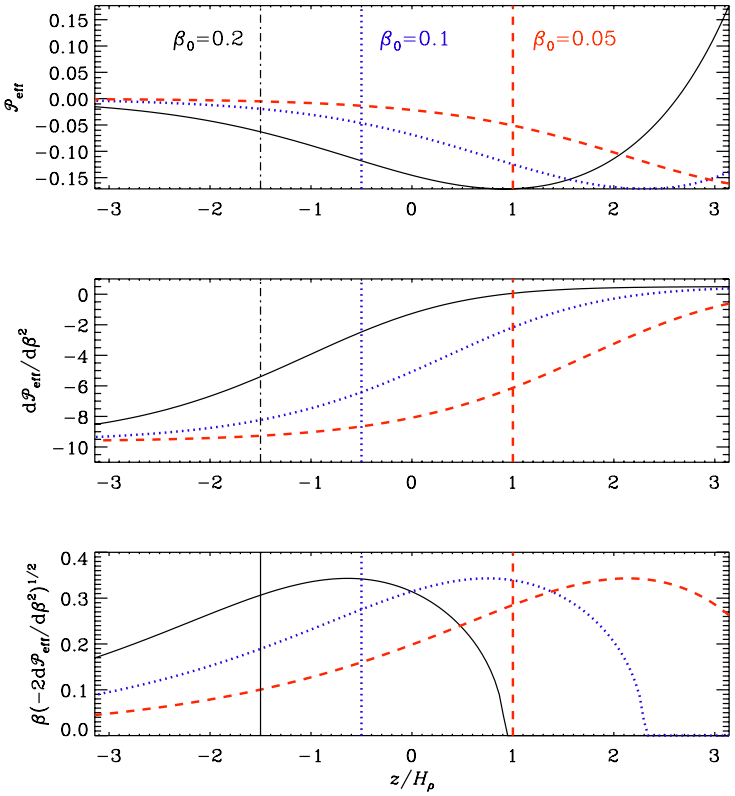


Figure 1 Profiles of $\mathcal{P}_{\rm eff}$, $d\mathcal{P}_{\rm eff}/d\beta^2$, and $\beta(-2d\mathcal{P}_{\rm eff}/d\beta^2)^{1/2}$ for fit parameters $q_{\rm p0} = 20$ and $\beta_{\rm p} = 0.167$, and the three field strengths $B_0/B_{\rm eq0} = 0.05$, 0.1, and 0.2 within the z-range from $-\pi$ to π, which is consistent with some of the models discussed below. The vertical lines of similar line types give the location where the unstable eigenmodes reaches its peak.

i.e. $\mathcal{P}_{\rm eff} = p_{\rm eff}/B_{\rm eq}^2$, where $q_{\rm p}(\beta)$ is determined by Equation (11). Since $\beta = \beta(z) = \beta_0 \exp(z/2H_\rho)$ increases with z, $\mathcal{P}_{\rm eff}(z)$ is small at large depths, reaches a negative minimum at some depth, and then becomes positive and equal to β^2. In Figure 1 we show vertical profiles of $\mathcal{P}_{\rm eff}$, $\mathrm{d}\mathcal{P}_{\rm eff}/\mathrm{d}\beta^2$, and $\beta(-2\mathrm{d}\mathcal{P}_{\rm eff}/\mathrm{d}\beta^2)^{1/2}$ for the fit parameters $q_{\rm p0} = 20$ and $\beta_{\rm p} = 0.167$ derived later in this article, and the three field strengths $\beta_0 \equiv B_0/B_{\rm eq0} = 0.05$, 0.1, and 0.2 within the z/H_ρ-range from $-\pi$ to π, which is also consistent with the DNS and some of the MFS discussed below. Here, $B_{\rm eq0} = B_{\rm eq}(z=0)$, $q_{\rm p0} = q_{\rm p}(\beta=0)$, and $\beta_\star = \beta_{\rm p}\sqrt{q_{\rm p0}}$.

Notice first of all that all three curves of $\mathcal{P}_{\rm eff}$ have minima with left flanks (negative slopes) within the domain. As the imposed field is increased, these curves shift downward (smaller values of z). Thus, we should expect the peak of the unstable eigenmode to appear somewhere along the left flanks of these curves and that these peaks move further down as the imposed field is increased. This is qualitatively reproduced by the DNS and MFS discussed below, except that the location is consistently a certain distance below the position where the left flanks have their steepest gradient. On the other hand, as is evident from the middle panel of Figure 1, the largest value of $\mathrm{d}\mathcal{P}_{\rm eff}/\mathrm{d}\beta^2$ is always achieved at the bottom of the domain. However, the growth rate of NEMPI still has a factor proportional to $v_{\rm A} = u_{\rm rms}\beta$ in front of it; see Equation (9). This then confines the instability to a narrow strip within the domain. In the third panel of Figure 1 we also plot therefore $\beta(-2\mathrm{d}\mathcal{P}_{\rm eff}/\mathrm{d}\beta^2)^{1/2}$, and their extrema are now only slightly above the location where DNS and MFS show a peak in the eigenfunction. The reason for the remaining discrepancy is not well understood at present.

3. Onset and Saturation of NEMPI in DNS

3.1. Isothermal Setup in DNS

Following the earlier work of Brandenburg *et al.* (2011) and Kemel *et al.* (2012a), we solve the equations for the velocity [$\boldsymbol{U}$], the magnetic vector potential [$\boldsymbol{A}$], and the density [ρ]:

$$\rho\frac{\mathrm{D}\boldsymbol{U}}{\mathrm{D}t} = -c_{\rm s}^2\boldsymbol{\nabla}\rho + \boldsymbol{J}\times\boldsymbol{B} + \rho(\boldsymbol{f}+\boldsymbol{g}) + \boldsymbol{\nabla}\cdot(2\nu\rho\mathbf{S}), \tag{15}$$

$$\frac{\partial\boldsymbol{A}}{\partial t} = \boldsymbol{U}\times\boldsymbol{B} + \eta\nabla^2\boldsymbol{A}, \tag{16}$$

$$\frac{\partial\rho}{\partial t} = -\boldsymbol{\nabla}\cdot\rho\boldsymbol{U}, \tag{17}$$

where ν is the kinematic viscosity, η is the magnetic diffusivity due to Spitzer conductivity of the plasma, $\boldsymbol{B} = \boldsymbol{B}_0 + \boldsymbol{\nabla}\times\boldsymbol{A}$ is the magnetic field, $\boldsymbol{B}_0 = (0, B_0, 0)$ is the imposed uniform field, $\boldsymbol{J} = \boldsymbol{\nabla}\times\boldsymbol{B}/\mu_0$ is the current density, μ_0 is the vacuum permeability, $\mathsf{S}_{ij} = \frac{1}{2}(\partial_j U_i + \partial_i U_j) - \frac{1}{3}\delta_{ij}\boldsymbol{\nabla}\cdot\boldsymbol{U}$ is the traceless rate-of-strain tensor. The forcing function [$\boldsymbol{f}$] consists of random, white-in-time, plane, non-polarized waves with a certain average wavenumber [$k_{\rm f}$]. The turbulent rms velocity is approximately independent of z with $u_{\rm rms} = \langle\boldsymbol{u}^2\rangle^{1/2} \approx 0.1\,c_{\rm s}$. The gravitational acceleration [$\boldsymbol{g} = (0,0,-g)$] is chosen such that $k_1 H_\rho = 1$, so the density contrast between bottom and top is $\exp(2\pi) \approx 535$. Here, $H_\rho = c_{\rm s}^2/g$ is the density scale height and $k_1 = 2\pi/L$ is the smallest wavenumber that fits into the cubic domain of size L^3. In most of our calculations, structures develop whose horizontal wavenumber [k_x] is close to k_1. We consider a domain of size $L_x \times L_y \times L_z$ in Cartesian coordinates (x, y, z), with periodic boundary conditions in the x- and y-directions and stress-free, perfectly conducting boundaries at the top and bottom ($z = \pm L_z/2$). In all

cases, we use a scale separation ratio $[k_f/k_1]$ of 30, a fluid Reynolds number $\mathrm{Re} \equiv u_{\rm rms}/\nu k_f$ of 18, and a magnetic Prandtl number $[\mathrm{Pr}_M = \nu/\eta]$ of 0.5. In our units, $\mu_0 = 1$ and $c_s = 1$. The value of B_0 is specified in units of the volume-averaged value $[B_{\rm eq0} = \sqrt{\mu_0 \rho_0}\, u_{\rm rms}]$ where $\rho_0 = \langle \rho \rangle$ is the volume-averaged density, which is constant in time. In addition to visualizations of the actual magnetic field, we also monitor $\overline{B}_y$, which is an average over y and a certain time interval $[\Delta t]$. Time is sometimes specified in terms of turbulent-diffusive times $[t\, \eta_{t0} k_1^2]$, where $\eta_{t0} = u_{\rm rms}/3k_f$ is the estimated turbulent diffusivity. Since the simulations are periodic in the x- and y-directions, we sometimes shift the images such that the peak field strength of NEMPI appears in the middle of the frame.

The simulations are performed with the PENCIL CODE [http://pencil-code.googlecode.com] which uses sixth-order explicit finite differences in space and a third-order accurate time stepping method. We use a numerical resolution of 256^3 mesh points.

3.2. Results

In Figure 2 we demonstrate that NEMPI can work over a range of field strengths. As we increase the strength of the imposed field, NEMPI develops at progressively greater depth. This result was recently obtained for MFS, but is now for the first time demonstrated in DNS. Figure 3 shows that the growth of the large-scale field $[\overline{B}_1]$ of the magnetic structure is similar for three different field strengths. Here, $\overline{B}_1$ has been determined by taking the maximum value of the mean field in the neighborhood of the position where the flux concentration later develops. Note that there is a range over which $\overline{B}_1$ grows approximately exponentially, independent of the value of B_0.

In Figure 4 we show $\overline{B}_y$ at early, intermediate, and late stages of the saturation process (left), and compare with visualizations of $\mathcal{P}_{\rm eff}$ at the same times. Here, $\mathcal{P}_{\rm eff} = \frac{1}{2}(1 - q_{\rm p})\beta^2$, where $q_{\rm p}(\beta)$ with $\beta = \overline{B}/B_{\rm eq}$ is evaluated from

$$q_{\rm p} = -2\Delta \overline{\Pi}^{\rm f}_{xx} / \overline{\boldsymbol{B}}^2, \tag{18}$$

for $\boldsymbol{B}_0 = (0, B_0, 0)$, and

$$\Delta \overline{\Pi}^{\rm f}_{ii} = \overline{\rho}\left(\overline{u_i^2} - \overline{u_{0i}^2}\right) + \frac{1}{2}\left(\overline{\boldsymbol{b}^2} - \overline{\boldsymbol{b}_0^2}\right) - \left(\overline{b_i^2} - \overline{b_{0i}^2}\right), \tag{19}$$

is applied to the xx-component of the total stress from the fluctuating velocity and magnetic fields. In Equation (19) no summation over the repeated index i is assumed.

In Figure 4, blue shades correspond to low values of $\mathcal{P}_{\rm eff}$ and occur around the minimum line (marked in white) where $\mathcal{P}_{\rm eff} = \mathcal{P}_{\rm min}$. As time progresses, low values of $\mathcal{P}_{\rm eff}$ are also found at greater depth as the magnetic-flux concentration descends. The fact that there is a clear spatial correlation between $\overline{B}_y$ and $\mathcal{P}_{\rm eff}$ provides strong evidence that the interpretation of the formation of structures in the stratified turbulence simulations in terms of NEMPI is indeed the correct one.

The descending structures have previously been referred to as "potato sack" structures (Brandenburg *et al.*, 2011), because of their widening cross-section with greater depth. When such structures were first seen in MFS (Brandenburg, Kleeorin, and Rogachevskii, 2010), they were originally thought to be artifacts of the model that one would not expect to see in the Sun. However, such structures were later also found in DNS (Brandenburg *et al.*, 2011), highlighting therefore the strong predictive power of MFS.

A visualization of the resulting mean flow $[\overline{\boldsymbol{U}}]$ is shown in Figure 5 as vectors. The flow shows a convergent shape toward the magnetic structures. It is interesting to note that such convergent flow structures are now also seen in local-helioseismic flow measurements

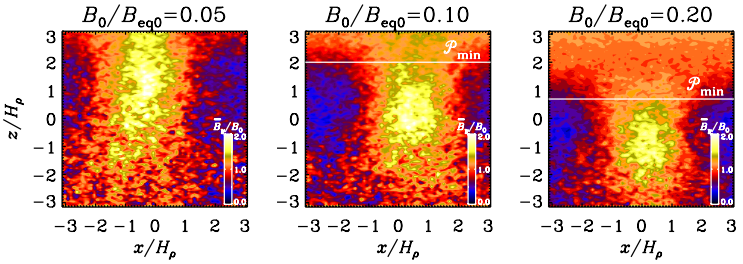


Figure 2 $\overline{B}_y/B_0$ from DNS for three values of the imposed field strength at the end of the linear growth phase of NEMPI for $\mathrm{Re}_M = 18$ and $\mathrm{Pr}_M = 0.5$. The times are $t\,\eta_{t0}/H_\rho^2 \approx 0.93$, 1.09, and 1.26 for $B_0/B_{eq0} = 0.05$, 0.1, and 0.2, respectively. The location of the $\mathcal{P}_{min}$ line is indicated in panels 2 and 3, while for panel 1 it lies above the computational domain.

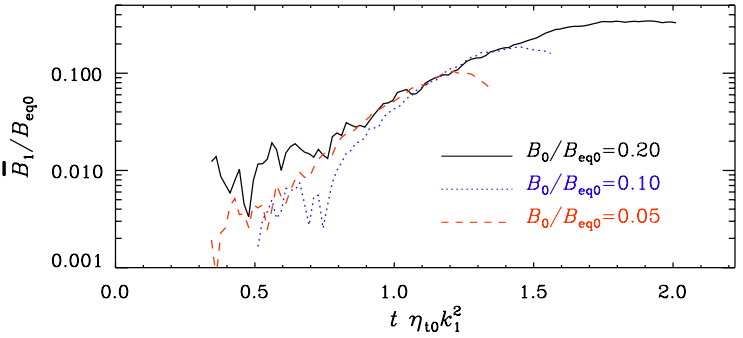


Figure 3 Growth of the large-scale field strength $[\overline{B}_1]$ at the center of the magnetic structure for three field strengths.

around active regions (Hindman, Haber, and Toomre, 2009). In this connection it is instructive to discuss the somewhat peculiar shape of such a structure that widens as it descends. Normally, in a strongly stratified atmosphere, descending structures get compressed and become narrower, but this is not seen in the present visualizations. As already argued by Brandenburg *et al.* (2012), this is because the boundaries of these structures do not coincide with material lines, so the mass is not conserved inside them and can leak through the boundaries. Indeed, these structures grow as they descend, and may become amenable to helioseismic detection; *cf.* Ilonidis, Zhao, and Kosovichev (2011). This phenomenon is well known in the description of turbulent plumes as a model of turbulent downdrafts in convection (Rieutord and Zahn, 1995). Such structures are known to widen as a result of entrainment. The sinking behavior of these apparently disconnected flow structures can be explained as follows: while inflows dominate downflows throughout the whole lifespan of the field concentration, in the initial stage the former can drag in a large fraction of the surrounding magnetic field, overcompensating the losses by downflows. However, as the environment gets depleted, this dynamical balance shifts and the structures start moving downwards.

3.3. Mean-Field Coefficients from DNS

In earlier work by Kemel *et al.* (2012b), the parameters $q_{p0} = 40$ and $\beta_p = 0.05$, corresponding to $\beta_\star = 0.32$, were used. Those values are compatible with work by Brandenburg *et al.*

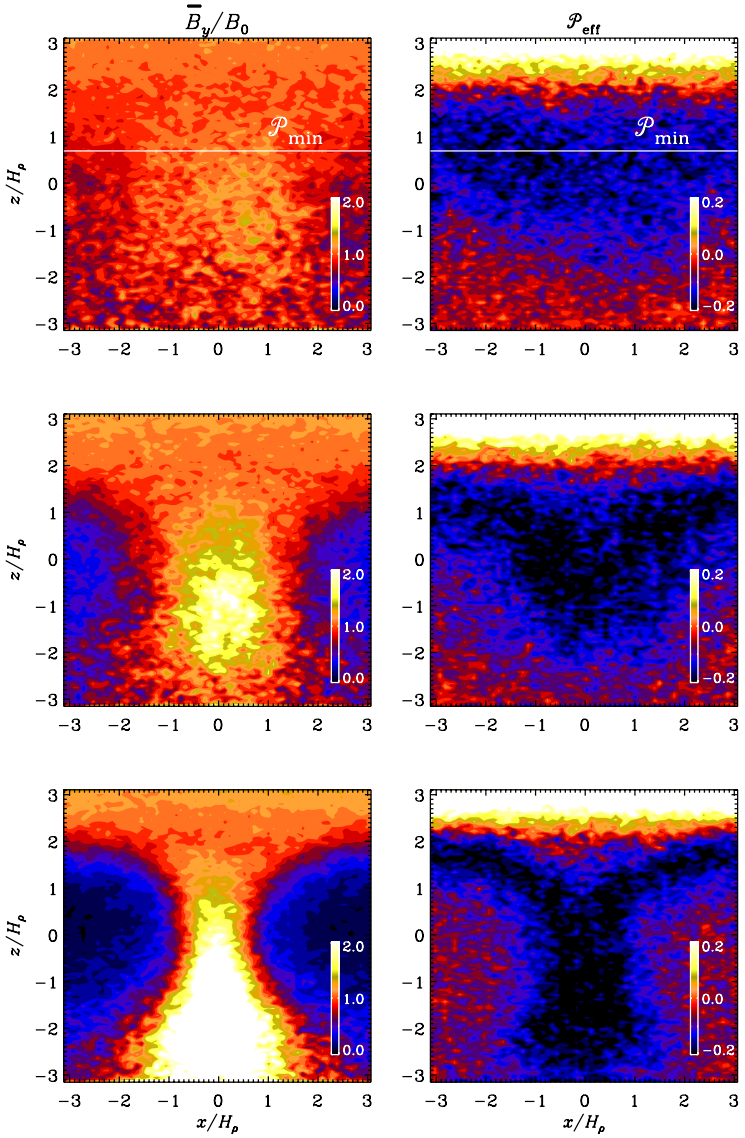


Figure 4 $\overline{B}_y$ and $\mathcal{P}_{\rm eff}$ from DNS at three times ($t\,\eta_{\rm t0}/H_\rho^2 \approx 0.84$, 1.34 and 1.84) showing the descent of the potato sack feature for $\mathrm{Re}_M = 18$, $\mathrm{Pr}_M = 0.5$, and $B_0/B_{\rm eq0} = 0.20$.

(2012) and Kemel *et al.* (2012a). However, in the present case we have a larger scale separation ratio, $k_{\rm f}/k_1 = 30$, for which these parameters have not yet been determined. In Figure 6 we show the functional form of $\mathcal{P}_{\rm eff}(\beta)$ for the present case with $k_{\rm f}/k_1 = 30$, $\mathrm{Re}_M = 18$, and $\mathrm{Pr}_M = 0.5$. Here we have followed the method described by Brandenburg *et al.* (2012); see their Equation (17). For the present model we find as fit parameters $q_{\rm p0} = 20$ and $\beta_{\rm p} = 0.167$, which corresponds to $\beta_\star = 0.75$.

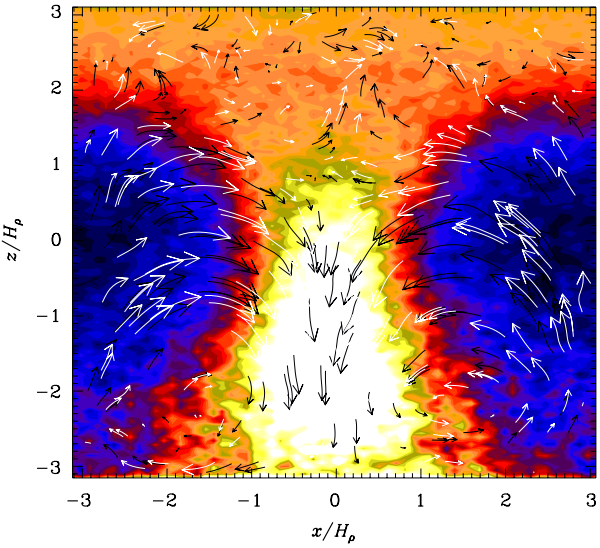


Figure 5 Vectors of $\overline{U}$ together with a color/grey-scale representation of $\overline{B}_y$ from DNS at a late time ($t\,\eta_{\rm t0}/H_\rho^2 \approx 1.7$) for $\mathrm{Re}_M = 18$, $\mathrm{Pr}_M = 0.5$, and $B_0/B_{\rm eq0} = 0.20$. The color scale is the same as in the lower left panel of Figure 4, and the longest vector corresponds to a flow speed of $0.04c_{\rm s}$ or $0.4u_{\rm rms}$.

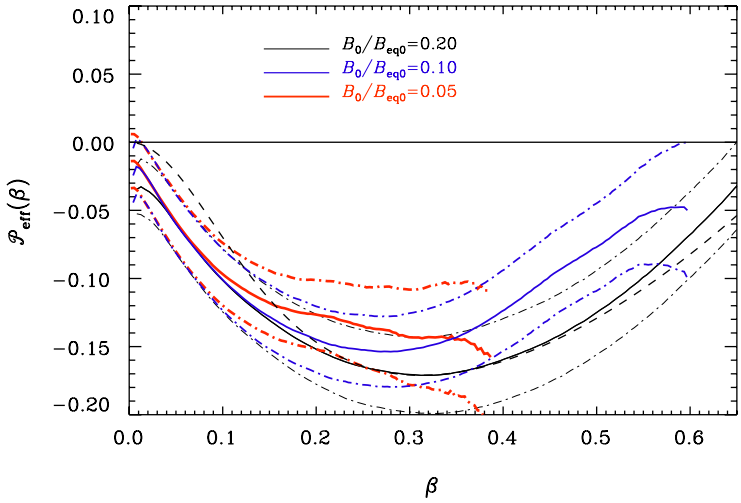


Figure 6 $\mathcal{P}_{\rm eff}(\beta)$ for the DNS used in this article with $k_{\rm f}/k_1 = 30$, $\mathrm{Re}_M = 18$, and $\mathrm{Pr}_M = 0.5$. The dashed line is a fit based on Equation (11) with $q_{\rm p0} = 20$ and $\beta_{\rm p} = 0.167$, which corresponds to $\beta_\star = 0.75$, and the dashed-dotted lines show the error-bar range.

4. Comparison with MFS

Recently, many aspects of NEMPI seen in the DNS have also been detected in MFS. Establishing the usefulness and limitations of MFS is important, because such models are easier to solve and allow one to explore parameters in regimes where DNS are harder to apply or have not yet been applied in the limited time since the close correspondence between DNS and MFS was first noted.

In the following we consider two-dimensional mean-field models, in which the presence of $q_{\rm s}$ has no effect on the solutions (Kemel *et al.*, 2012b). Furthermore, we ignore other

effects connected with the anisotropy of the turbulence. These effects have previously been found to be weak (Brandenburg *et al.*, 2012; Käpylä *et al.*, 2012).

4.1. Isothermal Setup in MFS

In this section we solve the evolution equations for mean velocity $\overline{\boldsymbol{U}}$, mean density $\overline{\rho}$, and mean vector potential $\overline{\boldsymbol{A}}$, in the form

$$\frac{\partial \overline{\boldsymbol{U}}}{\partial t} = -\overline{\boldsymbol{U}} \cdot \boldsymbol{\nabla} \overline{\boldsymbol{U}} - c_{\rm s}^2 \boldsymbol{\nabla} \ln \overline{\rho} + \boldsymbol{g} + \overline{\boldsymbol{\mathcal{F}}}_{\rm M} + \overline{\boldsymbol{\mathcal{F}}}_{\rm K}, \tag{20}$$

$$\frac{\partial \overline{\rho}}{\partial t} = -\overline{\boldsymbol{U}} \cdot \boldsymbol{\nabla} \overline{\rho} - \overline{\rho} \boldsymbol{\nabla} \cdot \overline{\boldsymbol{U}}, \tag{21}$$

$$\frac{\partial \overline{\boldsymbol{A}}}{\partial t} = \overline{\boldsymbol{U}} \times \overline{\boldsymbol{B}} - (\eta_{\rm t} + \eta) \overline{\boldsymbol{J}}, \tag{22}$$

where $\overline{\boldsymbol{\mathcal{F}}}_{\rm M}$ is given by

$$\overline{\rho} \overline{\boldsymbol{\mathcal{F}}}_{\rm M} = -\frac{1}{2} \boldsymbol{\nabla} \big[(1 - q_{\rm p}) \overline{\boldsymbol{B}}^2 \big], \tag{23}$$

and

$$\overline{\boldsymbol{\mathcal{F}}}_{\rm K} = (\nu_{\rm t} + \nu) \left(\nabla^2 \overline{\boldsymbol{U}} + \frac{1}{3} \boldsymbol{\nabla} \boldsymbol{\nabla} \cdot \overline{\boldsymbol{U}} + 2 \overline{\mathsf{S}} \boldsymbol{\nabla} \ln \overline{\rho} \right) \tag{24}$$

is the total (turbulent plus microscopic) viscous force. Here, $\overline{\mathsf{S}}_{ij} = \frac{1}{2} (\overline{U}_{i,j} + \overline{U}_{j,i}) - \frac{1}{3} \delta_{ij} \boldsymbol{\nabla} \cdot \overline{\boldsymbol{U}}$ is the traceless rate-of-strain tensor of the mean flow and $q_{\rm p}$ is approximated by Equation (11), which is only a function of the ratio $\beta \equiv |\overline{\boldsymbol{B}}| / B_{\rm eq}$. In Equation (23) we have taken into account that the mean magnetic is independent of y, so the mean magnetic tension vanishes.

4.2. Aspects of the MFS

We begin by showing $\overline{B}_y$ for three values of the imposed field strength at the end of the linear growth phase of NEMPI. The results are shown in Figures 7 and 8 for two different setups. In the former we use $q_{\rm p0} = 20$ and $\beta_{\rm p} = 0.167$ for the same z-range ($-\pi \leq z/H_\rho \leq \pi$) as in the DNS, while in the latter we use $q_{\rm p0} = 40$ and $\beta_{\rm p} = 0.05$ for somewhat stronger fields and a deeper z-range ($0 \leq z/H_\rho \leq 2\pi$), which is also the fiducial model used by Kemel *et al.* (2012b). In the former case the growth rate is $\approx 11 H_\rho^2 / \eta_t$, while in the latter it is $\approx 5.0 H_\rho^2 / \eta_t$.

Unlike the DNS, the MFS show that in the former series of models with $q_{\rm p0} = 20$ and $\beta_{\rm p} = 0.167$ the x-extent is slightly larger than the optimal horizontal wavelength of the instability, because one sees that some of the structures begin to split into two (Figure 7). This is not the case for the second model with $q_{\rm p0} = 40$ and $\beta_{\rm p} = 0.05$ (Figure 8).

Next, we compare $\overline{B}_y$ with $\mathcal{P}_{\rm eff} = \frac{1}{2}(1 - q_{\rm p}) \beta^2$. Again, there is a close correspondence between the $\overline{B}_y$ field and the resulting distribution of $\mathcal{P}_{\rm eff}$; see Figure 9. Here, $q_{\rm p}(\beta)$ is evaluated using Equation (11). Furthermore, there is a close correspondence between regions of enhanced magnetic field and enhanced density; see the right-most column of Figure 9. Note also that the relative variation of the mean density $[\Delta\overline{\rho}/\overline{\rho}]$ is of the order of 10^{-4}. This small value is a consequence of the plasma β being very large, *i.e.* $\overline{B}^2 / 2\mu_0 \rho c_{\rm s}^2 \ll 1$. Indeed, we expect $\Delta\overline{\rho}/\overline{\rho} \approx \overline{B}^2 / 2\mu_0 \rho c_{\rm s}^2 = \beta^2 M_{\rm a}^2$, where β at $z/H_\rho = -5$ is about 0.1 and $M_{\rm a} = u_{\rm rms}/c_{\rm s}$ is the rms Mach number, which is also about 0.1. We note that similar values are also seen in DNS, but here the variations from the turbulence are typically much larger.

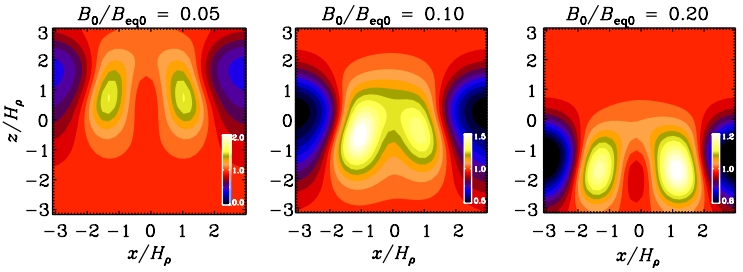


Figure 7 $\overline{B}_y/B_0$ from mean-field models for three values of the imposed field strength at the end of the linear growth phase of NEMPI. Here, $q_{p0} = 20$ and $\beta_p = 0.167$, which corresponds to $\beta_\star = 0.75$. The times are $t\,\eta_t/H_\rho^2 \approx 0.84$, 0.99, and 1.14 for $B_0/B_{eq0} = 0.05$, 0.1, and 0.2, respectively.

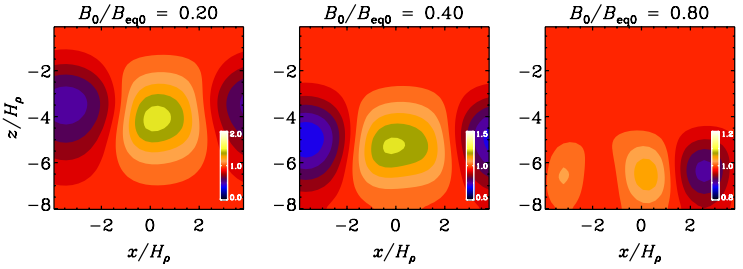


Figure 8 $\overline{B}_y/B_0$ from mean-field models for three values of the imposed field strength at the end of the linear growth phase of NEMPI, but for larger field strengths than in Figure 7 and vertical domain boundaries that are deeper down, so the magnetic field maxima of the instability fit better into the domain. Here, $q_{p0} = 40$ and $\beta_p = 0.05$, corresponding to $\beta_\star = 0.32$.

5. Application to the Sun

The motivation for studying an isothermal model lies in its simplicity, allowing a more thorough investigation of all possible aspects of NEMPI. The Sun is obviously not isothermal nor is it stably stratified. The purpose of this section is to discuss what is known about model dependencies of NEMPI and what else can be learn from the isothermal model. We here discuss the aspects of stratification, scale separation, and the combination of both.

i) *Stratification.* As we have shown above, NEMPI requires strong stratification: decreasing H_ρ increases the growth rate; see Equation (13). The restoring force associated with stable stratification gives rise to Brunt–Väisälä oscillations, which are however eliminated by using an isothermal equation of state. Including it tends to suppress NEMPI, as was shown by Käpylä *et al.* (2012). In an isentropic layer with marginal stability, this stabilizing effect is absent. The temperature increases now linearly with depth. The scale height is then smallest near the top, so NEMPI always tends to develop in the upper layers. Nevertheless, unstable structures still sink; see Brandenburg, Kleeorin, and Rogachevskii (2010).

ii) *Scale separation.* The growth rate increases with the scale separation ratio; see Equation (13). This suppresses NEMPI in deeper layers, so sinking structures must eventually dissolve. Such a variable scale-separation ratio is not yet included in our model. Realistic models of solar convection, such as those of Kitiashvili *et al.* (2010) or Stein *et al.* (2011), may not have enough scale separation; in our models we used $k_f/k_1 = 30$. Also, it is important that the models are run for long enough; see the discussion in Kemel *et al.* (2012a).

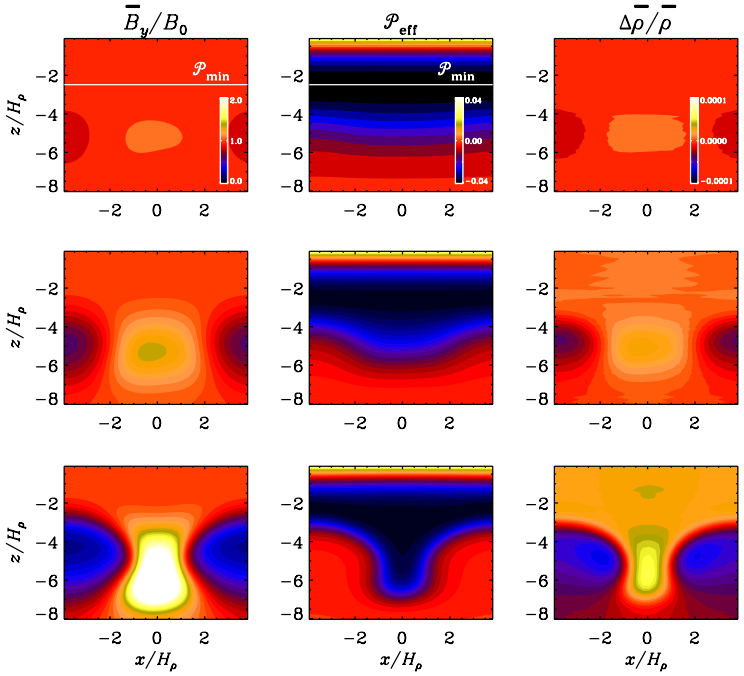


Figure 9 $\overline{B}_y/B_0$, $\mathcal{P}_{\rm eff}$, and $\Delta\overline{\rho}/\overline{\rho}$ from mean-field models at three times: $t\,\eta_{\rm t}/H_\rho^2 \approx 2.2$, 2.5, and 2.8. Here, $B_0/B_{\rm eq0} = 0.40$.

iii) *Relation between stratification and scale separation.* In the model we have varied H_ρ and $k_{\rm f}$ separately, but in the Sun the scale of the turbulent eddies is believed to be proportional to the local scale height, so $k_{\rm f}H_\rho \approx$ constant. Furthermore, in an isothermal domain, the value of k_1 has no physical significance, because the solution is independent of the position of the boundaries provided they are sufficiently far away from the location where the large-scale structures develop. In most of the models presented above, k_1 was equal to the horizontal wavenumber $[k_x]$ of the structures. In Figure 10 we show that, as we decrease H_ρ by a factor of three, the horizontal wavelength also decreases by a factor of three, and so $k_x H_\rho$ is approximately constant. In the following we refer to wavenumbers normalized by $1/H_\rho$ as κ, and since the horizontal wavelength is in the range $6-8\,H_\rho$, we have $\kappa_x \equiv k_x H_\rho \approx 1.0-0.8$. As H_ρ decreases, the vertical extent of the structures (in absolute units) shrinks by the same factor, so the effective k_z/k_x is approximately unity and thus $k \approx \sqrt{2}k_x$, so $\kappa \equiv kH_\rho \approx 1.5-1.1$. This shows that the growth rate for an isothermal atmosphere is fully determined by the coefficient $k_{\rm f}H_\rho$. If it is too small with respect to the strength of turbulent diffusion, the instability will not work.

In stellar mixing-length theory one assumes that the typical scale of the turbulent eddies is the mixing length $[\ell_{\rm mix}]$, given by

$$\ell_{\rm mix} \approx \alpha_{\rm mix} H_p, \tag{25}$$

where $\alpha_{\rm mix} \approx 1.6$ is an empirical dimensionless mixing-length parameter and $H_p \approx \gamma H_\rho$ is the pressure scale height with $\gamma \approx 5/3$ being the ratio of specific heats. Thus, using $k_{\rm f} =$

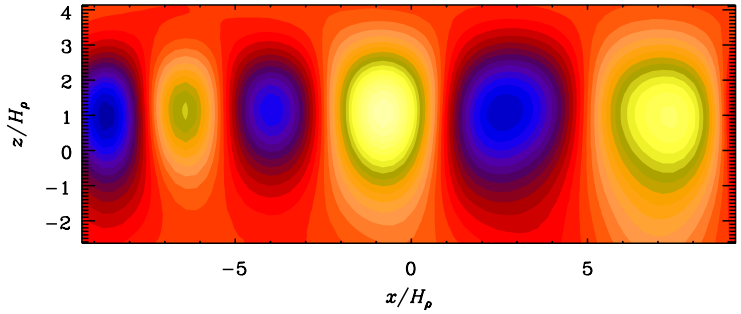


Figure 10 $\overline{B}_y$ in a setup similar to the first simulation from Figure 7 but with gravity three times stronger (or scale height three times smaller), taken at an early time.

$2\pi/\ell_{\rm mix}$, we have $\kappa_{\rm f} \equiv k_{\rm f} H_\rho \approx 2.4$. In that case the growth rate given in Equation (13) becomes

$$\frac{\lambda}{\eta_{\rm t} k^2} \approx 3\beta_\star \frac{k_{\rm f} H_\rho}{(k H_\rho)^2} - 1 = 3\beta_\star \frac{\kappa_{\rm f}}{\kappa^2} - 1. \tag{26}$$

Using $\beta_\star \approx 0.23$ (Kemel *et al.*, 2012a), $\kappa_{\rm f} = 2.4$, and $\kappa = 1.5-1.1$, we find $\lambda/\eta_{\rm t} k^2 = -0.3\ldots+0.4$. On the other hand, using the convection-zone model of Spruit (1974) uses a larger value of $\alpha_{\rm mix}$ such that $\kappa_{\rm f} \approx 3.3$ and thus $\lambda/\eta_{\rm t} k^2 = 0\ldots+0.9$, so the possibility of NEMPI under solar conditions hinges on details of the model. Next, for the near-surface shear layer at a depth of 20 – 40 Mm we estimate

$$\eta_{\rm t} k^2 \approx \frac{u_{\rm rms}}{3k_{\rm f}} \frac{\kappa^2}{H_\rho^2} \approx \frac{u_{\rm rms}}{3H_\rho} \frac{\kappa^2}{\kappa_{\rm f}} = 0.2\text{–}0.3\, \frac{u_{\rm rms}}{H_\rho} = \frac{1}{2\text{–}5\ \text{days}}, \tag{27}$$

while for the Spruit model the e-folding time would be 1 – 3 days. In this depth range the pressure scale height is 7 – 13 Mm, so the density scale height is 12 – 22 Mm, and thus the typical scale of flux concentrations is 80 – 150 Mm. (We recall that in our isothermal model NEMPI can occur at any depth, just depending on the field strength.)

In the Sun, many other effects play a role, for example the convective turbulence is strongly anisotropic with vertical motions dominating over horizontal ones, thus causing negative radial shear in the surface layers (Kitchatinov and Rüdiger, 2005). This may cause new effects that could either enhance or diminish the vertical coupling of different layers. Whether the final outcome still leads to structures resembling active regions can only be answered by performing calculations in which such vertical variations of the stratification are taken into account.

6. Conclusions

The present work has demonstrated that NEMPI is able to concentrate the magnetic field into large patches encompassing the size of many turbulent eddies. The physics of this instability is principally different from the usual interchange instability of continuous magnetic fields (Parker, 1966, 1979; Hughes and Proctor, 1988; Cattaneo and Hughes, 1988; Wissink *et al.*, 2000; Isobe *et al.*, 2005; Kersalé, Hughes, and Tobias, 2007), because here the energy source of the instability is turbulent energy as opposed to gravitational energy.

Furthermore, we are here concerned with a turbulent medium, and so the sign of the buoyancy force is reversed in a regime of intermediate field strength. We have here re-examined the simple case of an isothermal layer in which NEMPI can in principle occur at any depth whose value is determined by the strength of the imposed field. Our new DNS have verified that the growth rate is indeed independent of the strength of the imposed field, provided that the peak of the instability fits still comfortably within the domain. During the subsequent nonlinear evolution of the instability, the overall density stratification readjusts, allowing the magnetic-field concentrations to move further down. It is important to realize that the resulting structures are subject to significant turbulent entrainment (Rieutord and Zahn, 1995), so their boundaries are not closed. The agreement with corresponding mean-field models is remarkable and much more convincing than what has been possible to demonstrate in mean-field dynamo theory. Mean-field models provide therefore a strong source of guidance when designing new setups for DNS.

While NEMPI now begins to be fairly well understood for isothermal models, more work is required for non-isothermal ones. In that case, the density scale height is no longer constant and the degree of stratification is much stronger at the top than in deeper layers. The simple result that the instability can occur at any height, depending just on the strength of the imposed field, is then no longer so obvious. At the same time, there is another, perhaps more important, aspect: the possibility of other instabilities. One of them is connected with the suppression of turbulent convective-energy flux by the mean magnetic field. As shown by Kitchatinov and Mazur (2000), this effect can also lead to magnetic-flux concentrations and it may be sustained for much stronger magnetic-field strengths, thus allowing the formation of structures in which the magnetic pressure becomes comparable to the ambient gas pressure. There may also be a connection with flux-segregation events seen in simulations of magneto-convection at large aspect ratios (Tao *et al.*, 1998; Kitiashvili *et al.*, 2010), which have already been shown to produce bipolar regions in simulations with radiation transfer (Stein *et al.*, 2011). The study of the possibility of producing sunspots similar to those of Rempel (2011a, 2011b), but without initial flux structures, is now of high priority in the quest for solving the solar dynamo problem in terms of distributed dynamo models in which magnetic activity is explained as a surface phenomenon.

Acknowledgements We thank the referees for useful comments that have improved the presentation of the results. Computing resources provided by the Swedish National Allocations Committee at the Center for Parallel Computers at the Royal Institute of Technology in Stockholm and the High Performance Computing Center North in Umeå. This work was supported in part by the European Research Council under the AstroDyn Research Project No. 227952 and the Swedish Research Council under the project grant 621-2011-5076.

Appendix A: Growth rate of NEMPI

In this appendix we derive the growth rate of NEMPI, neglecting for simplicity dissipation processes, using the anelastic approximation, and assuming $H_\rho = $ constant and $\nabla_y = 0$. However, in contrast to the simplistic derivation of Equation (8), we employ here a consistent treatment of lateral pressure variations in the equation of motion, ignoring in Equation (20) the $\overline{\boldsymbol{U}} \cdot \boldsymbol{\nabla} \overline{\boldsymbol{U}}$ nonlinearity,

$$\frac{\partial \overline{\boldsymbol{U}}(t,x,z)}{\partial t} = -\frac{1}{\overline{\rho}} \boldsymbol{\nabla} p_{\rm tot} + \boldsymbol{g}, \tag{28}$$

where $p_{\rm tot} = \overline{p} + p_{\rm eff}$ is the total pressure (the sum of the mean gas pressure $[\overline{p}]$ and the effective magnetic pressure $[p_{\rm eff}]$). We shall take into account that the mean magnetic field

is independent of y, so the mean magnetic tension vanishes. Taking twice the curl of Equation (28), and noting that $\hat{\boldsymbol{z}} \cdot \boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \overline{\boldsymbol{U}} = -\Delta \overline{U}_z + \nabla_z \boldsymbol{\nabla} \cdot \overline{\boldsymbol{U}}$, we obtain

$$\frac{\partial}{\partial t}\left[\Delta \overline{U}_z + \nabla_z (\overline{\boldsymbol{U}} \cdot \boldsymbol{\nabla} \ln \overline{\rho})\right] = \hat{\boldsymbol{z}} \cdot \boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \left(\frac{p_{\rm tot}}{\overline{\rho}^2} \boldsymbol{\nabla} \overline{\rho}\right), \tag{29}$$

where we have used the anelastic approximation in the form $\boldsymbol{\nabla} \cdot \overline{\boldsymbol{U}} = -\overline{\boldsymbol{U}} \cdot \boldsymbol{\nabla} \ln \overline{\rho}$ [see the derivation of Equation (4)] and the fact that under the curl the gradient can be moved to $\overline{\rho}$. If we were to ignore variations of the density, the right-hand side would reduce to $\nabla_x^2 p_{\rm eff}/\overline{\rho} H_\rho$. In the following, however, we retain such variations, which result from the fact that $q_{\rm p} = q_{\rm p}(\beta)$ depends both on $\overline{B}$ and on $\overline{\rho}$. The right-hand side of Equation (29) can be simplified to give

$$\hat{\boldsymbol{z}} \cdot \boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \left(\frac{p_{\rm tot}}{\overline{\rho}^2} \boldsymbol{\nabla} \overline{\rho}\right) = \nabla_x \left[\left(\nabla_z \frac{p_{\rm tot}}{\overline{\rho}}\right) \frac{\nabla_x \overline{\rho}}{\overline{\rho}} - \left(\nabla_x \frac{p_{\rm tot}}{\overline{\rho}}\right) \frac{\nabla_z \overline{\rho}}{\overline{\rho}}\right]. \tag{30}$$

We linearize, indicating small changes by δ or subscript 1, as in Section 1, and note that

$$\delta\left(\frac{p_{\rm tot}}{\overline{\rho}}\right) = \frac{1}{2} v_{\rm A}^2 \left(1 - q_{\rm p} - {\rm d}q_{\rm p}/{\rm d}\ln\beta^2\right)\left(2\frac{\delta \overline{B}_y}{B_0} - \frac{\delta \overline{\rho}}{\overline{\rho}}\right), \tag{31}$$

while

$$\nabla_z \left(\frac{p_{\rm tot}}{\overline{\rho}}\right) = \frac{1}{2} v_{\rm A}^2 \left(1 - q_{\rm p} - {\rm d}q_{\rm p}/{\rm d}\ln\beta^2\right) \frac{1}{H_\rho}. \tag{32}$$

Inserting this into Equation (30), the $\delta\rho$ term in Equation (31) cancels the linearized form of $\nabla_x \overline{\rho}$ in Equation (30), and we are left with

$$\frac{\partial}{\partial t}\left(\Delta - \frac{1}{H_\rho}\nabla_z\right)\delta \overline{U}_z = 2\frac{v_{\rm A}^2}{H_\rho}\frac{{\rm d}\mathcal{P}_{\rm eff}}{{\rm d}\beta^2}\frac{\nabla_x^2 \delta \overline{B}_y}{B_0}, \tag{33}$$

where we have used Equation (14). Introducing a new variable $V_z = \sqrt{\overline{\rho}}\,\delta \overline{U}_z$, and rewriting Equation (33) for V_z, we obtain

$$\frac{\partial}{\partial t}\left(\Delta - \frac{1}{4H_\rho^2}\right) V_z = 2\sqrt{\overline{\rho}}\,\frac{v_{\rm A}^2}{H_\rho}\frac{{\rm d}\mathcal{P}_{\rm eff}}{{\rm d}\beta^2}\frac{\nabla_x^2 \delta \overline{B}_y}{B_0}. \tag{34}$$

Using the linearized form of Equation (6), we arrive at the following equation:

$$\frac{\partial^2}{\partial t^2}\left(\Delta - \frac{1}{4H_\rho^2}\right) V_z(t,x,z) = -\frac{2v_{\rm A}^2}{H_\rho^2}\frac{{\rm d}\mathcal{P}_{\rm eff}}{{\rm d}\beta^2}\nabla_x^2 V_z. \tag{35}$$

It follows from Equation (35) that a necessary condition for the large-scale instability is

$$\frac{{\rm d}\mathcal{P}_{\rm eff}}{{\rm d}\beta^2} < 0. \tag{36}$$

For instance, in the WKB approximation when $k_z H_\rho \gg 1$, *i.e.* when the characteristic scale of the spatial variation of the perturbations of the magnetic and velocity fields are much smaller than the density height length $[H_\rho]$, the growth rate of the instability reads

$$\lambda = \frac{v_{\rm A}}{H_\rho}\left(-2\frac{{\rm d}\mathcal{P}_{\rm eff}}{{\rm d}\beta^2}\right)^{1/2}\frac{k_x}{k}. \tag{37}$$

For an arbitrary vertical inhomogeneity of the density, we seek a solution of Equation (35) in the form $V_z(t,x,z) = V(z)\exp(\lambda t + ik_x x)$ and obtain an eigenvalue problem:

$$\lambda_0^{-2}(z)\left[1 + k_x^{-2}\left(\frac{1}{4H_\rho^2} - \nabla_z^2\right)\right] V_z = \lambda^{-2} V_z, \tag{38}$$

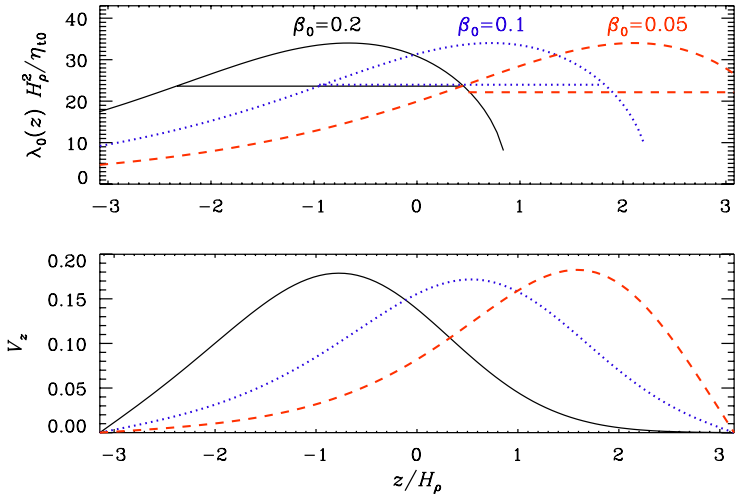


Figure 11 Upper panel: comparison between the graphs of $\lambda_0(z)$ and the corresponding eigenvalues [λ] (horizontal lines segments) for $B_0/B_{\rm eq0} = 0.05$, 0.1, and 0.2. Lower panel: eigenfunctions [$V_z(z)$] for the same three cases obtained by solving Equation (38).

where λ^{-2} is the eigenvalue and

$$\lambda_0(z) = \frac{v_{\rm A}(z)}{H_\rho}\left(-2\frac{{\rm d}\mathcal{P}_{\rm eff}(z)}{{\rm d}\beta^2}\right)^{1/2}. \tag{39}$$

We are interested in the fastest growing solutions, corresponding to the maximum value of λ. We find the maximum value of λ by discretizing Equation (38) on a grid, using a second-order finite-difference scheme for the derivatives, and solving the resultant eigenvalue problem numerically. In Figure 11 the resulting values of λ are compared with $\lambda_0(z)$ and the profiles of $V_z(z)$. The values of λ turn out to be about twice as large as the actual growth rates found in the fully compressible mean-field models, where viscosity and magnetic diffusion are included. A reasonable improvement would be to subtract the damping rate $[\eta_{\rm t}k^2]$ from the ideal growth rates $[\lambda]$; see Equation (9). In Figure 12 we compare $\lambda(k_x)$ with attenuated growth rates $[\lambda(k_x) - \eta_{\rm t}k^2]$ for different values of $\eta_{\rm t}$, assuming that the effective wavenumber obeys $k^2 \approx 2k_x^2$, as explained in paragraph (iii) of Section 5.

It is customary to obtain approximate analytic solutions to Equation (38) as marginally bound states of an associated Schrödinger equation, $\Psi'' - \tilde{U}(R)\,\Psi = 0$, via the transformation

$$\Psi(R) = \sqrt{R}\,V(z), \qquad R(z) = \frac{v_{A0}^2}{u_{\rm rms}^2\beta_{\rm p}^2}\,e^{z/H_\rho}, \tag{40}$$

where $\overline{\rho} = \overline{\rho}_0\,e^{-z/H_\rho}$ is used as mean density profile, $v_{\rm A0} = B_0/\sqrt{\overline{\rho}_0}$ is the Alfvén speed based on the averaged density, and

$$\tilde{U}(R) = \frac{k_x^2}{R}\left[\frac{H_\rho^2}{R} + a\left(1 - \frac{q_{\rm p0}}{(1+R)^2}\right)\right] \tag{41}$$

is the potential with

$$a = \frac{u_{\rm rms}^2\beta_{\rm p}^2}{\lambda^2} \tag{42}$$

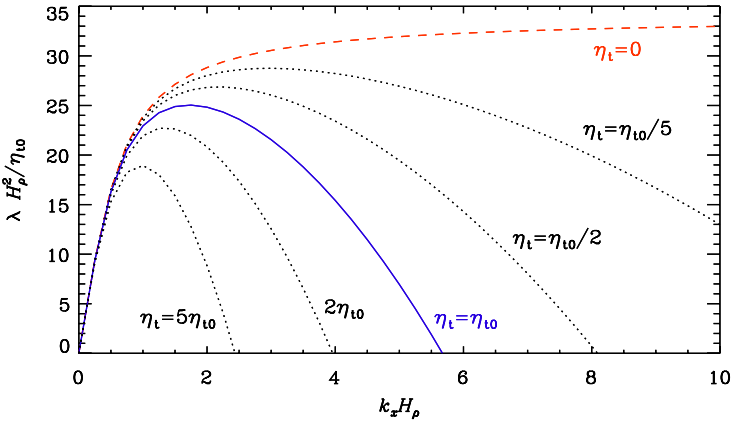


Figure 12 Growth rates obtained by solving Equation (38) for $B_0/B_{\rm eq0} = 0.1$ as a function of k_x for the case $\eta_{\rm t} = 0$ (dashed lines) compared with cases with different values of $\eta_{\rm t}$ (dotted lines). The solid line applies to the case $\eta_{\rm t} = \eta_{\rm t0}$.

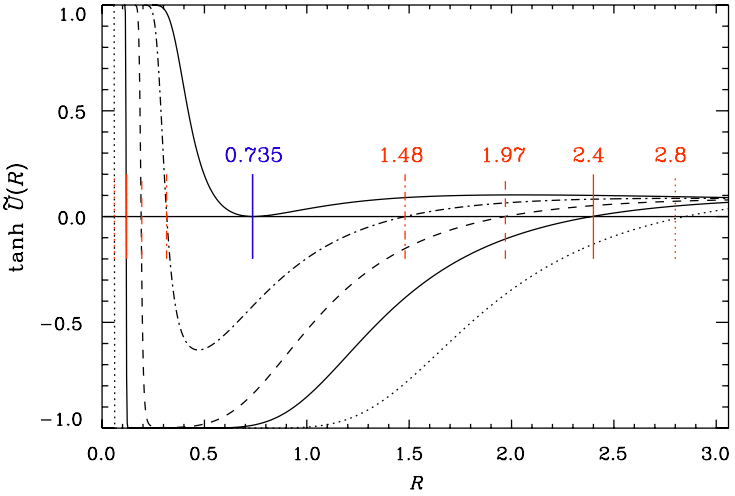


Figure 13 $\tanh\tilde{U}(R)$ for $a = 0.24$ (dash–triple–dotted line, $\tilde{\lambda} = 0.034$, $R_1 = R_2 = 0.735$), $a = 0.3$ [dash–dotted line, $\tilde{\lambda} = 0.030$, $(R_1, R_2) = (0.32, 1.5)$], $a = 0.4$ [dashed line, $\tilde{\lambda} = 0.026$, $(R_1, R_2) = (0.19, 1.9)$], $a = 0.56$ [solid line, $\tilde{\lambda} = 0.022$, $(R_1, R_2) = (0.12, 2.4)$], and $a = 1$ [dotted line, $\tilde{\lambda} = 0.017$, $(R_1, R_2) = (0.06, 2.8)$].

being a new eigenvalue. The potential $[\tilde{U}(R)]$ has the following asymptotic behavior: $\tilde{U}(R \to 0) = k_\perp^2 H_\rho^2/R^2$ and $\tilde{U}(R \to \infty) = a/R$. For the existence of the instability, the potential $\tilde{U}(R)$ should have a negative minimum. For example, for a long wavelength instability ($k_\perp^2 H_\rho^2 \ll 1$) and when $q_{\rm p0} > 1$, the potential $\tilde{U}(R)$ has a negative minimum, and the instability can be excited.

When the potential $\tilde{U}(R)$ has a negative minimum and since $\tilde{U}(R \to 0) > 0$ and $\tilde{U}(R \to \infty) > 0$, there are two points R_1 and R_2 (the so-called turning points) at which $\tilde{U}(R) = 0$. We have computed $\tilde{U}(R)$ for several values of a/H_ρ; see Figure 13, were we plot $\tanh\tilde{U}(R)$ to show more clearly the position of the turning points. Here we quote the nondimensional

values of $\tilde{\lambda} \equiv \lambda H_\rho / c_s$. It turns out for $a \approx 0.241\, H_\rho$ the turning points collapse to $R_1 = R_2 \approx 0.735$. The numerically computed maximum value of λ corresponds to $a \approx 0.56\, H_\rho$, where the turning points are $R_1 \approx 0.735$ and $R_2 \approx 2.4$. Using the equations $\tilde{U}(R_{1,2}) = 0$ together with Equations (41) and (42), we obtain the growth rate of the instability as

$$\lambda = \frac{\beta_\star u_{\rm rms}}{H_\rho} \frac{[R_1 R_2 (2 + R_1 + R_2)]^{1/2}}{(1 + R_1)(1 + R_2)}, \tag{43}$$

where we have used $\beta_\star = \beta_{\rm p} \sqrt{q_{\rm p0}}$. Note that Equation (43) is consistent with the simple estimate (12). In the critical case where $R = R_1 = R_2$, we find an upper bound of $\lambda \leq \sqrt{2} \beta_\star u_{\rm rms} R / H_\rho (1 + R)^{3/2}$.

References

Arlt, R., Sule, A., Rüdiger, G.: 2007, *Astron. Astrophys.* **461**, 295.
Bai, T.: 1987, *Astrophys. J.* **314**, 795.
Bai, T.: 1988, *Astrophys. J.* **328**, 860.
Benevolenskaya, E.E., Hoeksema, J.T., Kosovichev, A.G., Scherrer, P.H.: 1999, *Astrophys. J. Lett.* **517**, L163.
Bigazzi, A., Ruzmaikin, A.: 2004, *Astrophys. J.* **604**, 944.
Bogart, R.S.: 1982, *Solar Phys.* **76**, 155. doi:10.1007/BF00214137.
Brandenburg, A.: 2005, *Astrophys. J.* **625**, 539. doi:10.1086/429584.
Brandenburg, A., Kemel, K., Kleeorin, N., Mitra, D., Rogachevskii, I.: 2011, *Astrophys. J. Lett.* **740**, L50. doi:10.1088/2041-8205/740/2/L50.
Brandenburg, A., Kemel, K., Kleeorin, N., Rogachevskii, I.: 2012, *Astrophys. J.* **749**, 179. doi:10.1088/0004-637X/749/2/179.
Brandenburg, A., Kleeorin, N., Rogachevskii, I.: 2010, *Astron. Nachr.* **331**, 5. doi:10.1002/asna.201111638.
Cally, P.S., Dikpati, M., Gilman, P.A.: 2003, *Astrophys. J.* **582**, 1190.
Cattaneo, F., Hughes, D.W.: 1988, *J. Fluid Mech.* **196**, 323.
Charbonneau, P.: 2010, *Living Rev. Solar Phys.* **7**, 3. http://www.livingreviews.org/lrsp-2010-3.
Choudhuri, A.R., Gilman, P.A.: 1987, *Astrophys. J.* **316**, 788.
D'Silva, S., Choudhuri, A.R.: 1993, *Astron. Astrophys.* **272**, 621.
Epstein, T.: 1904, *Astron. Nachr.* **166**, 333.
Fan, Y.: 2009, *Living Rev. Solar Phys.* **6**, 4. http://www.livingreviews.org/lrsp-2009-4.
Golub, L., Vaiana, G.S.: 1980, *Astrophys. J. Lett.* **235**, L119.
Golub, L., Rosner, R., Vaiana, G.S., Weiss, N.O.: 1981, *Astrophys. J.* **243**, 309.
Hindman, B.W., Haber, D.A., Toomre, J.: 2009, *Astrophys. J.* **698**, 1749.
Hughes, D.W., Proctor, M.R.E.: 1988, *Ann. Rev. Fluid Dyn.* **20**, 187.
Ilonidis, S., Zhao, J., Kosovichev, A.: 2011, *Science* **333**, 993.
Isobe, H., Miyagoshi, T., Shibata, K., Yokoyama, T.: 2005, *Nature* **434**, 478.
Käpylä, P.J., Brandenburg, A., Kleeorin, N., Mantere, M.J., Rogachevskii, I.: 2012, *Mon. Not. Roy. Astron. Soc.* doi:10.1111/j.1365-2966.2012.20801.x, arXiv:1105.5785.
Kemel, K., Brandenburg, A., Kleeorin, N., Mitra, D., Rogachevskii, I.: 2012a, *Solar Phys.* doi:10.1007/s11207-012-9949-0, arXiv:1112.0279.
Kemel, K., Brandenburg, A., Kleeorin, N., Rogachevskii, I.: 2012b, *Astron. Nachr.* **333**, 95. doi:10.1002/asna.201111638.
Kersalé, E., Hughes, D.W., Tobias, S.M.: 2007, *Astrophys. J. Lett.* **663**, L113.
Kitchatinov, L.L., Mazur, M.V.: 2000, *Solar Phys.* **191**, 325. doi:10.1023/A:1005213708194.
Kitchatinov, L.L., Rüdiger, G.: 2005, *Astron. Nachr.* **326**, 379.
Kitiashvili, I.N., Kosovichev, A.G., Wray, A.A., Mansour, N.N.: 2010, *Astrophys. J.* **719**, 307.
Kleeorin, N., Mond, M., Rogachevskii, I.: 1996, *Astron. Astrophys.* **307**, 293.
Kleeorin, N., Rogachevskii, I.: 1994, *Phys. Rev. E* **50**, 2716.
Kleeorin, N.I., Rogachevskii, I.V., Ruzmaikin, A.A.: 1989, *Sov. Astron. Lett.* **15**, 274.
Kleeorin, N.I., Rogachevskii, I.V., Ruzmaikin, A.A.: 1990, *Sov. Phys. JETP* **70**, 878.
Kosovichev, A.G.: 2002, *Astron. Nachr.* **323**, 186.
Kosovichev, A.G., Stenflo, J.O.: 2008, *Astrophys. J. Lett.* **688**, L115.
Parker, E.N.: 1966, *Astrophys. J.* **145**, 811.
Parker, E.N.: 1979, *Cosmical Magnetic Fields*, Oxford University Press, New York.
Pipin, V.V., Kosovichev, A.G.: 2011, *Astrophys. J. Lett.* **727**, L45.

Rempel, M.: 2011a, *Astrophys. J.* **729**, 5.
Rempel, M.: 2011b, *Astrophys. J.* **740**, 15.
Rieutord, M., Zahn, J.-P.: 1995, *Astron. Astrophys.* **296**, 127.
Rogachevskii, I., Kleeorin, N.: 2007, *Phys. Rev. E* **76**, 056307.
Ruzmaikin, A.: 1998, *Solar Phys.* **181**, 1. doi:10.1023/A:1016563632058.
Sanford, F.: 1941, *Science* **94**, 18.
Schatten, K.H.: 2007, *Astrophys. J. Suppl.* **169**, 137.
Schüssler, M., Caligari, P., Ferriz-Mas, A., Moreno-Insertis, F.: 1994, *Astron. Astrophys. Lett.* **281**, L69.
Spruit, H.C.: 1974, *Solar Phys.* **34**, 277. doi:10.1007/BF00153665.
Stein, R.F., Lagerfjärd, A., Nordlund, Å., Georgobiani, D.: 2011, *Solar Phys.* **268**, 271.
Stein, R.F., Leibacher, J.: 1974, *Ann. Rev. Astron. Astrophys.* **12**, 407.
Stenflo, J.O., Kosovichev, A.G.: 2012, *Astrophys. J.* **745**, 129.
Tao, L., Weiss, N.O., Brownjohn, D.P., Proctor, M.R.E.: 1998, *Astrophys. J. Lett.* **496**, L39.
Tayler, R.J.: 1973, *Mon. Not. Roy. Astron. Soc.* **161**, 365.
Vitinskij, J.I.: 1969, *Solar Phys.* **7**, 210. doi:10.1007/BF00224899.
Wissink, J.G., Hughes, D.W., Matthews, P.C., Proctor, M.R.E.: 2000, *Mon. Not. Roy. Astron. Soc.* **318**, 501.
Zhao, J., Kosovichev, A.G., Duvall, T.L. Jr.: 2001, *Astrophys. J.* **557**, 384.

DOI 10.1007/978-1-4899-8005-2_20
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-013-0235-6

Infrared Observations from the *New Solar Telescope* at Big Bear

Philip R. Goode · Wenda Cao

Received: 20 February 2012 / Accepted: 21 January 2013 / Published online: 21 February 2013

Abstract The 1.6 m clear aperture solar telescope in Big Bear is operational and with its adaptive optics (AO) system it provides diffraction limited solar imaging and polarimetry in the near-infrared (NIR). While the AO system is being upgraded to provide diffraction limited imaging at bluer wavelengths, the instrumentation and observations are concentrated in the NIR. The *New Solar Telescope* (NST) operates in campaigns, making it the ideal ground-based telescope to provide complementary/supplementary data to SDO and *Hinode*. The NST makes photometric observations in Hα (656.3 nm) and TiO (705.6 nm) among other lines. As well, the NST collects vector magnetograms in the 1565 nm lines and is beginning such observations in 1083.0 nm. Here we discuss the relevant NST instruments, including AO, and present some results that are germane to NASA solar missions.

Keywords Observational near infrared astronomy

1. Overview

The 1.6 m clear aperture solar telescope ("NST" or *New Solar Telescope*) in Big Bear Solar Observatory had its first light in January 2009 with first science grade data in the Summer of 2009 and first data corrected by adaptive optics (AO) in the Summer of 2010. The NST is the first facility-class solar telescope built in the US in a generation. As shown in Figure 1, the NST is a modern, off-axis 1.6 m clear aperture instrument (Goode *et al.*, 2010; Cao *et al.*, 2010; http://www.bbso.njit.edu/nst_project.html) that offers a significant improvement in ground-based, high angular resolution and polarimetric capabilities. Early observations have been concentrated in the near infrared (NIR) because of limitations in AO, which

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

P.R. Goode (✉) · W. Cao
Big Bear Solar Observatory, New Jersey Institute of Technology, Big Bear City, CA, USA
e-mail: pgoode@bbso.njit.edu

W. Cao
e-mail: wcao@bbso.njit.edu

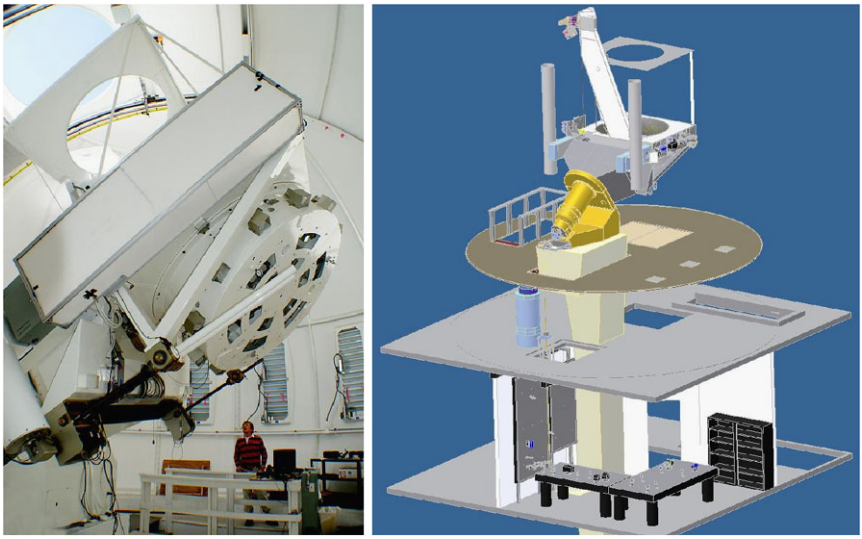


Figure 1 Left picture shows the 1.6 m off-axis NST with the primary mirror cell shown in the foreground. The design was driven by competing needs for rigidity atop the secondary mirror tower and an un-obscured light path throughout the telescope. Right schematic shows the full optical path through the NST with the bottom floor housing all instruments that are fed light from the AO.

Table 1 Milestones in NST commissioning.

Date	NST Milestone	Observations Enabled
January 2009	First Light with NST	Visible Light Photometry (0″.50)
Summer 2009	First Diffraction Limited Observations	Visible Light Photometry (0″.12)
Summer 2010	AO Corrected Observations	Visible & NIR Photometry (0″.10 – 0″.20)
Summer 2010	AO Corrected Magnetograms	NIR Vector **B** (0″.40)
Fall 2012	AO-308 Observations	Photometry in Bluest Light (0″.05)
Spring 2013	IRIM Upgrade to Dual FPI System	Higher Cadence Vector **B** in NIR (0″.20)
Spring 2013	VIM Installation	High Spatial Resolution Vector **B** (0″.10)
Summer 2013	First Observations with MCAO	AO Correction over Full 70″ × 70″ FOV
Fall 2013	CYRA First Light	Spectroscopy Covering 1.0 – 5.0 μm

are now being solved by larger format deformable mirrors (DMs), as will be discussed. Key milestones of NST commissioning are listed in Table 1. The telescope is well into its commissioning phase.

The telescope is configured as an off-axis Gregorian system consisting of a parabolic primary, prime focus field stop and heat reflector (heat-stop), elliptical secondary and diagonal flats. The primary mirror (PM) is 1.7 m with a clear aperture of 1.6 m. The focal ratio of the PM is f/2.4, and the final ratio is f/52. The 100″ circular opening in the field stop defines a 70″ × 70″ maximal square field of view (FOV). The working wavelength range covers 0.4 to 1.7 μm in the Coudé Laboratory two floors beneath the telescope and all wavelengths including far infrared are at the Nasmyth focus optical bench attached to the side of the telescope structure. In fact all wavelengths are also available at a focus immediately before the

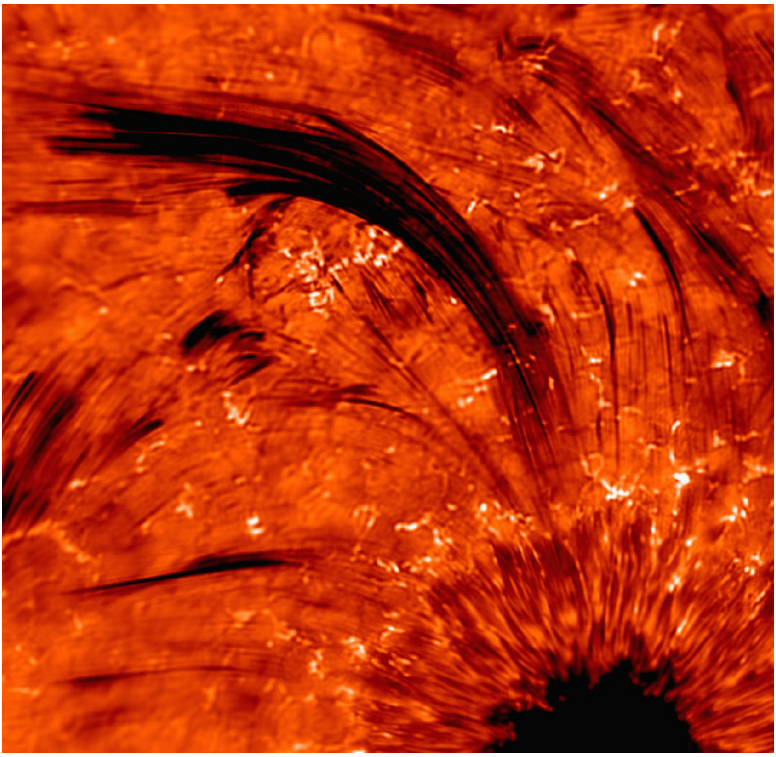


Figure 2 Off-band Hα −0.75 Å image taken on 2 July 2010 revealing the dynamical layer overlying the surface region above a sunspot. Most apparent are the long dark streaks, so-called "jets" arising from the bright magnetic regions on the edge of the penumbra.

light is fed to the Coudé Laboratory. An off-axis design was chosen principally because of its vastly reduced stray light, since there is no central obscuration to degrade the telescope's MTF (modulation transfer function) at high spatial frequencies.

The wavefront sensing system for alignment and for the PM active optics resides before M3. The polarization modulator also resides there. Having the modulator so far forward is a design advantage that helps us obtain vector magnetograms, which we were never able to do with the old 0.6 m telescope because, in part, there were always many (oblique) mirrors between the sunlight and the polarimeter's modulator.

The NST has a fully operational, nearly achromatic adaptive optics (AO) system on a vertical bench in the Coudé Laboratory for diffraction limited imaging, as illustrated in Figure 2. The Hα image in Figure 2 reveals the dynamical layer overlying the surface region. Most apparent are the long dark streaks, so-called "jets" arising from the bright magnetic regions on the edge of the penumbra. The image benefited from adaptive optics, which enabled real-time correction for atmospheric distortion. The AO system incorporates a 97 actuator deformable mirror, a Shack–Hartmann wave front sensor with 76 subapertures, and a digital signal processor system. This is structurally the same "AO-76" used on the now-retired 0.6 m BBSO telescope and the *Dunn Solar Telescope* (DST) of the National Solar Observatory (NSO). NST images also routinely undergo post-facto reconstruction. The Kiepenheuer-Institut für Sonnenphysik's software package for speckle interferometry of AO corrected solar data (KISIP) is a post-AO reconstruction algorithm allowing one to

achieve the diffraction limit of the telescope over a larger field of view than the isoplanatic patch for most observations (Wöger, von der Lühe, and Reardon, 2008).

The AO-76 system tracks on granulation for the NST, while with the old 0.6 m telescope we were only able to track on pores, which reflects to some extent the problems of retrofitting an AO system on a powerful telescope that was not designed to support AO. The success of AO-76 on the NST was no surprise, rather it was foreseen in a detailed and realistic error budget analysis (Rimmele, 2008, private communication) that included AO residuals, telescope and instrument error budgets, and the steady, good BBSO seeing (an AO correctable $r_0 \sim 9$ cm at 550 nm all day long during the summer observing season, which is easily sufficient for AO-76 to lock on granulation). This Fried parameter is in the mid-visible wavelength range and was determined from measurements of the atmospheric turbulence ("C_n^2") as a function of altitude above BBSO (Kellerer *et al.*, 2012). These are the first such measurements at a US solar observatory. The Fried parameter, r_0, is the coherence length of the atmospheric turbulence. For telescope apertures larger than r_0, resolution is seeing limited. In the summers of 2010 and 2011, AO was able to lock on granulation for extended periods of time. In the winter, attaining AO lock for extended periods was more problematic because of generally poorer seeing than in summer. The Strehl, generally used to quantify the performance of an AO system, is the ratio of the maximum intensity in the AO corrected image in the detector plane to that from a theoretical, perfect imaging system operating at the diffraction limit. The error budget analysis of Rimmele (2008, private communication) using the more pessimistic annual average $r_0 \sim 6$ cm (the 9 cm value comes primarily from summertime when the seeing is much better) from the ATST site survey (http://atst.nso.edu/site) in Big Bear showed that AO-76 will yield, in the detector plane, a high Strehl of about 0.7 in the near infrared (1 μm) under median, annualized BBSO seeing conditions. However, in the visible (0.5 μm), AO-76 will deliver a sufficient Strehl ($\sim$ 0.3, which would imply diffraction limited imaging of the solar disk) only under well-above average summer seeing conditions. However, most of the time the Strehl would be more like $\sim$0.1. Thus, in visible light our observations using AO-76 are better at the red end of the visible spectrum. The low Strehl matters especially for polarimetric observations, which range from difficult to impossible to interpret. The problem here is that the requisite opposite polarity measurements, which are subtracted from each other, are often closely (tenths of arcsecs) spaced together. Thus, truly diffraction limited observations, especially of the magnetic field will be rare in the visible spectrum, even though they would be critical for the highest possible spatial resolution.

Basically, to make AO corrections, at least one subaperture of the wavefront sensor (WFS) is required for every resolution element of atmospheric wavefront distortion. Roughly, if $r_0 \sim 10$ cm a 60 cm aperture telescope needs a WFS with $6 \times 6 = 36$ (36 subapertures) and the 97 actuator DM of AO-76 is more than sufficient for diffraction limited observations with a superb optical telescope. However, for the 1.6 m NST, $16 \times 16 = 256$ subapertures would be required. However, r_0 scales as $\lambda^{\frac{6}{5}}$, which corresponds to 20 cm in the NIR at 1 μm under the reasonable and standard assumption of Kolmogorov turbulence. In this case, the minimum WFS format would be $8 \times 8 = 64$, so AO-76 if fine. For Hα, AO-76 is inadequate, which means polarimetry in visible light is not possible except under extraordinary seeing conditions. However, imaging with speckle reconstruction can yield good images, like Figure 2. But G-band imaging would require a higher order AO system even for imaging. Thus, so far NST usage has been concentrated on the NIR. For these reasons, we have been working on a higher order AO system.

We have been developing, in collaboration with NSO, an AO system based on a commercially available 357 actuator DM (AO-308), which, based on the same kind of error budget analysis, we realistically expect to achieve a Strehl of 0.3 in the detector plane in the visible (0.5 μm) even for median, annualized BBSO seeing conditions (Rimmele, 2008, private communication).

The AO-308 project is well under way. We have purchased two 357 actuator DMs from Xinetics. Each has a special faceplate made of silicon instead of ULE (ultra-low expansion glass) because silicon has about 100 times the thermal conductivity of ULE to conduct away the heat on the DM from the Sun. The other elements of AO-308 are a more complex DSP system from Bittware (DSPs are about 10 times faster than those for AO-76) and a faster wavefront sensing camera (Phantom V7.3 from Vision Research). Each hardware element and its connections to its partners has been successfully tested. What remains is DSP control and the graphical user interface, with the former being built on the programming from AO-76. The optical design of AO-308 is close to that of AO-76 requiring only the replacement of two optical elements in the AO-76 feed. The AO-308 system is undergoing its commissioning phase as of this publishing and have full operation shortly thereafter.

With AO-308 feeding light to our optical instruments in the Coudé Laboratory, BBSO will collect a richer spectrum of data. The next, and on-going AO project is Multi-Conjugate Adaptive Optics (MCAO), which will follow shortly thereafter. In collaboration with NSO and with *Advanced Technology Solar Telescope* (ATST) in mind, we are developing a MCAO system for the NST that will expand the effective isoplanatic patch of diffraction limited observations in the field of view (FOV) by nearly an order of magnitude to at least 70″. The expansion of the diffraction limited FOV will be sufficient to cover entire active regions enabling photometric, spectroscopic and polarimetric observations, which include, for instance, flares that might occur at anytime and anywhere in an active region. Therefore, the MCAO system will eliminate a major limitation of conventional AO systems, the insufficiently large isoplanatic patch, and, thus provide the tool to effectively study, with the requisite high temporal cadence and with the NST's unprecedented spatial resolution, fundamental scientific questions like the onset and evolution of flares, flare triggering mechanisms, magnetic reconnection events and many other dynamic solar phenomena, which would be best performed with diffraction limited observations over an extended FOV. Such data will provide an excellent supplement/complement to SDO and *Hinode* data. Of course, for slow (about 5 s) cadence observations, the NST will use AO-76, then AO-308 combined with speckle reconstruction, while for appreciably higher cadence we will suffer some from the gradual roll-off in resolution as we move further from the isoplanatic patch. Of course, the NST is a 2 m^2 "light-bucket" that can be used to accumulate sufficient photons for very high temporal cadence observations of rapid phenomena in visible light, as well as in the NIR.

2. NIR Photometry and Polarimetry

The IRIM (*InfraRed Imaging vector Magnetograph*) system developed by BBSO is in operation (Cao *et al.*, 2011) on the NST using light fed from AO-76. We show the Q, U, and V components of a vector magnetogram in Figure 3. We note that severe polarization problems precluded obtaining vector magnetograms from the old, 0.6 m BBSO telescope and that BBSO vector magnetograms from that era were from a 20 cm aperture telescope. The superior optical design of the NST and good polarization calibration have already enabled us to obtain vector magnetograms with minimal cross-talk, as is clear from Figure 3 where

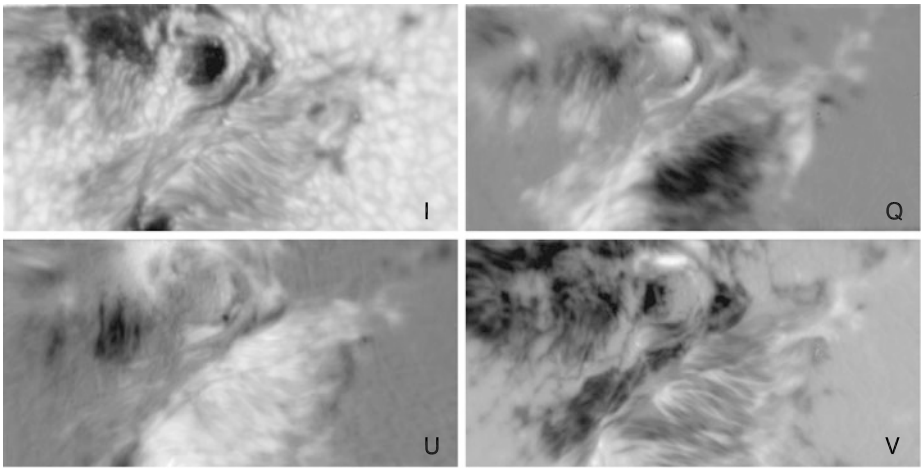


Figure 3 Calibrated IRIM vector magnetogram without any image reconstruction near AR 11283 at 463″W 119″E, taken on 7 September 2011. Stokes I and V are sliced in the blue wing of the Fe I 1564.85 nm line while Q and U are sliced in line center. Dual-beam technology was employed to minimize seeing-induced noise. The field of view of IRIM is about $50'' \times 25''$.

there is no apparent cross-talk of the line-of-sight component within the two transverse components.

The NST magnetograms in Figure 3 are from IRIM, which is one of the first imaging spectropolarimeters working at 1565 nm, and is used for observations of the Sun at its opacity minimum, exposing the deepest photospheric layers that can be seen. The temporal cadence for the IRIM vector magnetograms is 30 s, but it will be about five times faster when its Lyot filter is replaced by a second Fabry-Pérot Interferometer (FPI) in the Spring of 2012, so the second generation IRIM will be available shortly thereafter. IRIM's first imaging polarimetric observations at 1565 nm were made at the diffraction limit on 1 July 2005 using BBSO's retired 0.6 m telescope (Cao *et al.*, 2006). Stokes-V profiles were obtained from the FPI scan, and the data provide access to both the true magnetic field strength and the filling factor of the small-scale magnetic flux elements. The strength of IRIM is its extreme sensitivity and high spatial resolution, which allows us to study weak and small-scale magnetic fields in the quiet Sun. One might think that the quality of magnetograms from the NST with triple the aperture of the old telescope would be compensated by roughly tripling of the wavelength of the observations (ignoring the factor of eight difference in aperture of the NST and that of the old BBSO magnetograph). But this is not true, after all, Zeeman splitting increases quadratically with wavelength, so IRIM can more precisely detect magnetic flux. Thus, we will now be able to extract the true magnetic field strength and the filling factor for the small-scale magnetic flux elements (Cao *et al.*, 2006). Additionally, the terrestrial atmosphere is more benign in the NIR with the mean, annualized Fried parameter at BBSO being about 25 cm (under the standard assumption of Kolmogorov turbulence), which is four times that at 0.5 μm. Even with AO-76 corrected light, this much larger Fried parameter is essential in the observations of faint features, which otherwise would tend to drift in and out of sharpness. The focal plane Strehl in the NIR is sufficient to enable sustained diffraction limited images with AO-76 under typical BBSO observing conditions. Depending on the number of spectral positions chosen for the line scans in IRIM operations, we will have as short as a 30 s cadence for vector magnetograms. The expected polarization accuracy will

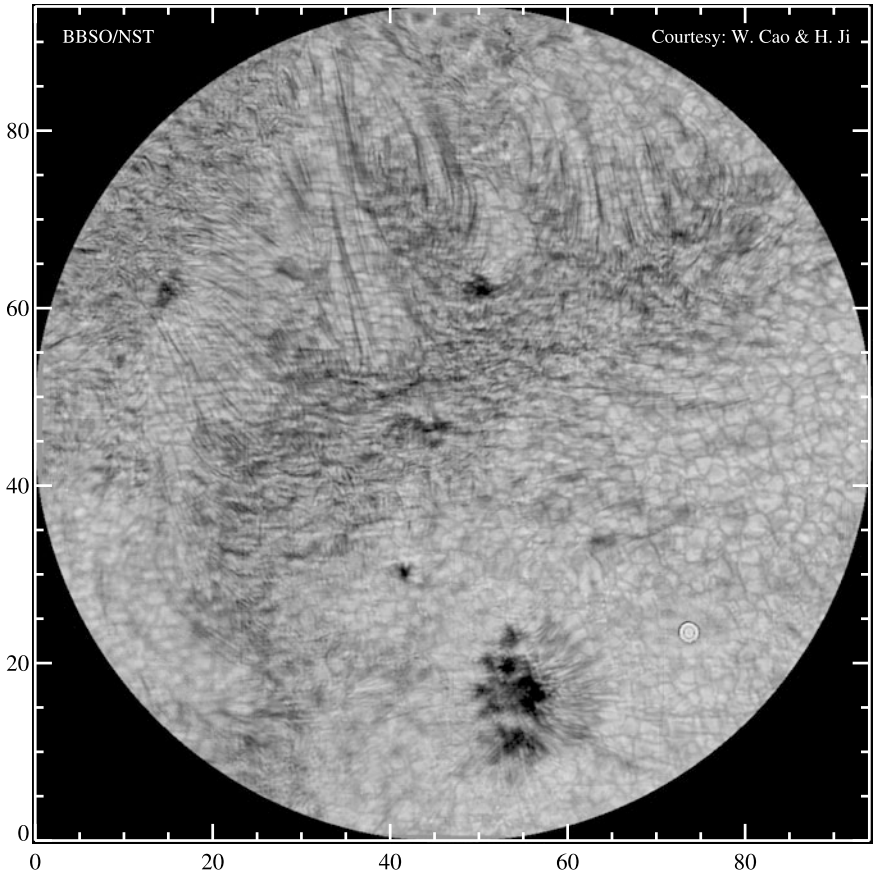


Figure 4 He I 10830 Å image take with IRIM on 22 July 2011 revealing a previously unknown, hyperfine magnetic loop structure in active region NOAA 11259. In this example, the loops primarily connect areas of opposite polarity. The field of view is $94'' \times 94''$.

be our nominal 10^{-3}. The spatial resolution will be around $0''.2$ for direct imaging, and $0''.4$ for magnetograms.

We are upgrading IRIM to a dual FPI system (replace Lyot filter with second FPI), so that we can fine tune to any wavelength between 1.0 and 1.7 μm, rather than being confined by the Lyot filter to the 1.6 μm regime with IRIM. The upgraded instrument is called NIRIS (NIR *Imaging Spectropolarimeter*). NIRIS (Cao *et al.*, 2012) will have all the capabilities of IRIM with five times the cadence. It has a broader wavelength coverage over the NIR that will enable NIRIS to measure the magnetic field in the upper chromosphere and base of the corona using the He I 10830 Å line. Such measurements have only been made with spectrographs, like SOLIS until now. NIRIS will be on-line in the Spring of 2013 (see Table 1).

With IRIM, we occasionally substitute a 0.5 Å narrow bandpass Lyot filter centered on 10830 Å. With this, we have obtained images like that in Figure 4 in which one can see a hyperfine system of magnetic loops reaching up to the base of the corona. The loops have a typical cross-section of about 100 km, which is about the same as the diffraction limit of the NST at 1 μm. The He I 10830 Å line is formed in the upper chromosphere and transition

region. A striking feature is that the thickness of the loops remains almost constant over long distances. We have also attained preliminary magnetograms from IRIM with the 10830 Å Lyot filter. With NIRIS, we will have about five times more light and a tighter bandpass.

The weakness of IRIM is the spatial resolution, but this problem will be solved using our *Visible Imaging Magnetograph* (VIM). VIM was in use on the old BBSO telescope and obtained spectroscopic data and line-of-sight magnetograms. IRIM and VIM have the same basic design and operational modes and both magnetograph systems were ready to be integrated into the NST at first light, but since fully diffraction limited operation of VIM will require AO-308, we chose to bring IRIM on line first. For the work described here, we will have AO-308 and VIM for science that requires vector magnetograms of the highest spatial resolution (0″.1).

2.1. Mid-IR Spectroscopy

The *CrYogenic infRAred Spectrograph* (CYRA) spanning 1.0 to 5.0 μm will help the NST achieve its scientific potential of a new and improved probing of the fundamentals of the Sun's atmosphere with its dynamic magnetic field, the origin of space weather. CYRA will be a substantial improvement over the two current solar IR spectrographs (Pierce, 1964; Penn *et al.*, 1991) – one operating at the National Solar Observatory (on the McMath-Pierce telescope) and the other at the Institute for Astronomy (Mees Solar Observatory, University of Hawaii), both of which are based on warm optics except for the detectors and order sorting filters, whereas CYRA will be fully cryogenic. CYRA will be a significant advance, particularly for high-spatial and high-cadence and high Zeeman sensitivity observations of the Sun's atmosphere from the photosphere through the chromosphere and into the corona, as well as observations of dimmer targets such as sunspot umbrae and off-limb features.

Acknowledgements We thank BBSO observing staff and instrument team for their support. W. Cao and P.R. Goode acknowledge the support of the US NSF (AGS-0847126 and AGS-0745744), NASA (NNX08BA22G), and AFOSR (FA2386-12-1-3018 and FA9550-09-1-0655).

References

Cao, W., Jing, J., Ma, J., Xu, Y., Wang, H., Goode, P.R.: 2006, *Publ. Astron. Soc. Pac.* **118**, 838.
Cao, W., Gorceix, N., Coulter, R., Ahn, K., Rimmele, T.R., Goode, P.R.: 2010, *Astron. Nachr.* **331**, 636.
Cao, W., Ahn, K., Goode, P.R., Shumko, S., Gorceix, N., Coulter, R.: 2011, *ASP Conf. Ser.* **437**, 345.
Cao, W., Goode, P.R., Ahn, K., Gorceix, N., Schmidt, W., Lin, H.: 2012, *ASP Conf. Ser.* **463**, 291.
Goode, P.R., Coulter, R., Gorceix, N., Yurchyshyn, V., Cao, W.: 2010, *Astron. Nachr.* **331**, 620.
Kellerer, A., Gorceix, N., Marino, J., Cao, W., Goode, P.R.: 2012, *Astron. Astrophys.* **542**, A2.
Penn, M.J., Mickey, D.L., Canfield, R.C., Labonte, B.J.: 1991, *Solar Phys.* **135**, 163.
Pierce, A.K.: 1964, *Appl. Opt.* **3**, 1337.
Wöger, F., von der Lühe, O., Reardon, K.: 2008, *Astron. Astrophys.* **488**, 3751.

DOI 10.1007/978-1-4899-8005-2_21
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0069-7

A Nonlinear Force-Free Magnetic Field Approximation Suitable for Fast Forward-Fitting to Coronal Loops. I. Theory

Markus J. Aschwanden

Received: 23 November 2011 / Accepted: 7 July 2012 / Published online: 2 August 2012

Abstract We derive an analytical approximation of nonlinear force-free magnetic field solutions (NLFFF) that can efficiently be used for fast forward-fitting to solar magnetic data, constrained either by observed line-of-sight magnetograms and stereoscopically triangulated coronal loops, or by 3D vector-magnetograph data. The derived NLFFF solutions provide the magnetic field components $B_x(\mathbf{x})$, $B_y(\mathbf{x})$, $B_z(\mathbf{x})$, the force-free parameter $\alpha(\mathbf{x})$, the electric current density $\mathbf{j}(\mathbf{x})$, and are accurate to second-order (of the nonlinear force-free α-parameter). The explicit expressions of a force-free field can easily be applied to modeling or forward-fitting of many coronal phenomena.

Keywords Sun: corona · Sun: magnetic fields

1. Introduction

The coronal magnetic field can be constrained in a number of ways, such as by extrapolation of photospheric magnetograms or vector-magnetograph data, by radio observations of gyroresonance layers above sunspots, of by coronal seismology of oscillating loops. Before the advent of the STEREO mission, attempts were made to model observed coronal loops with stretched potential field solutions (Gary and Alexander, 1999), to fit a linear force-free model with solar-rotation stereoscopy (Wiegelmann and Neukirch, 2002; Feng *et al.*, 2007), by tomographic reconstruction with magnetohydrostatic constraints (Wiegelmann and Inhester, 2003; Ruan *et al.*, 2008), by magnetic modeling applied to spectropolarimetric loop detections (Wiegelmann *et al.*, 2005), or by magnetic field supported stereoscopic loop triangulation (Wiegelmann and Inhester, 2006; Conlon and Gallagher, 2010). Recently, stereoscopic triangulation of coronal loops with the STEREO mission became

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

M.J. Aschwanden (✉)
Lockheed Martin Advanced Technology Center, Dept. ADBS, Solar and Astrophysics Laboratory, Bldg. 252, 3251 Hanover St., Palo Alto, CA 94304, USA
e-mail: aschwanden@lmsal.com

available, which constrains the 3D geometry of coronal magnetic field lines (Aschwanden *et al.*, 2008; Aschwanden, 2009). The plethora of coronal high-resolution data allows us now to compare different magnetic models and to test whether they are self-consistent. A critical assessment of nonlinear force-free field (NLFFF) codes revealed the disturbing fact that different NLFFF codes yield incompatible results among themselves, and exhibit significant misalignments with stereoscopically triangulated loops (DeRosa *et al.*, 2009; Sandman *et al.*, 2009; Aschwanden and Sandman, 2010; Sandman and Aschwanden, 2011; Aschwanden *et al.* 2012a, 2012b). The discrepancy was attributed to uncertainties in the boundary conditions as well as to the non-force-freeness of the photosphere and lower chromosphere. Earlier tests with the virial theorem already indicated that the magnetic fields in the lower chromosphere at altitudes of $h \lesssim 400$ km are not force-free (Metcalf *et al.*, 1995). Constraints by coronal tracers thus have become an important criterion to bootstrap a self-consistent magnetic field solution. The misalignment between theoretical extrapolation models and stereoscopically triangulated loops could be minimized by using potential field models with forward-fitted unipolar magnetic charges (Aschwanden and Sandman, 2010) or dipoles (Sandman and Aschwanden, 2011).

In this Paper we go a step further by deriving a simple analytical approximation of nonlinear force-free field solutions that is suitable for fast forward-fitting to stereoscopically triangulated loops or to some other coronal observations. While accurate solutions of force-free magnetic fields have been known for special mathematical functions (Low and Lou, 1990) that have been used to reconstruct the local twist of coronal loops (Malanushenko, Longcope, and McKenzie, 2009; Malanushenko, Yusuf, and Longcope, 2011), they are not suitable for forward-fitting to entire active regions. In contrast, our theoretical framework entails the representation of a potential or non-potential field by a superposition of a finite number of elementary field components that are associated with buried unipolar magnetic charges at arbitrary locations, each one being divergence-free and force-free to a good approximation, as we test numerically. While this Paper contains the analytical framework of the magnetic field model, the numerical forward-fitting code with applications to observations will be presented in a Paper II (Aschwanden and Malanushenko, 2012), and applications to stereoscopically observed active regions in Aschwanden *et al.* (2012a, 2012b).

2. Theory

2.1. Potential Field Parameterization

The simplest representation of a magnetic potential field that fulfills Maxwell's divergence-free condition ($\nabla \cdot \mathbf{B} = 0$) is a unipolar magnetic charge j that is buried below the solar surface, which predicts a magnetic field $\mathbf{B}_j(\mathbf{x})$ that points away from the buried unipolar charge and whose field strength falls off with the square of the distance r_j,

$$\mathbf{B}_j(\mathbf{x}) = B_j \left(\frac{d_j}{r_j} \right)^2 \frac{\mathbf{r}_j}{r_j}, \tag{1}$$

where B_j is the magnetic field strength at the solar surface above a buried magnetic charge, (x_j, y_j, z_j) is the subphotospheric position of the buried charge, d_j is the depth of the magnetic charge,

$$d_j = 1 - \sqrt{x_j^2 + y_j^2 + z_j^2}, \tag{2}$$

and $\mathbf{r}_j = [x - x_j, y - y_j, z - z_j]$ is the vector between an arbitrary location $\mathbf{x} = (x, y, z)$ in the solar corona (were we desire to calculate the magnetic field) and the location (x_j, y_j, z_j) of the buried charge. We choose a Cartesian coordinate system (x, y, z) with the origin in the Sun center and are using units of solar radii, with the direction of z chosen along the line-of-sight from Earth to Sun center. For a location near disk center ($x \ll 1, y \ll 1$), the magnetic charge depth is $d_j \approx (1 - z_j)$. Thus, the distance r_j from the magnetic charge is

$$r_j = \sqrt{(x - x_j)^2 + (y - y_j)^2 + (z - z_j)^2}. \tag{3}$$

The absolute value of the magnetic field $B_j(r_j)$ is simply a function of the radial distance r_j (with B_j and d_j being constants for a given magnetic charge),

$$B(r_j) = B_j \left(\frac{d_j}{r_j} \right)^2. \tag{4}$$

In order to obtain the Cartesian coordinates (B_x, B_y, B_z) of the magnetic field vector $\mathbf{B}_j(\mathbf{x})$, we can rewrite Equation (1) as

$$\begin{aligned} B_x(x, y, z) &= B_j (d_j / r_j)^2 (x - x_j)/r_j, \\ B_y(x, y, z) &= B_j (d_j / r_j)^2 (y - y_j)/r_j, \\ B_z(x, y, z) &= B_j (d_j / r_j)^2 (z - z_j)/r_j. \end{aligned} \tag{5}$$

We progress now from a single magnetic charge to an arbitrary number N_m of magnetic charges and represent the general magnetic field with a superposition of N_m buried magnetic charges, so that the potential field can be represented by the superposition of N_m fields $\mathbf{B}_j$ from each magnetic charge $j = 1, \ldots, N_\mathrm{m}$,

$$\mathbf{B}(\mathbf{x}) = \sum_{j=1}^{N_\mathrm{m}} \mathbf{B}_j(\mathbf{x}) = \sum_{j=1}^{N_\mathrm{m}} B_j \left(\frac{d_j}{r_j} \right)^2 \frac{\mathbf{r}_j}{r_j}. \tag{6}$$

As an example we show the representation of a dipole with two magnetic unipolar charges ($N_\mathrm{m} = 2$) of opposite polarity ($B_2 = -B_1$) in Figure 1. Each of the unipolar charges has a radial magnetic field (dotted lines), but the superposition of the two vectors of both unipolar charges in every point of space, $\mathbf{B}(\mathbf{x}) = \mathbf{B}_1(\mathbf{x}) + \mathbf{B}_2(\mathbf{x})$, reproduces the familiar dipole field. For the case shown in Figure 1 we used the parameterization of two subphotospheric unipolar magnetic charges at positions $x_1 = -0.5$ and $x_2 = +0.5$, which produces dipole-like field lines (solid curves), while they converge to the classical solution of a dipole field in the limit of $x_1 \mapsto 0$ and $x_2 \mapsto 0$, as can be shown analytically (Jackson, 1962).

2.2. Force-Free Field Solution of a Uniformly Twisted Fluxtube

A common geometrical concept is to characterize coronal loops with cylindrical fluxtubes. For thin fluxtubes, the curvature of coronal loops and the related forces can be neglected, so that a cylindrical geometry can be applied. Because the footpoints of coronal loops are anchored in the photosphere, where a random velocity field creates vortical motion on the coronal fluxtubes, they are generally twisted. We consider now such twisted fluxtubes in a cylindrical geometry and derive a relation between the helical twist and the force-free parameter α. The analytical solution of a uniformly twisted flux tube is described in several textbooks (*e.g.* Gold and Hoyle, 1960; Priest, 1982; Sturrock, 1994; Boyd and Sanderson, 2003; Aschwanden, 2004), but we summarize the derivation here to provide physical insights for the generalized derivation of nonlinear force-free magnetic field solutions derived in Section 2.3 in a self-consistent notation.

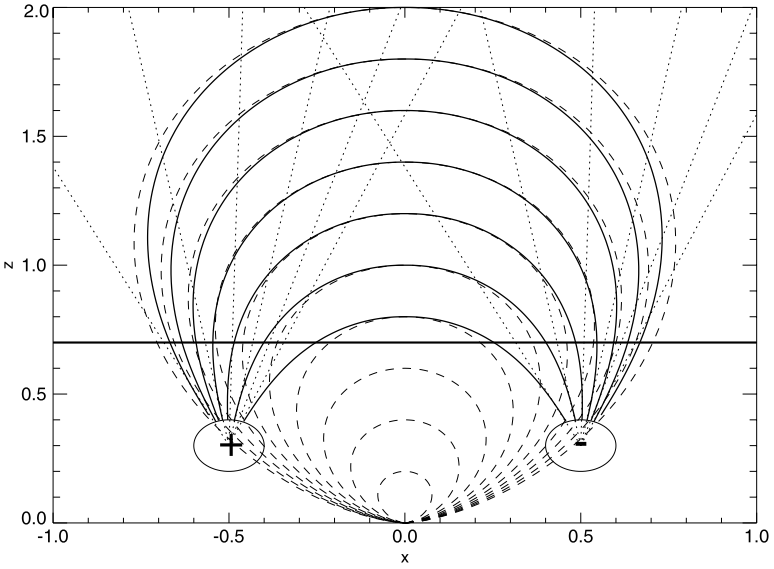


Figure 1 The magnetic field of a symmetric dipole (dashed lines) is shown, together with the field resulting from the superposition of two unipolar magnetic charges (solid lines). The two field models become identical once the two unipolar charges are moved towards the location of the dipole moment at position $(x, y) = (0, 0)$. The radial field of each unipolar (positive and negative) charge is also shown for comparison (dotted lines).

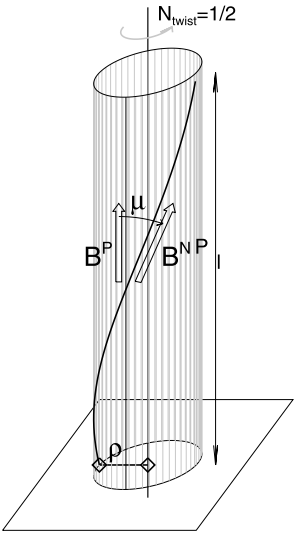


Figure 2 The basic 3D geometry of a cylindrical flux tube with uniform twist is defined by the length l of the cylinder axis, the number of twisting turns along this length, $N_{\rm twist}$, or by the misalignment angle μ at the flux tube radius ρ between the potential field line $\mathbf{B}^{\rm P}$ (aligned with the cylindrical axis) and the non-potential field line $\mathbf{B}^{\rm NP}$ (aligned with the twisted loop). The non-potential field line $\mathbf{B}^{\rm NP}$ can be decomposed into a longitudinal field component B_s and an azimuthal field component B_φ.

We consider a straight cylinder where a uniform twist is applied, so that an initially straight field line $\mathbf{B} = (0, 0, B_s)$, aligned with a field line coordinate s, is rotated by a number $N_{\rm twist}$ of full turns over the cylinder length l, yielding an azimuthal field component B_φ at radius ρ,

$$\frac{B_\varphi}{B_s} = \frac{\rho \partial \varphi}{\partial s} = \frac{2\pi \rho N_{\rm twist}}{l} = b\rho, \tag{7}$$

with the constant b defined in terms of the number of full twisting turns $N_{\rm twist}$ over a (loop) length l.

The cylindrical geometry of a twisted flux tube is visualized in Figure 2. The longitudinal component of the untwisted magnetic field corresponds to a potential field vector $\mathbf{B}^{\rm P}$, while the twisted non-potential field line $\mathbf{B}^{\rm NP}$ has a helical geometry with an angle μ at a radius ρ, which can be described by the longitudinal component B_s and the azimuthal component B_φ. The fluxtube can be considered as a sequence of cylinders with radii ρ, each one twisted by the same twist angle $\partial\varphi/\partial s = 2\pi N_{\rm twist}/l$. For uniform twisting, the magnetic components B_φ and B_s depend only on the radius ρ, but not on the length coordinate s or azimuth angle φ. Thus, the functional dependence in cylindrical coordinates (ρ, φ, s) is

$$\mathbf{B} = [B_\rho, B_\varphi, B_s] = \big[0, B_\varphi(\rho), B_s(\rho)\big]. \tag{8}$$

Consequently, the general expression of $\nabla \times \mathbf{B}$ in cylindrical coordinates,

$$\nabla \times \mathbf{B} = \left[\frac{1}{\rho}\frac{\partial B_s}{\partial \varphi} - \frac{\partial B_\varphi}{\partial s}, \frac{\partial B_\rho}{\partial s} - \frac{\partial B_s}{\partial \rho}, \frac{1}{\rho}\left(\frac{\partial}{\partial \rho}(\rho B_\varphi) - \frac{\partial B_\rho}{\partial \varphi}\right)\right], \tag{9}$$

is simplified with $B_\rho = 0$ and the sole dependencies of $B_\varphi(\rho)$ and $B_s(\rho)$ on the radius ρ (Equation (7)), yielding a force-free current density $\mathbf{j}$ of

$$\mathbf{j} = [j_\rho, j_\varphi, j_s] = \frac{c}{4\pi}(\nabla \times \mathbf{B}) = \frac{c}{4\pi}\left[0, -\frac{\partial B_s}{\partial \rho}, \frac{1}{\rho}\left(\frac{\partial}{\partial \rho}(\rho B_\varphi)\right)\right]. \tag{10}$$

Requiring that the Lorentz force is zero for a force-free solution, $\mathbf{F} = \mathbf{j} \times \mathbf{B} = 0$, we obtain a single non-zero component in the radial ρ-direction, since $j_\rho = 0$ and $B_\rho = 0$ for the two other components,

$$\mathbf{F} = \mathbf{j} \times \mathbf{B} = [B_s j_\varphi - B_\varphi j_s, 0, 0], \tag{11}$$

yielding a single differential equation for B_s and B_φ,

$$B_s \frac{dB_s}{d\rho} + B_\varphi \frac{1}{\rho}\frac{d}{d\rho}(\rho B_\varphi) = 0. \tag{12}$$

By substituting $B_\varphi = b\rho B_s$ from Equation (7) into Equation (12), this simplifies to

$$\frac{d}{d\rho}\big[\big(1 + b^2\rho^2\big)B_s\big] = 0. \tag{13}$$

A solution is found by making the expression inside the derivative to a constant (B_0), which yields B_φ and B_s,

$$\mathbf{B} = [B_\rho, B_\varphi, B_s] = \left[0, \frac{B_0\, b\rho}{1 + b^2\rho^2}, \frac{B_0}{1 + b^2\rho^2}\right]. \tag{14}$$

[This equation also corrects a misprint in Equation (5.5.8) of Aschwanden (2004), where a superfluous zero component has to be eliminated.] With the definition of the force-free α-parameter,

$$(\nabla \times \mathbf{B}) = \frac{4\pi}{c}\mathbf{j} = \alpha(\rho)\mathbf{B}, \tag{15}$$

we can now verify that the α-parameter for a uniformly twisted fluxtube depends only on the radius ρ,

$$\alpha(\rho) = \frac{2b}{(1 + b^2\rho^2)}, \tag{16}$$

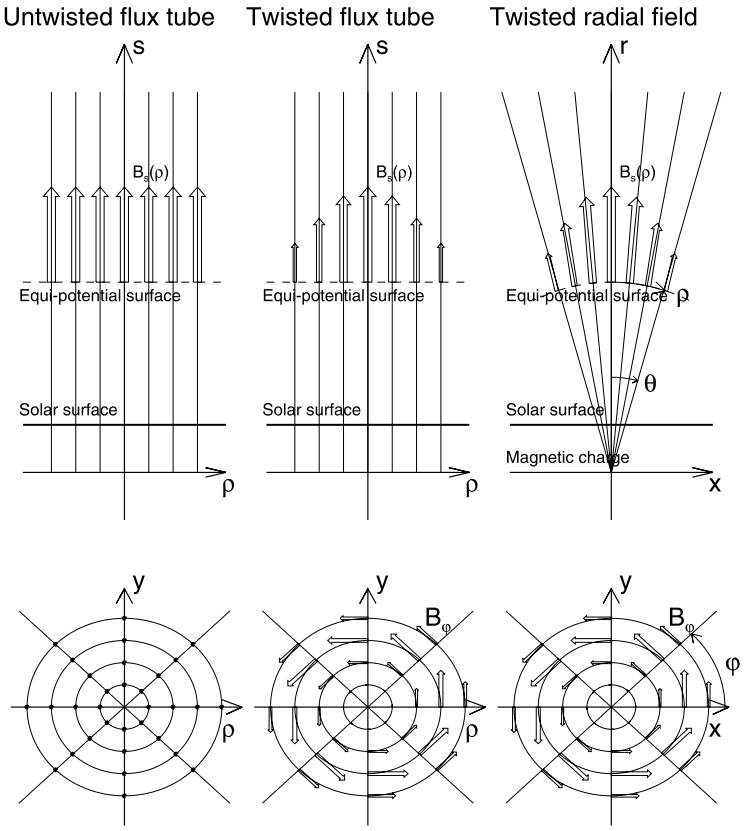


Figure 3 The field line geometry is shown for an untwisted cylindrical flux tube (left), a twisted cylindrical flux tube (middle), and for a twisted radial field (right), from the side view in the xz-plane (top) and from the top view in the xy-plane (bottom). The top panels show the longitudinal magnetic field component $B_s(\rho)$ and the bottom panels show the azimuthal magnetic field component $B_\varphi(\rho, \varphi)$.

with the constant b defined in terms of the number of full twisting turns $N_{\rm twist}$ over a (loop) length l (see Equation (7)),

$$b = \frac{2\pi N_{\rm twist}}{l}. \tag{17}$$

The geometry of a twisted flux tube is visualized in Figure 3 (top middle), where the parallel field lines are aligned with the coordinate axis s in the vertical direction, the cross-sectional radius ρ is defined in the direction perpendicular to s, and the twist angle φ is indicated in the horizontal projection (Figure 3, bottom middle). According to Equations (8) and (14), the variations of the longitudinal $B_s(\rho)$ and of the azimuthal component $B_\varphi(\rho)$ with radius ρ are

$$B_s(\rho) = \frac{B_0}{1 + b^2\rho^2}, \tag{18}$$

$$B_\varphi(\rho) = \frac{B_0 b\rho}{1 + b^2\rho^2}. \tag{19}$$

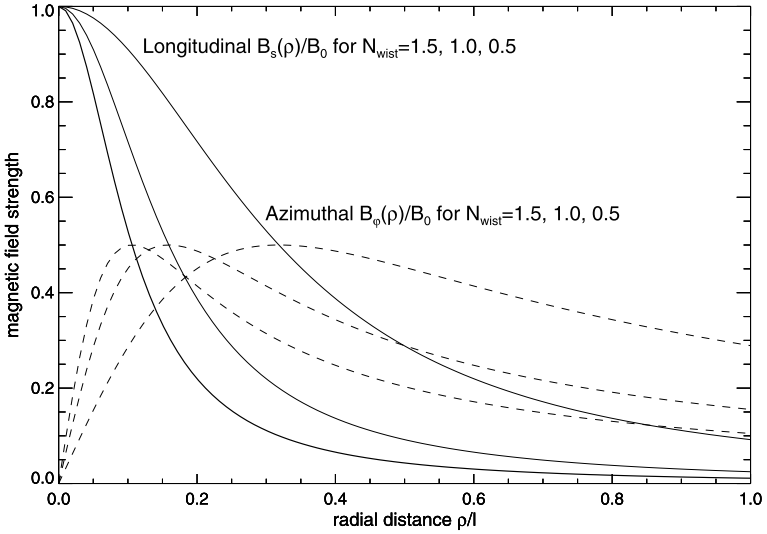


Figure 4 The dependence of the longitudinal (solid lines) and azimuthal magnetic field component (dashed lines) as a function of the distance r/l from the twist axis field is shown for three different amounts of twist ($N_{\text{twist}} = 1.5, 1.0, 0.5$ full turns per loop length l).

These radial dependencies are shown in Figure 4 for different numbers of twist ($N_{\text{twist}} = 0.5, 1.0, 1.5$). In the limit of vanishing twist ($N_{\text{twist}} = 0 \mapsto b = 0$), we have an untwisted flux tube (Figure 3, left) with a constant longitudinal field $B_s(\rho) = B_0$ and a vanishing azimuthal component $B_\varphi(\rho) = 0$. The dependence of the azimuthal field component $B_\varphi(\rho)$ and the longitudinal field component $B_s(\rho)$ as a function of the radius ρ from the twist axis (Figure 4) shows that the longitudinal component falls off monotonically with radius ρ, while the azimuthal component increases first for small distances $\rho \ll l$, but falls off at larger distances. Thus, the twisting causes a smaller cross-section of a fluxtube compared with the potential field situation, as widely known (*e.g.* Klimchuk, Antiochos, and Norton, 2000).

2.3. Nonlinear Force-Free Field Parameterization

We are now synthesizing the concept of point-like buried magnetic charges that we used to parameterize a potential field (Section 2.1) with the uniformly twisted flux tube concept that represents an exact solution of a nonlinear force-free field (Section 2.2). The geometric difference between the two concepts is the spherical symmetry of a point charge versus the parallel field configuration of an untwisted flux tube. However, we can synthesize the two geometries by considering the parallel field as a far-field approximation of a radial field. In an Euclidean parallel field, the equi-potential surface is a plane perpendicular to the parallel field vector, while a radial field has spherical equi-potential surface. We can make the transformation of a parallel field in cylindrical coordinates (s, ρ, φ) into a radial field with spherical coordinates (r, θ, φ) by mapping (see Figure 3)

$$\begin{aligned} s &\mapsto r, \\ \rho &\mapsto r \sin(\theta). \end{aligned} \tag{20}$$

This transformation from cylindrical to spherical coordinates preserves the orthogonality of the longitudinal field component ($B_s \mapsto B_r$) to the equi-potential surface ($s = \text{const} \mapsto$

$r = \text{const}$) and conserves the magnetic flux $\Phi(r)$ along a bundle of field lines with area $A(r) = \rho^2(r)$,

$$\Phi(r) = B(r)A(r) = B(r)\rho^2(r) = B(r)r^2 \sin^2\theta = \text{const}, \tag{21}$$

if the longitudinal component $B(r) \propto r^{-2}$ (Equation (1)) decreases quadratically with distance from the magnetic charge. Thus, applying the transformation into spherical coordinates (Equation (20)) and the magnetic flux conservation (Equation (21)) to the straight flux tube solution (Equations (18) and (19)), we can already guess the approximate nonlinear force-free solution in spherical coordinates,

$$B_r(r,\theta) \propto r^{-2} \frac{1}{(1 + b^2 r^2 \sin^2\theta)}, \tag{22}$$

$$B_\varphi(r,\theta) \propto r^{-2} \frac{br\sin\theta}{(1 + b^2 r^2 \sin^2\theta)}. \tag{23}$$

More rigorously, we can derive a nonlinear force-free field solution by writing the divergence-free condition $(\nabla \cdot \mathbf{B}) = 0$ and the force-free condition $(\nabla \times \mathbf{B}) = (4\pi/c)\mathbf{j} = \alpha(\rho)\mathbf{B}$ (Equation (15)) of a magnetic field vector $(B_r, B_\theta, B_\varphi)$ in spherical coordinates (r, θ, φ) (with the origin at the location of the magnetic charge and the spherical symmetry axis aligned with the vertical direction to the local solar surface),

$$(\nabla \cdot \mathbf{B}) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 B_r\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(B_\theta \sin\theta) + \frac{1}{r\sin\theta}\frac{\partial B_\varphi}{\partial\varphi} = 0, \tag{24}$$

$$[\nabla \times \mathbf{B}]_r = \frac{1}{r\sin\theta}\left[\frac{\partial}{\partial\theta}(B_\varphi \sin\theta) - \frac{\partial B_\varphi}{\partial\varphi}\right] = \alpha B_r, \tag{25}$$

$$[\nabla \times \mathbf{B}]_\theta = \frac{1}{r}\left[\frac{1}{\sin\theta}\frac{\partial B_r}{\partial\varphi} - \frac{\partial}{\partial r}(r B_\varphi)\right] = \alpha B_\theta, \tag{26}$$

$$[\nabla \times \mathbf{B}]_\varphi = \frac{1}{r}\left[\frac{\partial}{\partial r}(r B_\theta) - \frac{\partial B_r}{\partial\theta}\right] = \alpha B_\varphi. \tag{27}$$

For a simple approximative nonlinear force-free solution we require axi-symmetry with no azimuthal dependence ($\partial/\partial\varphi = 0$) and neglect components that contribute only to second order ($B_\theta \propto [br\sin\theta]^2 \approx 0$), in analogy to the uniformly twisted flux tubes on cylindrical surfaces (Figure 2). This requirement simplifies Equations (24) – (27) to

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 B_r\right) \approx 0, \tag{28}$$

$$\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(B_\varphi \sin\theta) = \alpha B_r, \tag{29}$$

$$-\frac{1}{r}\frac{\partial}{\partial r}(r B_\varphi) \approx 0, \tag{30}$$

$$-\frac{1}{r}\frac{\partial B_r}{\partial\theta} \approx \alpha B_\varphi. \tag{31}$$

Eliminating α from Equations (29) and (31) and using the analog *ansatz* as for cylindrical fluxtubes (Equation (7)),

$$B_\varphi = B_r br\sin\theta, \tag{32}$$

we obtain a similar differential equation as in Equation (13),

$$\frac{\partial}{\partial\theta}\left[B_r\left(1 + b^2 r^2 \sin^2\theta\right)\right] = 0. \tag{33}$$

A solution of this differential equation is obtained by setting the expression inside the bracket to the constant $B_0(d^2/r^2)$, which fulfills the divergence-free condition (Equation (28)), and we obtain a solution for B_r and B_φ (using Equation (32)), for α (using Equation (29)),

$$B_r(r,\theta) = B_0\left(\frac{d^2}{r^2}\right)\frac{1}{(1+b^2r^2\sin^2\theta)}, \tag{34}$$

$$B_\varphi(r,\theta) = B_0\left(\frac{d^2}{r^2}\right)\frac{br\sin\theta}{(1+b^2r^2\sin^2\theta)}, \tag{35}$$

$$B_\theta(r,\theta) \approx 0, \tag{36}$$

$$\alpha(r,\theta) \approx \frac{2b\cos\theta}{(1+b^2r^2\sin^2\theta)}. \tag{37}$$

This solution fulfills both the force-free condition (Equations (29) – (31)) and the divergence-free condition (Equation (28)) to second-order accuracy ($\propto [br\sin\theta]^2$). We see that this solution is identical with the simplified derivation of Equations (22) and (23). At locations near the twist axis ($\theta \mapsto 0$), the general solution (Equations (34) – (37)) converges to the cylindrical flux tube geometry solution (Equations (18) and (19)). Furthermore, in the limit of vanishing twist ($b \mapsto 0$) we retrieve the potential-field solution (Equation (4)), since the force-free parameter becomes $\alpha \mapsto 0$, the azimuthal field component becomes $B_\varphi = 0$, and the radial component reproduces the potential-field solution $B_r \mapsto B_0(d^2/r^2)$.

2.4. Cartesian Coordinate Transformation

In the derivation in the last section we derived the solution in terms of spherical coordinates (r,θ,φ) in a coordinate system where the rotational symmetry axis is aligned with the vertical to the solar surface intersecting a magnetic charge j. Since we are going to model a number of magnetic charges at arbitrary positions on the solar disk, we have to transform an individual coordinate system (r_j,θ_j,φ_j) associated with magnetic charge j into a Cartesian coordinate system (x,y,z) that is given by the observers line-of-sight (in z-direction) and the observer's image coordinate system (x,y) in the plane-of-sky. The variables of the Cartesian coordinate transformation are shown in Figure 5.

The radial magnetic field vector $\mathbf{B}_r$ (which is pointing radially away from a magnetic charge j located in the solar interior at (x_j,y_j,z_j) is simply given by the difference of the Cartesian coordinates from an arbitrary location (x,y,z),

$$\frac{\mathbf{B}_r}{B_r} = \left[\frac{x-x_j}{r_j}, \frac{y-y_j}{r_j}, \frac{z-z_j}{r_j}\right] = [\cos_{r,x}, \cos_{r,y}, \cos_{r,z}], \tag{38}$$

where B_r is the absolute value of the radial magnetic field component $B_r(r_j,\theta_j)$ (Equation (34)), r_j is the spatial length of the radial vector $\mathbf{r}_j$ (Equation (3)), defining the directional cosines $\cos_{r,i}$ (for the 3D coordinates $i = x,y,z$) of the radial magnetic field vector $\mathbf{B}_r$.

The azimuthal component $\mathbf{B}_\varphi$ (with the absolute value $B_\varphi(r_j,\theta_j)$ defined in Equation (35)) of the twisted magnetic field is orthogonal to the direction of the twist axis $\mathbf{R}$ (aligned with the local vertical),

$$\mathbf{R} = [x_j, y_j, z_j], \tag{39}$$

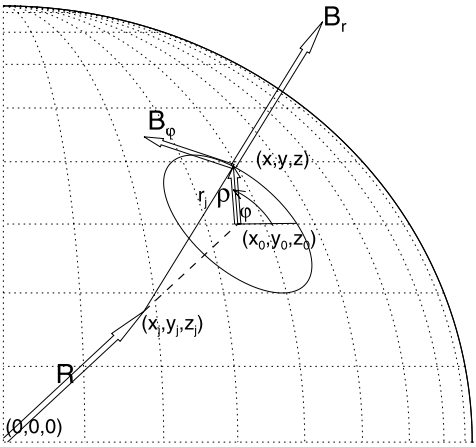


Figure 5 The geometry of a twisted radial field of a magnetic charge j buried at a subphotospheric position (x_j, y_j, z_j) is shown. The central twist axis (dashed line) intersects an equi-potential surface at position (x_0, y_0, z_0) and the longitudinal field vector B_r at position (x, y, z) has a radial distance ρ from the twist axis and an azimuth angle φ. The azimuthal magnetic field component B_φ at location (x, y, z) is orthogonal to the radial vector ρ and the longitudinal field component B_r, as well as to the direction of the twist axis $\mathbf{R}$.

and the radial magnetic field component $\mathbf{B}_r$ (Figure 5), and thus can be computed from the vector product of the two vectors $\mathbf{B}_r$ and $\mathbf{R}$,

$$\frac{\mathbf{B}_\varphi}{B_\varphi} = \frac{\mathbf{R} \times \mathbf{B}_r}{|\mathbf{R} \times \mathbf{B}_r|} = [\cos_{\varphi,x}, \cos_{\varphi,y}, \cos_{\varphi_z}], \tag{40}$$

which defines the directional cosines $\cos_{\varphi,i}$ of the azimuthal component in the Cartesian coordinate system. The vector product allows us also to extract the inclination angle θ_j between the radial magnetic field component $\mathbf{B}_r$ and the local vertical direction $\mathbf{R}$,

$$\theta_j = \sin^{-1}\left(\frac{|\mathbf{R} \times \mathbf{B}_r|}{|\mathbf{R}||\mathbf{B}_r|}\right). \tag{41}$$

Finally, the total non-potential magnetic field vector $\mathbf{B} = (B_x, B_y, B_z)$ is then the vector sum of the radial $\mathbf{B}_r$ and the azimuthal magnetic field component $\mathbf{B}_\varphi$,

$$\begin{aligned} B_x &= B_r(r_j, \theta_j)\cos_{r,x} + B_\varphi(r_j, \theta_j)\cos_{\varphi,x} \\ B_y &= B_r(r_j, \theta_j)\cos_{r,y} + B_\varphi(r_j, \theta_j)\cos_{\varphi,y} \\ B_z &= B_r(r_j, \theta_j)\cos_{r,z} + B_\varphi(r_j, \theta_j)\cos_{\varphi,z}, \end{aligned} \tag{42}$$

with the directional cosines $(\cos_{r,i}, \cos_{\varphi,i}, \cos_{\theta,i})$ defined by Equations (38) and (40). This is a convenient parameterization that allows us directly to calculate the magnetic field vector of the non-potential field $\mathbf{B}_j = (B_x, B_y, B_z)$ associated with a magnetic charge j that is characterized with five parameters: $(B_j, x_j, y_j, z_j, \alpha_j)$, where we define the force-free α-parameter from the twist parameter $b_j = 2\pi N_{\rm twist}/l$ (Equation (7)) at the location of the twist axis ($\theta_j = 0$),

$$\alpha_j = \alpha(\theta_j = 0) = 2\, b_j, \tag{43}$$

according to Equation (37).

2.5. Superposition of Twisted Field Components

The total non-potential magnetic field from all $j = 1, \ldots, N_\mathrm{m}$ magnetic charges can be approximately obtained from the vector sum of all components j (in an analog way as we applied in Equation (6) for the potential field),

$$\mathbf{B}(\mathbf{x}) = \sum_{j=1}^{N_\mathrm{m}} \mathbf{B}_j(\mathbf{x}), \tag{44}$$

where the vector components $\mathbf{B}_j = (B_{x,j}, B_{y,j}, B_{z,j})$ of the non-potential field of a magnetic charge j are defined in Equation (42), which can be parameterized with $5N_\mathrm{m}$ free parameters $(B_j, x_j, y_j, z_j, \alpha_j)$ for a non-potential field, or with $4N_\mathrm{m}$ free parameters for a potential field (with $\alpha_j = 0$). Of course, the sum of force-free magnetic field vectors is generally not force-free, but we will prove in the following (Equations (46) and (47)) that the sum of NLFFF solutions of the form of Equations (34) – (37), which are force-free to second-order accuracy in α (or, more strictly, in $[br\sin\theta]$), have the property that their sum is also force-free to second-order in α.

Let us first consider the condition of divergence-freeness. Since the divergence operator is linear, the superposition of a number of divergence-free fields is divergence-free also,

$$\nabla \cdot \mathbf{B} = \nabla \cdot \left(\sum_j \mathbf{B}_j \right) = \sum_j (\nabla \cdot \mathbf{B}_j) = 0. \tag{45}$$

While the divergence-free condition is exactly fulfilled for a potential field solution (Equation (4)), our quasi-force-free approximation (Equations (34) – (37)) matches this requirement to second order in α, as the insertion of the solutions (Equations (34) – (37)) into the divergence expression (Equation (24)) shows. For a quantitative measure of this level of accuracy we can also check numerical tests of the figure of merit (Section 3.3).

Now, let us consider the condition of force-freeness. A force-free field has to satisfy Maxwell's equation (Equation (15)). Since we parameterized both the potential field and the non-potential field with a linear sum of N_m magnetic charges, the requirement would be,

$$\nabla \times \mathbf{B} = \nabla \times \sum_{j=1}^{N_\mathrm{m}} \mathbf{B}_j = \sum_{j=1}^{N_\mathrm{m}} (\nabla_j \times \mathbf{B}_j) = \sum_{j=1}^{N_\mathrm{m}} \alpha_j(\mathbf{r}) \mathbf{B}_j = \alpha(\mathbf{r}) \mathbf{B}. \tag{46}$$

Generally, these three equations of the vector $\nabla \times \mathbf{B}$ cannot be fulfilled with a scalar function $\alpha(\mathbf{r})$ for a sum of force-free field components, unless the magnetic field volume consists of spatially separated force-free subvolumes. However, we can show the validity of the force-freeness equation (Equation (46)) to second-order accuracy in α. Note that the nonlinear force-free parameter α is proportional to b (Equations (37) and (43)), which is defined in Equation (17), and thus we set second-order accuracy in b equal to second-order accuracy in α. The argument goes as follows. If we use spherical coordinates, the NLFFF solution of the radial component is of zeroth order, $B_r(r,\theta) \propto O(\alpha^0)$ (Equation (34)), the azimuthal component is of first order, $B_\varphi(r,\theta) \propto O(\alpha^1)$ (Equation (35)), and the neglected third component magnetic field component is of second-order, $B_\theta(r,\theta) \propto O(\alpha^2)$ (as can be shown by inserting B_r and B_φ into Equation (26)). The curl of the magnetic field (Equations (25) – (27)) is then of first order for the radial component, $[\nabla \times \mathbf{B}]_r \propto \alpha B_r \propto O(\alpha^1)$ (Equation (25)), to second order for the azimuthal component, $[\nabla \times \mathbf{B}]_\varphi \propto (\alpha B_\varphi) \propto O(\alpha^2)$ (Equation (27)), and the remaining third component is of third-order, $[\nabla \times \mathbf{B}]_\theta \propto (\alpha B_\theta) \propto O(\alpha^3)$ (Equation (26)).

Therefore, if we neglect second-order and higher-order terms, the divergence-free condition (Equation (46)), which generally has three equations for the three curl components, *e.g.* $[\nabla \times \mathbf{B}]_r$, $[\nabla \times \mathbf{B}]_\varphi$, $[\nabla \times \mathbf{B}]_\theta$, reduces to one single equation for the radial component, $[\nabla \times \mathbf{B}]_r$, which can be fulfilled with a scalar function $\alpha(\mathbf{r})$,

$$\alpha(\mathbf{r}) \approx \frac{[\nabla \times \mathbf{B}]_r}{B_r} = \frac{[\nabla \times \sum_{j=1}^{N_\mathrm{m}} \mathbf{B}_j]_r}{\sum_{j=1}^{N_\mathrm{m}} \mathbf{B}_{j,r}}$$
$$= \frac{[\sum_{j=1}^{N_\mathrm{m}} \nabla \times \mathbf{B}_j]_r}{\sum_{j=1}^{N_\mathrm{m}} \mathbf{B}_{j,r}} = \frac{\sum_{j=1}^{N_\mathrm{m}} \alpha_j B_j}{\sum_{j=1}^{N_\mathrm{m}} B_j}. \quad (47)$$

Thus, we expect that the force-freeness is fulfilled to second-order accuracy $O(\alpha^2)$ (or strictly speaking $O(b^2)$). We will demonstrate the near force-freeness of simulated examples in the next section.

3. Simulations and Tests

We are now going to simulate examples of the analytical nonlinear force-free solutions in order to visualize the magnetic topology and to quantify the accuracy of the divergence-free and force-free conditions.

3.1. Numerical Examples

The simplest case is a single magnetic charge $j = 1$, which we illustrate as case A in Figure 6 (top row). We choose the following parameters: a magnetic field strength of $B_1 = 1000$ G (gauss) at the solar surface directly above the buried charge, the location $(x_1, y_1, z_1) = (0.1, 0.0, 0.95)$ for the buried charge, and a number of zero twist $b_1 = 0$ for the potential field case. We show the simulated line-of-sight magnetogram $B_z(x, y)$ in Figure 1 (top left), which mimics an isolated sunspot. The pixel size of the magnetogram and the stepping size in the extrapolation along a field line is $\Delta s = 0.004$ solar radii (2800 km $\approx 4''$, corresponding to the pixel size of SoHO/MDI magnetograms). We extrapolate the field lines for every pixel that has a footpoint magnetic field strength above a threshold of 50 % (>500 G). The field lines point in radial direction away from the center of the buried magnetic charge, as is expected for the potential field of an isolated sunspot (and defined in Equation (1)).

The next basic example is a magnetic dipole, which can be represented in our model by a superposition of a pair of two magnetic charges with opposite polarity, as sketched in Figure 1. The case B shown in Figure 6 is simulated with equal, but oppositely signed magnetic field strengths ($B_1 = 1000$ G, $B_2 = -1000$ G) at mirrored positions ($x_1 = 0.1, x_2 = -0.1$), otherwise we used the same parameters as in case A ($y_1 = y_2 = 0.0, r_1 = r_2 = 0.95, b_2 = b_1 = 0.0$). The magnetic field lines mimic the familiar structure of a dipole, which is parameterized here with eight free parameters (in the potential case).

A quadrupolar configuration is simulated in case C (Figure 6, bottom), with translational symmetry ($x_1 = 0.1, x_2 = 0.05, x_3 = -0.05, x_4 = -0.1$; $y_1 = 0.1, y_2 = 0.05, y_3 = 0.1, y_4 = 0.05$), equal depths ($r_1 = r_2 = r_3 = r_4 = 0.95$), and alternating field strengths ($B_1 = B_3 = 1000, B_2 = B_4 = -1000$ G). The quadrupolar configuration shows essentially two bipoles, each one with field lines that mostly connect within the same dipole domain, but a few intermediate field lines actually connect from one to the other domain.

In Figure 7 we show the same three configurations as for the potential field model (A, B, and C of Figure 6), but add electric currents caused by twisting, corresponding to

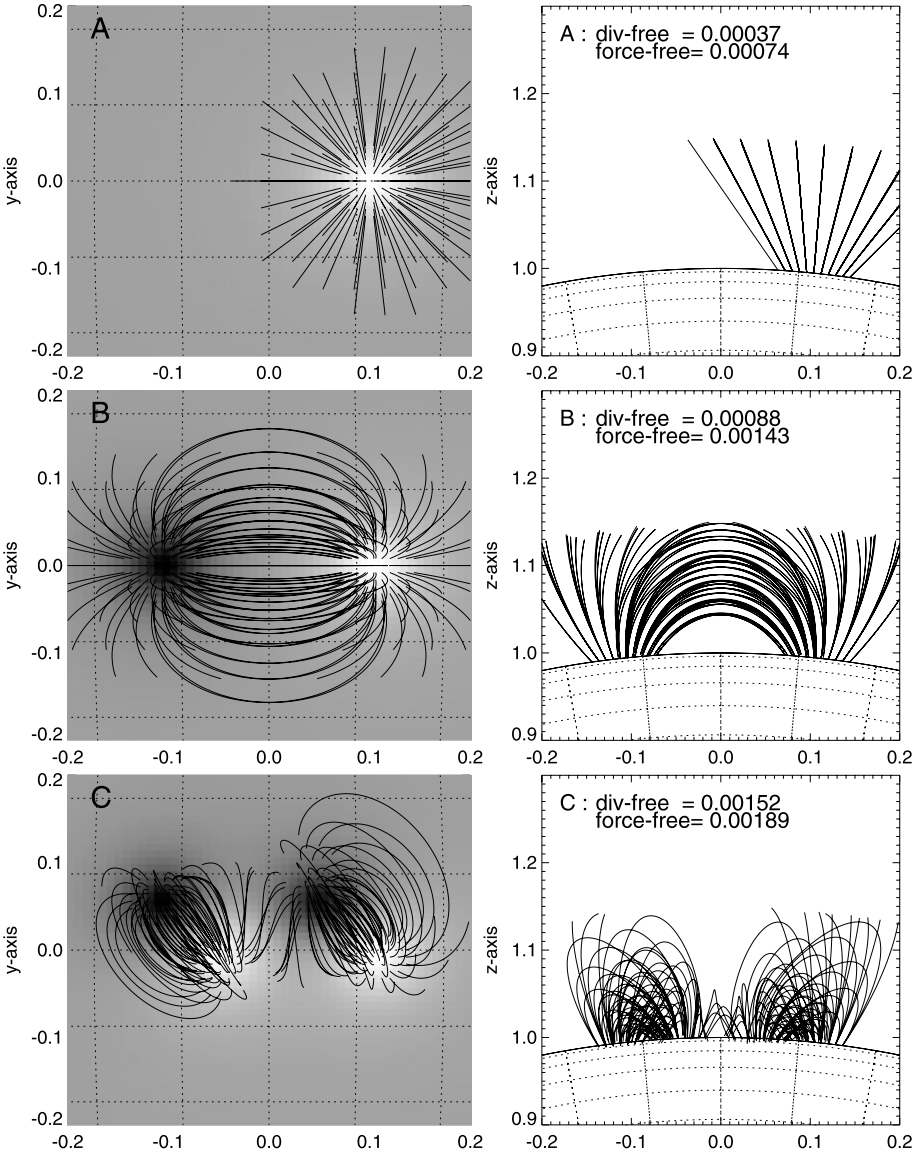


Figure 6 Simulations of three line-of-sight magnetograms (left) and magnetic field lines projected into the x–y plane (left) and into the vertical x–z plane (right). The three cases include: (A) a single positive magnetic charge (first row), (B) a dipole produced by two magnetic charges with opposite polarity (second row), and (C) a quadrupole configuration (third row). See parameters in Table 1. Only field lines with magnetic fields above a 50 % threshold of the maximum field strength are shown.

$N_{\rm twist} = -0.5$ turns for the single charge (case D) or first dipole (case E), and $N_{\rm twist} = 1.0$ for the second dipole (case F), defined for a loop length of $L = 0.1\pi$ solar radii. These amounts of twist correspond to force-free α-parameters of $\alpha = 2\pi N_{\rm twist}/L = -10$ and -20 solar radius^{-1} (*i.e.*, $\alpha = -1.43$ and -2.86×10^{-10} cm^{-1}). Comparing the potential (Figure 6) and non-potential cases (Figure 7) shows clearly the differences that result from the

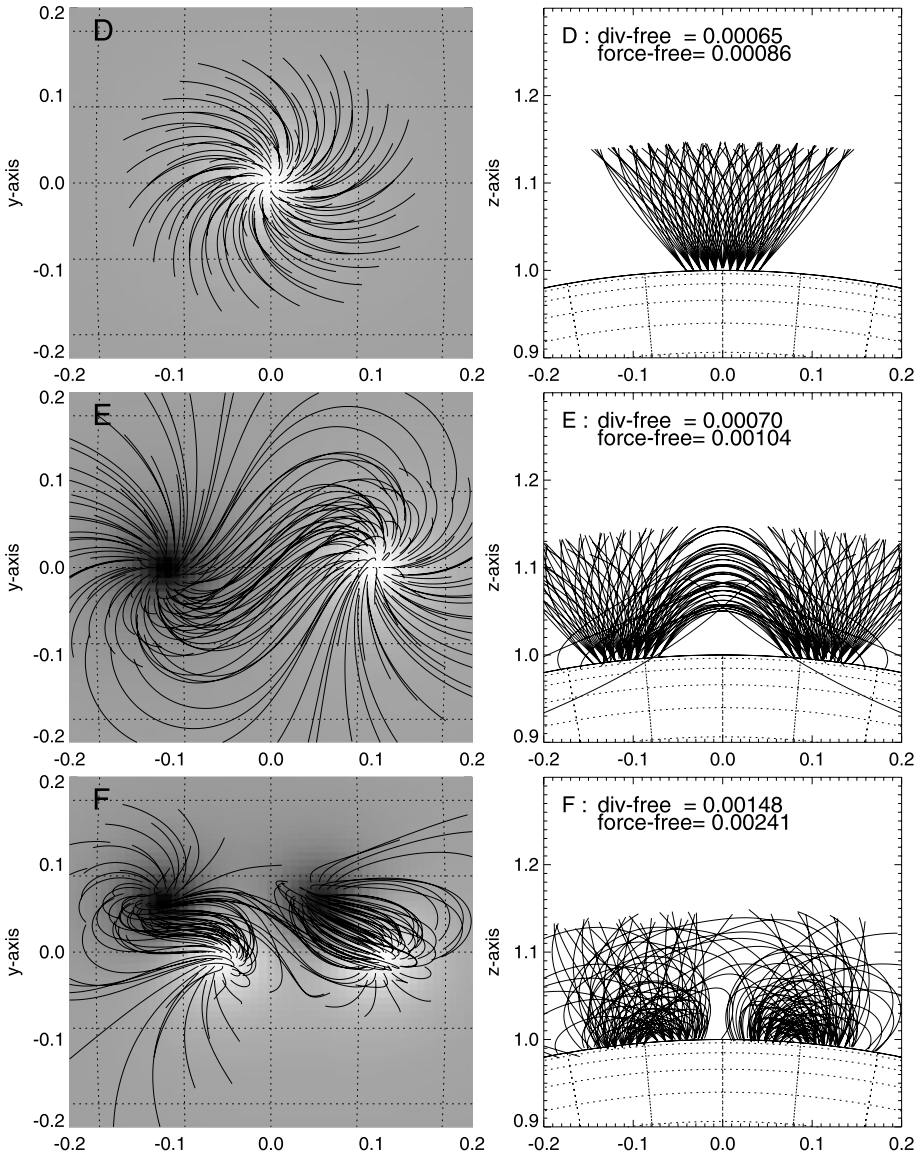


Figure 7 Simulations of three line-of-sight magnetograms (left) and magnetic field lines of a non-potential model with currents are shown, projected into the $x-y$ plane (left) and into the vertical $x-z$ plane (right). The parameters of the three cases (D), (E), and (F) are identical to thous of (A), (B), and (C), except for the addition of electric currents.

presence of electric currents. The force-free field lines of a sunspot become distorted into spiral shapes (case D), the straight dipole becomes distorted into a sigmoid shape (case E), and the quadrupolar configuration becomes also more distorted with sigmoid-like structures (case F).

In Figure 8 we show a few more complicated cases (G, H, and I), consisting of $N_{\rm m} = 10$ magnetic charges, with random values chosen in the magnetic field range -1000 G $<$

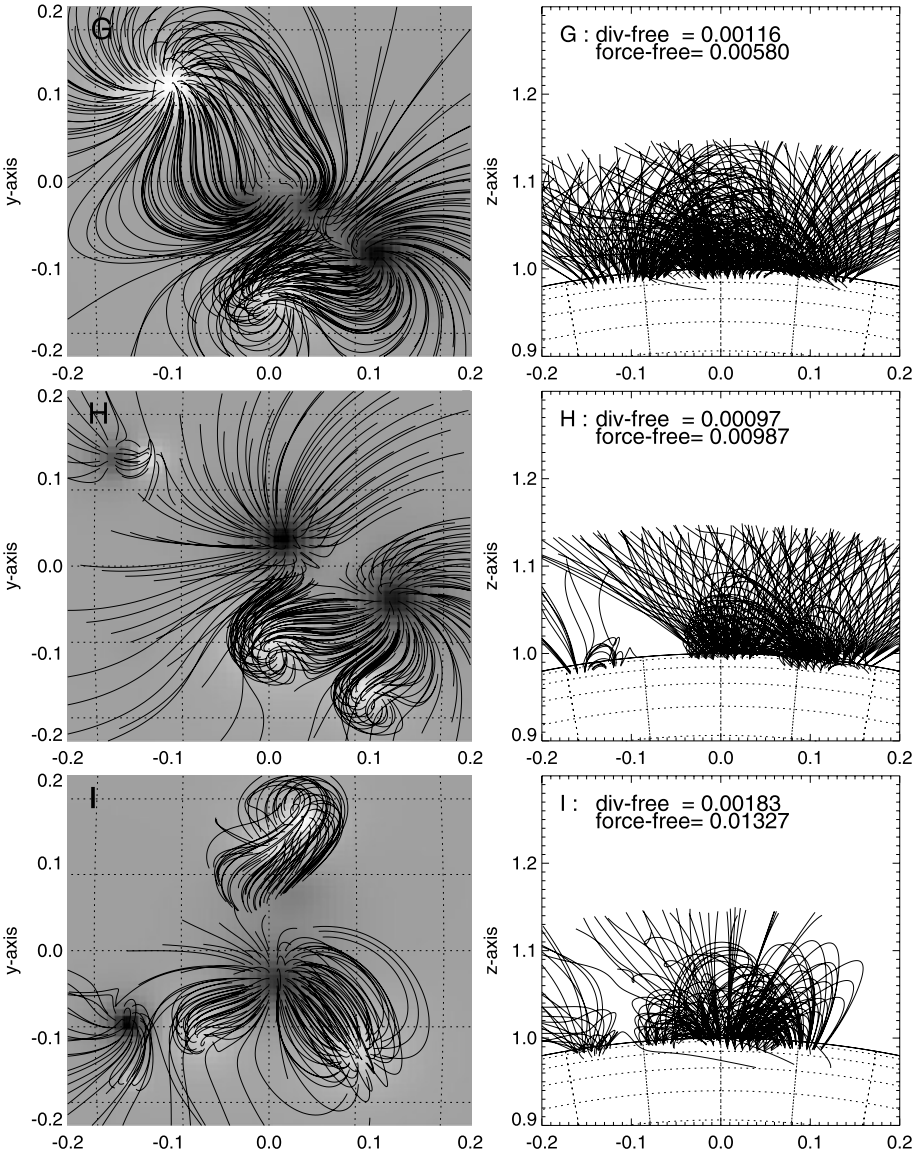


Figure 8 Simulations of three line-of-sight magnetograms (left) and magnetic field lines of a non-potential model with currents are shown, projected into the x–y plane (left) and into the vertical x–z plane (right). The three cases (G), (H), and (I) have each $N_{\rm m} = 10$ magnetic charges, with randomly chosen field strengths, locations, and electric currents.

$B_j < +1000$ G, in positions $-0.15 < x_j < 0.15$ solar radii, $-0.15 < y_j < 0.15$ solar radii, $0.95 < r_j < 0.97$ solar radii, and random twist in the range $-3 < N_{\rm twist} < +3$ per $L = 0.1\pi$ solar radii. The field lines displayed in Figure 8 demonstrate that a rich variety of sigmoid-shaped dipoles and inter-connecting multi-pole configurations can be generated with our quasi-force-free solutions, which mimic realistic active regions observed in the solar corona.

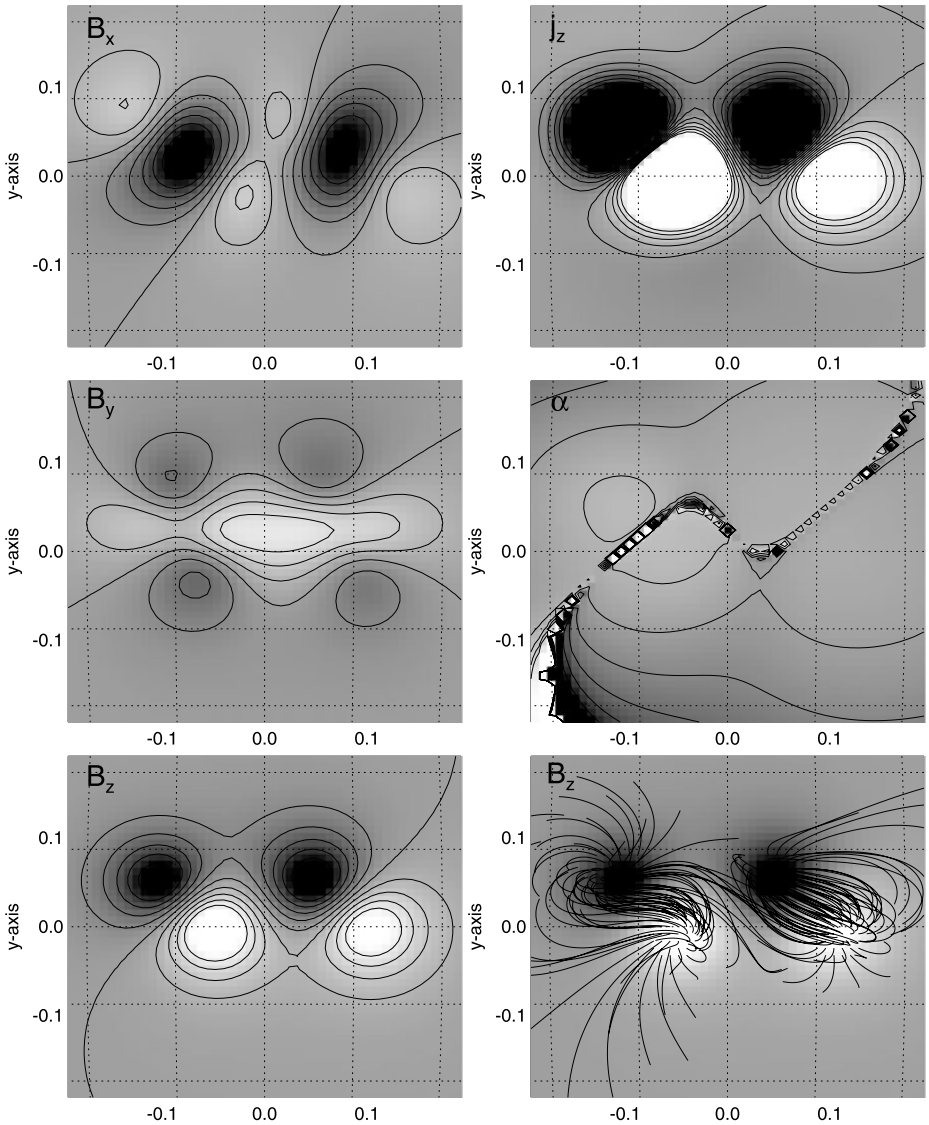


Figure 9 Maps of the magnetic field components $B_x(x, y)$, $B_y(x, y)$, $B_z(x, y)$ (left panels), the electric current density $j_z(x, y)$, and the force-free α-parameter (right panels).

3.2. Force-Free α-Parameter and Electric Current Maps

In Figure 9 we show examples of various maps that can be generated to visualize a 3D vector field solution, for the case F of a quadrupolar configuration with currents. We show the following quantities in the image plane $(x, y, z = 1 + \Delta s)$, which corresponds to an image plane near the solar surface: The three magnetic field vector component maps $B_x(x, y)$, $B_y(x, y)$, $B_z(x, y)$ (Figure 9, left panels, the vertical electric current map $j_z(x, y)$ (Figure 9, top right panel), the nonlinear α-parameter $\alpha(x, y)$ (Figure 9, middle right panel), and the

Table 1 Figures of merit for nine simulations of nonlinear force-free field solutions, detailing the size of the 3D data cube, the number of magnetic charges (N_m), potential or non-potential model (P and NP), the number of computed field lines N_f, the divergence-freeness L_d, the force-freeness L_f, and the computation times t_{CPU}.

Case	Data cube	Magnetic charges N_m	Field lines N_f	Divergence-freeness L_d	Force-freeness L_f	Computation time t_{CPU} (s)
A	$51 \times 51 \times 37$	1 (P)	87	0.0004	0.0007	0.078
B	$51 \times 51 \times 37$	2 (P)	160	0.0009	0.0014	0.309
C	$51 \times 51 \times 37$	4 (P)	159	0.0015	0.0019	0.351
D	$51 \times 51 \times 37$	1 (NP)	87	0.0006	0.0009	0.083
E	$51 \times 51 \times 37$	2 (NP)	160	0.0007	0.0010	0.314
F	$51 \times 51 \times 37$	4 (NP)	159	0.0015	0.0024	0.414
G	$51 \times 51 \times 37$	10 (NP)	336	0.0012	0.0058	2.462
H	$51 \times 51 \times 37$	10 (NP)	302	0.0010	0.0099	1.764
I	$51 \times 51 \times 37$	10 (NP)	217	0.0018	0.0133	1.370

LOS magnetogram $B_z(x, y)$ together with extrapolated field lines (Figure 9, bottom right panel). The B_z map shows most clearly the locations of the four buried magnetic charges that form two dipolar or a quadrupolar configuration. The magnetic polarization is also reflected in the j_z and α-map. The B_z and the α-map show also the location of the neutral line, where numerical effects due to the limited spatial resolution become visible.

3.3. Figures of Merit

The degree of convergence towards a divergence-free magnetic field model solution can be quantified by a measure that compares the average divergence $\nabla \cdot \mathbf{B}$, which should be close to zero, with the gradient $B/\Delta x$ of the magnetic field over a reference length scale Δx, for instance a pixel of the computational grid. The average deviation can then be defined by (see also Wheatland, Sturrock, and Roumeliotis (2000) or Schrijver *et al.* (2006))

$$L_d = \frac{1}{V} \int_V \frac{|(\nabla \cdot \mathbf{B})|^2}{|B/\Delta x|^2} \, dV. \tag{48}$$

Similarly, the force-freeness can be quantified by the ratio of the Lorentz force, $(\mathbf{j} \times \mathbf{B}) = (\nabla \times \mathbf{B}) \times \mathbf{B}$ to the normalization constant $B^2/\Delta x$,

$$L_f = \frac{1}{V} \int_V \frac{|(\nabla \times \mathbf{B}) \times \mathbf{B}|^2}{|B^2/\Delta x|^2} \, dV, \tag{49}$$

where $B = |\mathbf{B}|$.

We calculated these figure of merit quantities for the nine cases simulated in Figures 6–9. The values are listed in each of the panels in Figures 6–8 and listed in Table 1. The potential-field cases (A, B, and C) are found to have a figure of merit in the range of $L_d = 0.0009 \pm 0.0006$ for the divergence-freeness, and $L_f = 0.0014 \pm 0.0006$ for the force-freeness. The non-potential field cases (D, E, F, G, H, and I) have values in similar ranges of $L_d = 0.0009 \pm 0.0005$ for the divergence-freeness, and $L_f = 0.0100 \pm 0.0080$ for the force-freeness. We find no tendency that this figure of merit depends on the number of magnetic charges or some other model parameters. The fact that our quasi-force free analytical solutions perform equally well as standard NLFFF codes described in Schrijver *et al.*

(2006) tells us that the inaccuracy of the analytical approximation is commensurable or even smaller than the numerical uncertainty of other NLFFF codes. However, since our analytical solution provides an explicit formulation of nonlinear force-free fields, it can be computed much faster than the standard NLFFF codes, and still provides approximate solutions with acceptable accuracy (to second order). The computation time of the analytical solutions for the cases shown in Figures 6 – 8 amounts to about 1 s (on a recent Mac computer: Mac OS X, 2 × 3.2 GHz Quad-Core Intel Xeon, Memory 32 GB 800 MHz DDR2 FB-DIMM), while standard iterative NLFFF codes need several hours to converge to a single NLFFF solution.

4. Discussion and Conclusions

The coronal magnetic field has generally been computed by extrapolation from lower boundary data in form of photospheric magnetograms $B_z(x, y, z = z_{\rm ph})$ or vector-magnetograph data $\mathbf{B}(x, y)$, using a numerical extrapolation algorithm that fulfills the conditions of force-freeness $(\nabla \cdot \mathbf{B})$ and divergence-freeness $\nabla \times \mathbf{B} = \alpha(\mathbf{r})\mathbf{B}$, where $\alpha(\mathbf{r})$ is a scalar function in space $\mathbf{r}$. These extrapolation algorithms are very computing-intensive, because a good solution requires many iterations on a large computational 3D-grid that has sufficient spatial resolution to resolve the relevant magnetic field gradients. The accuracy of these numerical solutions depends very much on the noise in boundary vector magnetic field data as well as on deviations of photospheric fields from a force-free state. Recent stereoscopic triangulation of coronal loops has demonstrated a considerable mismatch between the extrapolated fields and the actual coronal loops, which cannot easily be reconciled with extrapolation algorithms, since they have only a very limited degree of freedom within the noise of the boundary data. Moreover, since each NLFFF solution is very time-consuming to compute, these algorithms are not suitable for forward-fitting.

The forward-fitting of magnetic field solutions to observed data requires a faster algorithm to compute many NLFFF solutions for variable boundary data or for coronal constraints as given by stereoscopic 3D reconstructions. The fastest computational way would be an explicit analytical solution for the coronal field vectors $\mathbf{B}(\mathbf{r})$ as a function of some suitable parameterization of the boundary data or coronal constraints. There exist some analytical solutions of nonlinear force-free fields, such as a class of solutions in terms of Legendre polynomials (Low and Lou, 1990), which is characterized by some spatial symmetry and has been used to test numerical extrapolation algorithms (*e.g.* DeRosa *et al.*, 2009; Malanushenko, Longcope, and McKenzie, 2009). However, to our knowledge, the class of analytical NLFFF solutions of Low and Lou (1990) has never been applied to forward-fitting of observed data, such as line-of-sight magnetograms, vector magnetograph 3D data, or to stereoscopically triangulated loops. Moreover, the special class of NLFFF solutions derived in Low and Lou (1990) correspond to harmonics of Legendre polynomials, which have a high degree of symmetry that does not match realistic observations of active regions, and thus is not suitable for forward-fitting to real data.

What we need to model observed solar magnetic data with high accuracy is:

i) an explicit formulation of an analytical NLFFF solution;
ii) a parameterization of the NLFFF solution with a sufficient large number of free parameters that can be forward-fitted to data and converges close to observations; and
iii) a fast computation algorithm that can perform many interactions without computing-intensive techniques.

Hence, such a project consists of developing a suitable analytical formulation first, and then to implement the analytical solutions into a forward-fitting code. In this paper we have undertaken the first step. We started with a potential-field parameterization in terms of $N_{\rm m}$ buried magnetic charges, which is defined by $4N_{\rm m}$ free parameters that can easily be extracted from an observed line-of-sight magnetogram $B_z(x, y)$ with arbitrary accuracy, as demonstrated in two recent studies (Aschwanden and Sandman, 2010; Aschwanden *et al.*, 2012a). The key concept of this potential-field representation is that an arbitrary complex 3D magnetic field can be decomposed into a finite number of elementary magnetic field components, where each one simply consists of a quadratically decreasing radial field of a buried magnetic charge. Divergence-freeness is conserved due to the linearity in the superposition of elementary field components. In a next step we extended the elementary potential-field component to a nonpotential-field component by adding a uniform twist that can be parameterized by the force-free α-parameter. Such an elementary nonpotential field component requires five free parameters, consisting of the four potential-field parameters plus the force-free α-parameter. We derived an explicit analytical formulation of the radial $B_r(r, \theta)$ and azimuthal field vector $B_\varphi(r, \theta)$ that represents an approximative solution of the divergence-free and force-free condition to second order ($\propto \alpha^2$). This solution is very accurate for weakly non-potential fields and converges to the potential field solution for $\alpha = 0$. In analogy to the potential-field representation, we represent a general non-potential field solution with a superposition of elementary non-potential field components and prove that the divergence-freeness and force-freeness is conserved to second-order accuracy in our NLFFF approximation.

We calculated some examples of potential and non-potential fields that mimic an isolated sunspot, a dipolar and a quadrupolar configuration, as well as more complex multi-polar configurations. The examples show that the magnetic field of arbitrary complex active regions can be represented with our parameterization. Increasing the force-free α-parameter distorts circular field lines into helical and sigmoid-shaped geometries. Our parameterization allows one to compute either field lines (starting from arbitrary locations), 3D data cubes of magnetic field vectors, of maps of the force-free α-parameter and electric current j_z (Figure 9). We tested the figures of merit for divergence-freeness and force-freeness, which amount to $L_{\rm d} \lesssim 10^{-3}$ and $L_{\rm f} \lesssim 10^{-2}$. The examples demonstrate also the computing speed of this algorithm, which amounts to the order of ≈ 1 s for a computation grid that encompasses a typical active region with the spatial resolution of MDI. Thus, we envision that a full-fletched forward-fitting code can converge within a few seconds to a few minutes, depending on the number of iterations and number of magnetic field components.

Where do we go from here? The next step is the development of a forward-fitting code that uses the magnetic field parameterization described here (see Paper II). We envision the applications to at least three different sets of constraints, requiring three different versions of forward-fitting codes: i) line-of-sight magnetograms $B_z(x, y)$ and 3D coordinates $[x(s), y(s), z(s)]$ of stereoscopically triangulated loops; ii) line-of-sight magnetograms $B_z(x, y)$ and 2D coordinates $[(x(s), y(s)]$ of traced loops; and iii) vector-magnetograph data $[B_x(x, y), B_y(x, y), B_z(x, y)]$. The first application requires STEREO data, while the second one can be obtained from any EUV imager (*e.g.* AIA/SDO, TRACE, EIT/SOHO). The third application can be conducted with the new HMI/SDO data and is equivalent to other NLFFF extrapolation codes without coronal constraints, while the first two use coronal tracers and alleviate the force-free assumption of photospheric data. We envision that these three applications will reveal insights into a number of crucial questions in a novel way.

There is a large number of physical problems and issues that can be addressed with the anticipated forward-fitting code, such as:

i) the force-freeness of the photosphere;
ii) the accuracy of NLFFF solutions;
iii) the spatial distribution of electric currents in active regions;
iv) the temporal evolution of currents before and during flares;
v) the spatial distribution of current dissipation and coronal heating;
vi) helicity injection;
vii) the 3D geometry of coronal loops which is needed for hydrodynamic modeling;
viii) scaling laws of the volumetric heating function with other physical parameters;
ix) tests of the magnetic field strength inferred from coronal seismology, *etc*.

There hardly exists a phenomenon in the solar corona that can be modeled without the knowledge of the coronal magnetic field.

Acknowledgements We thank Anny Malanushenko for helpful discussions. Part of the work was supported by NASA contract NNG 04EA00C of the SDO/AIA instrument and the NASA STEREO mission under NRL contract N00173-02-C-2035.

Appendix A: The Gold–Hoyle Flux Rope

A simple geometry of a force-free field structure is the Gold–Hoyle flux rope (Gold and Hoyle, 1960), which consists of a curved axis with helical field lines curved around the axis (Figure 10). While the stretched version of a flux rope with a straight twist axis has the exact force-free solution of a uniformly twisted flux tube (Section 2.2), the curved version of the Gold–Hoyle flux rope is subject to curvature forces due to the gradient of the magnetic field across the flux rope diameter and has a modified force-free solution.

In order to explore the limitations of our force-free field parameterization we attempt here to model such a Gold–Hoyle flux rope. We use the coordinates $(x_0, 0, z_0)$ and $(-x_0, 0, z_0)$ with $x_0 = 0.1$ and $z_0 = 0.985$ solar radii (marked with diamonds in Figure 11) and extrapolate field lines $B(s)$ with our method, starting from the apex position $(0, 0, z_a)$ with $z_a = 1.1$, for a set of six cases with various force-free parameters $\alpha_1 = \alpha_2$, where the α's associated with the twist axis of each buried charge are defined by $\alpha = 2\pi N_{\rm twist}/L$, with the loop length $L = 2\pi x_0 = 0.314$ and the number of twist turns $N_{\rm twist} = 0, 1, \ldots, 5$ (indicated with $N = 0, \ldots, 5$ in Figure 11). The case $N = 0$ corresponds to the potential field case, yielding a coplanar elliptical loop shape. The case $N = 1$ represents a slightly twisted field line that has a sigmoid shape and is a quasi-force-free solution. The cases with $N = 2, \ldots, 5$ are strongly twisted field lines and may be less force-free, since the neglected α^2 terms could be significant.

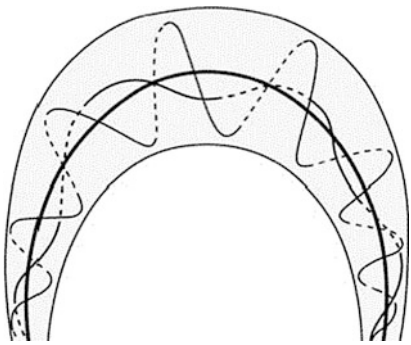


Figure 10 Cartoon of Gold–Hoyle flux rope.

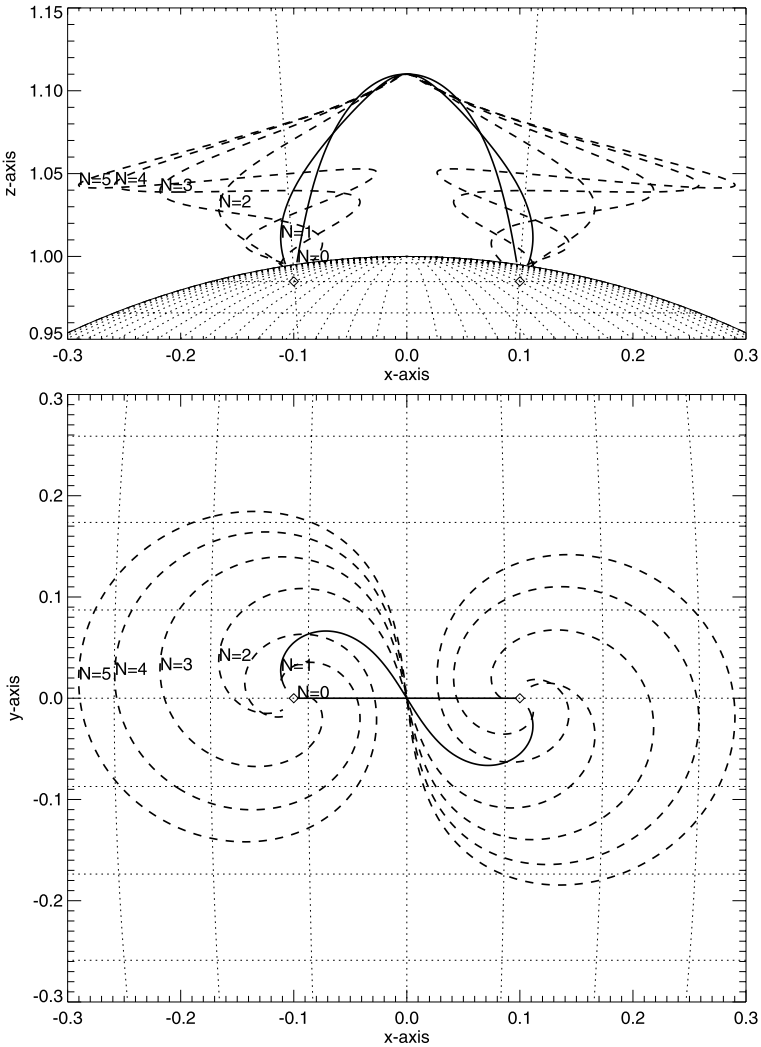


Figure 11 Dipolar field lines with various numbers of twisting turns: $N = 0$ (potential field line), stable sigmoid ($N = 1$; solid line), and unstable sigmoids ($N = 2, \ldots, 5$; dashed lines), according to our parameterization of point charges with twisted vertical axes. Note that the limit of large twist numbers does not turn into a Gold–Hoyle flux rope (Figure 10) with our parameterization.

Obviously we cannot reproduce the exact shape of the Gold–Hoyle flux rope as shown in Figure 10 (with about seven twist turns) with our choice of parameterization. The reason lies in the geometric constraints of the twist axis, which is semi-circular in the case of the Gold–Hoyle model, but consists of vertical twist axes in our parameterization. So, this counter-example clearly demonstrates the limitations of our parameterization. Nevertheless, although the cartoon with the Gold–Hoyle geometry is very popular, especially for interplanetary flux ropes and CMEs, it is not clear whether such Gold–Hoyle type geometries are found in loops in the lower corona, and whether the Gold–Hoyle geometry corresponds to an exact force-free solution. It is conceivable that the Sun exerts rotational stress mostly

in the photosphere (*i.e.* rotating sunspots), which propagates in vertical direction along the field lines, but does not necessarily lead to a uniformly twisted circular flux tube as shown in Figure 10, because the magnetic field drops rapidly with r^{-2} with height (for magnetic charges with small sub-photospheric depths), and thus the magnetic stress is not uniformly distributed along a semi-circular potential field line as envisioned in the Gold–Hoyle scenario. However, for a case with a near-constant magnetic field strength $B(s)$ along a potential field line, we would expect a uniform twist as outlined in the Gold–Hoyle case.

On the other side, strongly twisted flux tubes with a twist larger than about 1.25 full turns are unstable due to the kink instability and may erupt, which is another reason why multiply twisted flux tubes are unlikely to be found in active regions. Even Gold and Hoyle (1960) found a critical twist number of $\Phi_{\rm twist} \lesssim 2.49\pi$ ($N_{\rm twist} = \Phi_{\rm twist}/2\pi \lesssim 1.25$) above which no equilibrium exists, which is also confirmed by recent MHD simulations (*e.g.* Török and Kliem, 2003). Thus, the Gold and Hoyle flux rope case may not be relevant for modeling magnetic fields in stable active regions. Nevertheless, more general parameterizations could be anticipated in future work, such as twist axes that follow potential field lines, rather than vertical axes, as used in our parameterization to minimize the number of free parameters.

References

Aschwanden, M.J.: 2004, *Physics of the Solar Corona. An Introduction*, Praxis/Springer, Chichester/Berlin, 216.
Aschwanden, M.J.: 2009, *Space Sci. Rev.* **149**, 31.
Aschwanden, M.J., Malanushenko, A.: 2012, *Solar Phys.*, submitted (Paper II).
Aschwanden, M.J., Sandman, A.W.: 2010, *Astron. J.* **140**, 723.
Aschwanden, M.J., Wülser, J.P., Nitta, N., Lemen, J.: 2008, *Astrophys. J.* **679**, 827.
Aschwanden, M.J., Wülser, J.P., Nitta, N.V., Lemen, J.R., DeRosa, M., Malanushenko, A.: 2012a, *Astrophys. J.*, submitted.
Aschwanden, M.J., Wülser, J.P., Nitta, N.V., Lemen, J.R.: 2012b, *Solar Phys.*, in press.
Boyd, T.J.M., Sanderson, J.J.: 2003, *The Physics of Plasmas*, Cambridge University Press, Cambridge, 102.
Conlon, P.A., Gallagher, P.T.: 2010, *Astrophys. J.* **715**, 59.
DeRosa, M.L., Schrijver, C.J., Barnes, G., Leka, K.D., Lites, B.W., Aschwanden, M.J., *et al.*: 2009, *Astrophys. J.* **696**, 1780.
Feng, L., Inhester, B., Solanki, S., Wiegelmann, T., Podlipnik, B., Howard, R.A., Wülser, J.P.: 2007, *Astrophys. J. Lett.* **671**, L205.
Gary, A., Alexander, D.: 1999, *Solar Phys.*. **186**, 123.
Gold, T., Hoyle, F.: 1960, *Mon. Not. Roy. Astron. Soc.* **120**, 89.
Jackson, J.D.: 1962, *Classical Electrodynamics*, Wiley, New York, 184.
Klimchuk, J.A., Antiochos, S.K., Norton, D.: 2000, *Astrophys. J.* **542**, 504.
Low, B.C., Lou, Y.Q.: 1990, *Astrophys. J.* **352**, 343.
Malanushenko, A., Longcope, D.W., McKenzie, D.E.: 2009, *Astrophys. J.* **707**, 1044.
Malanushenko, A., Yusuf, M.H., Longcope, D.W.: 2011, *Astrophys. J.* **736**, 97.
Metcalf, T.R., Jiao, L., Uitenbroek, H., McClymont, A.N., Canfield, R.C.: 1995, *Astrophys. J.* **439**, 474.
Priest, E.R.: 1982, *Solar Magnetohydrodynamics*, Reidel, Dordrecht, 125.
Ruan, P., Wiegelmann, T., Inhester, B., Neukirch, T., Solanki, S.K., Feng, L.: 2008, *Astron. Astrophys.* **481**, 827.
Sandman, A.W., Aschwanden, M.J.: 2011, *Solar Phys.* **270**, 503.
Sandman, A., Aschwanden, M.J., DeRosa, M., Wülser, J.P., Alexander, D.: 2009, *Solar Phys.* **259**, 1.
Schrijver, C.J., DeRosa, M.L., Metcalf, T.R., Liu, Y., McTiernan, J., Regnier, S., Valori, G., Wheatland, M.S., Wiegelmann, T.: 2006, *Solar Phys.* **235**, 161.
Sturrock, P.A.: 1994, *Plasma Physics. – An Introduction to the Theory of Astrophysical, Geophysical and Laboratory Plasmas*, Cambridge University Press, Cambridge, 216.
Török, T., Kliem, B.: 2003, *Astron. Astrophys.* **406**, 1043.
Wiegelmann, T., Inhester, B.: 2003, *Solar Phys.* **214**, 287.
Wiegelmann, T., Inhester, B.: 2006, *Solar Phys.* **236**, 25.
Wiegelmann, T., Neukirch, T.: 2002, *Solar Phys.* **208**, 233.
Wiegelmann, T., Lagg, A., Solanki, S.K., Inhester, B., Woch, J.: 2005, *Astron. Astrophys.* **433**, 701.
Wheatland, M.S., Sturrock, P.A., Roumeliotis, G.: 2000, *Astrophys. J.* **540**, 1150.

DOI 10.1007/978-1-4899-8005-2_22
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0070-1

A Nonlinear Force-Free Magnetic Field Approximation Suitable for Fast Forward-Fitting to Coronal Loops. II. Numeric Code and Tests

Markus J. Aschwanden · Anna Malanushenko

Received: 28 February 2012 / Accepted: 7 July 2012 / Published online: 8 August 2012

Abstract Based on a second-order approximation of nonlinear force-free magnetic field solutions in terms of uniformly twisted field lines derived in Paper I, we develop here a numeric code that is capable to forward-fit such analytical solutions to arbitrary magnetogram (or vector magnetograph) data combined with (stereoscopically triangulated) coronal loop 3D coordinates. We test the code here by forward-fitting to six potential field and six nonpotential field cases simulated with our analytical model, as well as by forward-fitting to an exactly force-free solution of the Low and Lou (*Astrophys. J.* **352**, 343, 1990) model. The forward-fitting tests demonstrate: i) a satisfactory convergence behavior (with typical misalignment angles of $\mu \approx 1^\circ - 10^\circ$), ii) relatively fast computation times (from seconds to a few minutes), and iii) the high fidelity of retrieved force-free α-parameters ($\alpha_{\rm fit}/\alpha_{\rm model} \approx 0.9 - 1.0$ for simulations and $\alpha_{\rm fit}/\alpha_{\rm model} \approx 0.7 \pm 0.3$ for the Low and Lou model). The salient feature of this numeric code is the relatively fast computation of a quasi-force-free magnetic field, which closely matches the geometry of coronal loops in active regions, and complements the existing nonlinear force-free field (NLFFF) codes based on photospheric magnetograms without coronal constraints.

Keywords Sun: corona · Sun: magnetic fields

1. Introduction

This paper contains a description of a new numerical code that performs fast forward-fitting of nonlinear force-free magnetic fields (NLFFF). An alternative NLFFF forward-fitting code has been pioneered by Malanushenko, Longcope, and McKenzie (2009), which first fits

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

M.J. Aschwanden (✉) · A. Malanushenko
Solar and Astrophysics Laboratory, Lockheed Martin Advanced Technology Center, Dept. ADBS, Bldg. 252, 3251 Hanover St., Palo Alto, CA 94304, USA
e-mail: aschwanden@lmsal.com

separate linear force-free solutions to individual loops, and in a next step retrieves a self-consistent NLFFF solution from the obtained linear force-free α-values (Malanushenko, Yusuf, and Longcope, 2011). Since any calculation of a single NLFFF solution requires substantial computing time, we explore here a much faster NLFFF forward-fitting code that retrieves a self-consistent quasi-force-free magnetic field with somewhat reduced accuracy (*i.e.*, second order in α), but should be still sufficient for most practical applications.

The NLFFF models are thought to describe the magnetic field in the solar corona in a most realistic way, because the required force-freeness and divergence-freeness fulfill Maxwell's electrodynamic equations for a steady-state situation. Except for very dynamic episodes, such as flares or magnetic reconnection events, the magnetic field corona is thought to evolve close to a force-free steady state. NLFFF models reveal also the magnitude and topology of field-aligned currents, which are crucial for understanding energetic processes in the solar corona.

About a dozen NLFFF codes exist that have been described in detail and quantitatively compared (Schrijver *et al.*, 2006, 2008; Metcalf *et al.*, 1995, 2008; DeRosa *et al.*, 2009), which includes:

i) divergence-free and force-free optimization algorithms (Wheatland, Sturrock, and Roumeliotis, 2000; Wiegelmann, 2004),
ii) the evolutionary magneto-frictional method (Yang, Sturrock, and Antiochos, 1986; Valori, Kliem, and Fuhrmann, 2007), or a
iii) Grad–Rubin-style (Grad and Rubin, 1958) current-field iteration method (Amari, Boulmezaoud, and Aly, 2006; Wheatland, 2006; Malanushenko, Longcope, and McKenzie, 2009).

Most of these NLFFF algorithms use a photospheric boundary condition (in form of a magnetogram or 3D vector magnetograph data) and extrapolate the magnetic field in a coronal box above the photospheric boundary, by optimizing the conditions of divergence-freeness and force-freeness (for a general overview of non-potential field calculation methods see, *e.g.*, Aschwanden, 2004). Only the code of Malanushenko, Longcope, and McKenzie (2009) uses loop coordinates as additional constraints from the coronal volume. The methods have different degrees of accuracy, which can be quantified by an average misalignment angle between the theoretical model and observed (stereoscopically triangulated) coronal loops, which typically amounts to $\mu \approx 24^\circ - 44^\circ$ (see Table 1 in DeRosa *et al.*, 2009). These NLFFF codes are relatively computing-intensive (with typical computation times of several hours to a over a day), and thus are not suitable for forward-fitting, which requires many iterations.

In Paper I (Aschwanden, 2012) we derived an approximation of a general solution of a class of NLFFF models (with twisted magnetic fields) that is suitable for fast forward-fitting to coronal loops. The accuracy of this "quasi-NLFFF solution" is of second-order in the force-free parameter α. Obviously, we have a trade-off between accuracy and computation speed. This fast forward-fitting code can be applied to virtually every kind of simulated or observed magnetogram or 3D vector magnetograph data, combined with constraints from coronal loop coordinates, in form of 2D or 3D coordinates as they can be obtained by stereoscopic triangulation (*e.g.*, Feng *et al.*, 2007a; Aschwanden *et al.*, 2008a). In this Paper II we describe this first "fast" NLFFF forward-fitting code and test it with simulated data and analytical NLFFF solutions, such as obtained from the Low and Lou (1990) model.

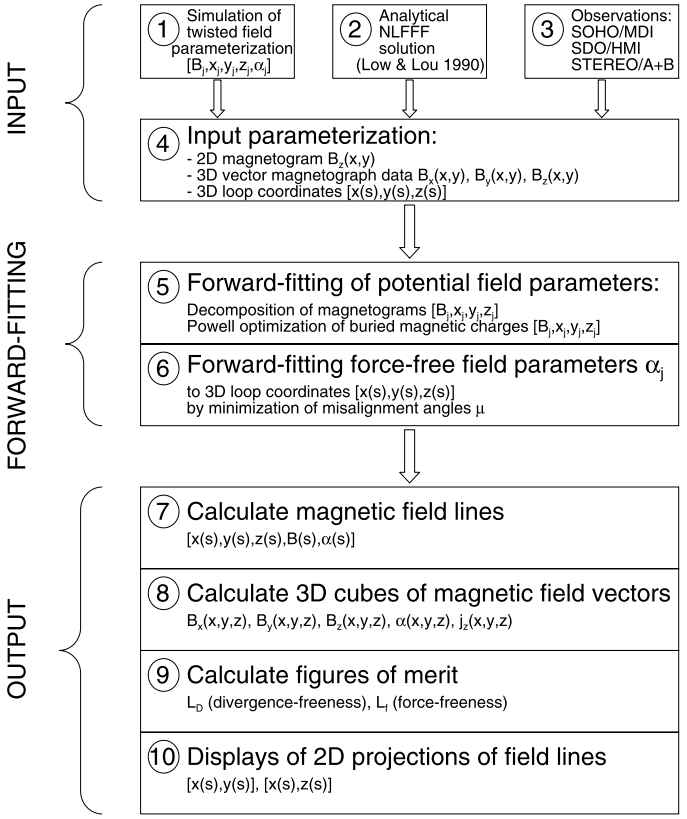


Figure 1 A flow chart of 10 modules of the forward-fitting code that calculates nonlinear force-free field solutions from various forms of inputs (simulations, analytical solutions, observational data). The 10 modules are described in Section 2.

2. Numeric Code

A scheme of the numeric code that performs forward-fitting of nonlinear force-free magnetic fields (NLFFF) is shown in Figure 1. The 10 different modules of the algorithm can be organized into three groups: Input modules (1 – 4), forward-fitting modules (5 – 6), and output modules (7 – 10), which we will describe in some more detail in the following.

(1) Simulated Input This module serves to create test cases and defines a 3D magnetic field model directly by $n = 5N_{\rm m}$ free parameters, which includes the surface magnetic field strength B_j and subphotospheric position (x_j, y_j, z_j) of the buried magnetic charges, as well as the force-free parameters α_j of the twisted magnetic field for every magnetic charge $j = 1, \ldots, N_{\rm m}$ (see definitions in Paper I). We will use models with $N_{\rm m} = 1 - 10$ magnetic charges, so we deal with $n = 5 - 50$ input parameters per test case. Our models will use unipolar ($N_{\rm m} = 1$), dipolar ($N_{\rm m} = 2$), quadrupolar ($N_{\rm m} = 4$), and random distributions of $N_{\rm m} = 10$ magnetic charges, where the models with multiple charges are grouped into pairs of opposite magnetic polarity with identical force-free parameters $\alpha_j = \alpha_{j+1}$ for pairs with conjugate magnetic polarization (to mimic a nearly force-free field). The purpose of this simulation module is mostly to test the convergence of the code (with a large number of

free parameters), so that the output can be compared with a known input, regardless of other problems, such as the suitability of our parameterization (which is unknown for external analytical or observational data) or the fulfillment of the divergence-free and force-free conditions (which define a NLFFF solution).

(2) Analytical NLFFF Solutions This module accesses external magnetic field data (in form of 3D cubes of magnetic field vectors) and extrapolated field lines (which serve as proxy for coronal loops) from a known analytical NLFFF solution. In our tests described here we will use solutions of a particular NLFFF model described in Low and Lou (1990), which is also summarized and used in Malanushenko, Longcope, and McKenzie (2009; Appendix A). The Low and Lou field depends on two free parameters in the Grad–Shafranov equation, which contains a constant a and the harmonic number n of the Legendre polynomial. We will use a model with $[a = 0.6, n = 2.0]$, which are also rendered in Malanushenko, Longcope, and McKenzie (2009). Since the Low and Lou model represents an exact analytical solution, we can test whether our code is capable to retrieve the correct force-free parameters $\alpha(\mathbf{x})$ in the 3D cube, as well as along individual loops, $\alpha(s)$. Furthermore, it will reveal whether our choice of magnetic field parameterization $(B_j, x_j, y_j, z_j, \alpha_j)$ is suitable to represent this particular NLFFF magnetic field, whether the forward-fitting code converges to the correct solution, and how divergence-free and force-free our analytical approximation of second order is compared with an exact NLFFF solution.

(3) Observational Data Input This module inputs external data directly, such as line-of-sight magnetograms $B_z(x, y)$ from SOHO/MDI or SDO/HMI, or alternatively vector fields $[B_x(x, y), B_y(x, y), B_z(x, y)]$ if available. In addition, constraints on coronal field lines can be obtained from stereoscopic triangulation from STEREO/A and B (*e.g.*, Feng *et al.*, 2007a; Aschwanden *et al.*, 2008a), in form of 3D field line coordinates $[x(s), y(s), z(s)]$, where s is a field line coordinate that extends from one loop footpoint $s = 0$ to the other loop footpoint at $s = L$, or to an open-field boundary of the 3D computation box. For future applications we envision also modeling with (automated) 2D loop tracings alone (*e.g.*, from SOHO/EIT, TRACE, *Hinode*/EIS, or SDO/AIA), without the necessity of STEREO observations. However, 2D loop tracings represent weaker constraints than 3D loop triangulations, and thus may imply larger ambiguities in the NLFFF forward-fitting solution.

(4) Input Coordinate System After we get input from one of the three options (Figure 1 top), we need to bring the input data into the same self-consistent coordinate system. Since magnetograms are measured in the photosphere, the curvature of the solar surface has to be taken into account. If a longitudinal magnetic field strength $B_z(x, y)$ is measured at image position (x, y), the corresponding line-of-sight coordinate z is defined by $x^2 + y^2 + z^2 = R_\odot^2$, which defines the 3D position of the magnetic field, $B_z(x, y, z)$. No correction of the coordinates of the magnetogram is needed for simulated and observed input data. However, the analytical NLFFF solution of Low and Lou (1990) neglects the curvature of the solar surface and yields the 3D magnetic field vectors $\mathbf{B}(\mathbf{x})$ in a Cartesian grid. Hence we place the Cartesian Low and Lou solution tangentially to the solar surface and extrapolate the magnetic field vectors to the exact position of the curved (photospheric) solar surface (assuming an r^{-2}-dependence). After we transformed all input into the same coordinate system, normalized to length units of solar radii ($R_\odot = 1$) from Sun center $[0, 0]$, we have magnetograms in form of $B_z(x, y, z_{\rm ph})$, or vector magnetograph data in the form of $[B_x(x, y, z_{\rm ph}), B_y(x, y, z_{\rm ph}), B_z(x, y, z_{\rm ph})]$, with the photospheric level at $z_{\rm ph} = \sqrt{1 - x^2 - y^2}$, and coronal loops in 3D coordinates of $[x(s), y(s), z(s)]$, with $0 < s < L$, and L being the length of a loop, or a segment of it.

(5) Forward-Fitting of Potential-Field Parameters We decompose first the line-of-sight magnetogram $B_z(x, y, z_{\rm ph})$ into a number of $N_{\rm m}$ buried magnetic charges, which produce 2D Gaussian-like local distributions $B_z(x, y)$ in the magnetogram, which are iteratively subtracted, while the maximum field strength B_j and 3D position (x_j, y_j, z_j) is measured for each component. An early approximate algorithm is shown with tests in Aschwanden and Sandman (2010; Equation (13) and Figure 3 therein). A more accurate inversion for the deconvolution of magnetic charges from a line-of-sight magnetogram is derived in Aschwanden *et al.* (2012a; Appendix A and Figure 4 therein). In order to obtain the maximum accuracy of this inversion, our code used the parameters (B_j, x_j, y_j, z_j) of the direct inversion as an initial guess and executes an additional forward-fitting optimization with the Powell method (Press *et al.*, 1986), where each of the $N_{\rm m}$ components is optimized by fitting the local magnetogram, repeated with four iterations for all magnetic sources. We found that the parameters converge already at the second iteration, given the relatively high accuracy of the initial guess. With this step we have already determined 80 % of the $n = 5N_{\rm m}$ free parameters $(B_j, x_j, y_j, z_j), \alpha_j,\ j = 1, \ldots, N_{\rm m}$, leaving only the force-free parameters α_j to be determined. If we set $\alpha_j = 0$, we have already an exact parameterization of the 3D potential field $\mathbf{B}_{\rm pot}(\mathbf{x})$, which also predicts the transverse field components $B_x(x, y, z_{\rm ph})$ and $B_y(x, y, z_{\rm ph})$ from the line-of-sight magnetogram $B_z(x, y, z_{\rm ph})$.

(6) Forward-Fitting of Non-potential-Field Parameters For the forward-fitting of the force-free parameters α_j for each magnetic charge $j = 1, \ldots, N_{\rm m}$ we can use either the constraints of the coronal loops ($q_{\rm v} = 0$), or the transverse components of the vector-magnetograph data ($q_{\rm v} = 1$), or a combination of both ($0 < q_{\rm v} < 1$), which we select with a weighting factor $q_{\rm v}$ in the optimization of the overall misalignment angle μ, *i.e.*,

$$\mu = q_{\rm v}\mu_{\rm loop} + (1 - q_{\rm v})\mu_{\rm vect}. \tag{1}$$

The forward-fitting of the best-fit force-free parameters α_j is performed by iterating the calculation of the 3D misalignment angle, which is defined for loops (or equivalently for a vector-magnetograph 3D field vector) by

$$\mu_{\rm loop} = \cos^{-1}\left(\frac{\mathbf{B}^{\rm theo} \cdot \mathbf{B}^{\rm obs}}{|\mathbf{B}^{\rm theo}| \cdot |\mathbf{B}^{\rm obs}|}\right), \tag{2}$$

between the theoretically calculated loop field lines $\mathbf{B}^{\rm theo}$ based on a trial set of parameters $(B_j, x_j, y_j, z_j), \alpha_j,\ j = 1, \ldots, N_{\rm m}$, and the observed field direction $\mathbf{B}^{\rm obs}$ of the observed loops. The overall misalignment angle is averaged (quadratically) from $N_{\rm seg} = 10$ loop positions in all $N_{\rm loop}$ loops. The variation of the trial sets of α_j is accomplished by a progressive subdivision of magnetic zones in subsequent iterations, starting from a single value for the entire active region (which corresponds to a linear force-free field model), and progressing with zones that become successively smaller by a factor of 2^{i-1}, with $i = 1, \ldots, N_{\rm iter}$ the number of iterations. The hierarchical subdivision of α-zones proceeds in order of decreasing magnetic field strength B_j. In each iteration all magnetic zones are successively varied, and for each zone the force-free parameter α_j is varied within a range of $|\alpha_j| < \alpha_{\rm max}$, until a minimum of the overall misalignment angle μ is found. An example of a hierarchical subdivision of α-zones in subsequent iterations is shown in Figure 2. For the test images we have chosen a dimension of $N_x = N_y = 60$, for which the subdivision of zone radii reaches a lower limit of one pixel after about five iterations (since $2^5 = 32 \approx N_x/2$). Thus, after five iterations, all magnetic sources are fitted individually in each iteration step. Convergence is generally reached for $N^{\rm iter} \lesssim 10-20$ iteration cycles. The computation scales linearly with the number $N_{\rm m}$ of magnetic sources and the number $N_{\rm loop}$ of fitted loops.

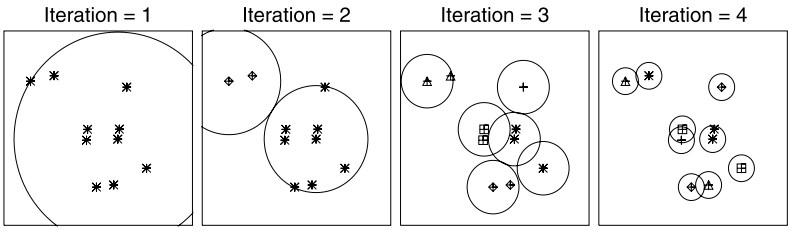


Figure 2 The scheme of hierarchical subdivision of α-zones (with a common force-free parameter α) is illustrated for four iteration cycles and $N_{\rm m} = 10$ magnetic charges. The number of α-zones increases with $2^{(i-1)}$ and the radius of an α-zone decreases with a factor $2^{(1-i)}$ in subsequent iterations $i = 1, \ldots, 4$. The number of α-zones becomes identical with the number of magnetic charges $j = 1, \ldots, N_{\rm m}$ after four iteration cycles. This number of free parameters α_i to be optimized is this way is reduced to 1, 2, 6, and 9 in subsequent iteration cycles for this example.

(7) Calculating Magnetic Field Lines Once our forward-fitting algorithm converged and determined a full set of $n = 5N_{\rm m}$ free parameters, $(B_j, x_j, y_j, z_j, \alpha_j)$, $j = 1, \ldots, N_{\rm m}$, we can calculate the magnetic field vector $\mathbf{B}(\mathbf{x})$ of the quasi-force-free field at any arbitrary location $\mathbf{x} = (x, y, z)$ in space (see Equations (34)–(42) in Paper I). To calculate the magnetic field along a particular field line $[x(s), y(s), z(s)]$, we just step iteratively by increments Δs,

$$\begin{aligned} x(s + \Delta s) &= x(s) + \Delta s\big[B_x(s)/B(s)\big]p, \\ y(s + \Delta s) &= y(s) + \Delta s\big[B_y(s)/B(s)\big]p, \\ z(s + \Delta s) &= z(s) + \Delta s\big[B_z(s)/B(s)\big]p, \end{aligned} \tag{3}$$

where $p = \pm 1$ represents the sign or polarization of the magnetic charge, and thus can be flipped to calculate a field line into opposite direction.

(8) Calculation of 3D Data Cubes By the same token we calculate 3D cubes of magnetic field vectors $\mathbf{B}(\mathbf{x}) = B_x(x_i, y_j, z_k)$, $B_y(x_i, y_j, z_k)$, $B_z(x_i, y_j, z_k)$, in a Cartesian grid (i, j, k) with $i = 1, \ldots, N_x$, $j = 1, \ldots, N_y$, $k = 1, \ldots, N_z$. The 3D cubes of force-free parameters $\alpha(x_i, y_j, z_k)$ can be calculated from the $B(x_i, y_j, z_k)$ cubes, for each of the three vector components,

$$\alpha_x(\mathbf{x}) = \frac{1}{4\pi}\frac{(\nabla \times \mathbf{B})_x}{\mathbf{B}_x} = \frac{1}{4\pi B_x}\left(\frac{\partial B_z}{\partial y} - \frac{\partial B_y}{\partial z}\right), \tag{4}$$

$$\alpha_y(\mathbf{x}) = \frac{1}{4\pi}\frac{(\nabla \times \mathbf{B})_y}{\mathbf{B}_y} = \frac{1}{4\pi B_y}\left(\frac{\partial B_x}{\partial z} - \frac{\partial B_z}{\partial x}\right), \tag{5}$$

$$\alpha_z(\mathbf{x}) = \frac{1}{4\pi}\frac{(\nabla \times \mathbf{B})_z}{\mathbf{B}_z} = \frac{1}{4\pi B_z}\left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right), \tag{6}$$

using a second-order scheme to compute the spatial derivatives, *i.e.*, $\partial B_x/\partial y = (B_{i+1j,k} - B_{i-1,j,k})/2(y_{i+1} - y_{i-1})$. In principle, the three values α_x, α_y, α_z should be identical, but the numerical accuracy using a second-order differentiation scheme is most handicapped for those loop segments with the smallest values of the B-component (appearing in the denominator), for instance in the α_z component $\propto (1/B_z)$ near the loop tops (where $B_z \approx 0$). It is therefore most advantageous to use all three parameters α_x, α_y, and α_z in a weighted mean,

$$\alpha = \frac{\alpha_x w_x + \alpha_y w_y + \alpha_z w_z}{w_x + w_y + w_z}, \tag{7}$$

but weight them by the magnitude of the (squared) magnetic field strength in each component,

$$w_x = B_x^2, \qquad w_y = B_y^2, \qquad w_z = B_z^2, \tag{8}$$

so that those segments have no weight where the B-component approaches zero. With this method, we can determine the force-free parameter $\alpha(x_i, y_j, z_k)$ at any given 3D grid point $[x_i, y_j, z_k]$, as well as along a loop coordinate, $\alpha(s)$.

The 3D cubes of current densities $\mathbf{j} = (j_x, j_y, j_z)$ follow from the definition $\mathbf{j}/c = (\nabla \times \mathbf{B})/(4\pi) = \alpha(\mathbf{x})\mathbf{B}$,

$$\mathbf{j}(x_i, y_j, z_k) = c\alpha(x_i, y_j, z_k)\mathbf{B}(x_i, y_j, z_k). \tag{9}$$

(9) Calculation of Figures of Merit Figures of merit (how physical a converged NLFFF solution is) can be computed for the divergence-freeness $\nabla \cdot \mathbf{B} = 0$ compared to the field gradient $B/\Delta x$ over a pixel length Δx,

$$L_\mathrm{d} = \frac{1}{V} \int_V \frac{|(\nabla \cdot \mathbf{B})|^2}{|B/\Delta x|^2} \,\mathrm{d}V. \tag{10}$$

Similarly, the force-freeness can be quantified by the ratio of the Lorentz force, $(\mathbf{j} \times \mathbf{B}) \propto (\nabla \times \mathbf{B}) \times \mathbf{B}$ to the normalization constant $B^2/\Delta x$,

$$L_\mathrm{f} = \frac{1}{V} \int_V \frac{|(\nabla \times \mathbf{B}) \times \mathbf{B}|^2}{|B^2/\Delta x|^2} \,\mathrm{d}V, \tag{11}$$

where $B = |\mathbf{B}|$. We calculate these quantities in agreement with the definitions given in Paper I.

(10) Display of 2D Projections For visualization purposes of the 3D field, of both the numerically calculated solution (of our quasi-NLFFF model) and for the observed loops, it is most practical to display the field lines in the three orthogonal projections, *i.e.*, $[x(s), y(s)]$ for a top-down view, or $[x(s), z(s)]$ and $[y(s), z(s)]$ for side views.

Control Parameter Settings The numeric forward-fitting code has a number of control parameter settings, which can be changed individually to optimize the performance or the computation speed of the code. We list the set of standard control parameter settings in Table 1, which are generally used in this paper if not mentioned otherwise. These parameters control: the selection of loop field lines (module 1 – 3: N_grid, Δx, Thresh), the decomposition of the magnetogram (module 4: N_mag, q_mag, nsm, i_opt), and the forward-fitting of the force-free α parameter (module 5: Meth, N_iter, Δs, N_seg, h_max, h_alt, α_max, acc, q_loop, q_zone, q_v, eps).

3. Potential Field Tests

A first set of six test cases consists of potential field models (with $\alpha_j = 0$), including a unipolar charge, a dipole, a quadrupole, and three cases with 10 randomly distributed magnetic sources, identical to Cases #1 – 3 in Paper I, and to Cases #7 – 9 (but with α_j set to zero). For each of these six cases we show in Figure 3 a set of field lines calculated from the model (Figure 3, red curves), and a set of field lines obtained from forward-fitting with our NLFFF code. The agreement between the two sets of field lines can be expressed by the mean 3D misalignment angle μ (Equation (2)), which is found to be very small, within a range of $\mu = 0.0° - 6.6°$, or $\mu = 3.4° \pm 2.1°$. The individual values are listed in Table 2

Table 1 Standard control parameter settings of the forward-fitting code used in the tests of this study.

Parameter	Description
$N_{\rm grid} = 8$	Grid size in pixels for loop footpoint selection
$\Delta x = 0.0034$	Pixel size of computation grid (in solar radii)
Thresh = 0	Threshold of magnetic field [gauss] for loop footpoint selection
$N_{\rm mag} = 10$	Maximum number of magnetic charges
$q_{\rm mag} = 0.001$	Residual limit $B/B_{\rm max}$ of magnetogram decomposition
nsm = 0	Smoothing of magnetogram (in number of boxcar pixels)
$i_{\rm opt} = 4$	Number of cycles for optimization of potential field parameters
Meth = A	Method of subdividing magnetic zones
$N_{\rm iter} = 20$	Maximum number of iteration cycles
$\Delta s = \Delta x$	Spatial resolution along field line (in solar radii)
$N_{\rm seg} = 10$	Number of loop segments for misalignment angle calculation
$h_{\rm max} = 3.5\Delta x$	Maximum altitude range for magnetogram calculation (solar radii)
$h_{\rm alt} = 0.15$	Maximum altitude range for field line extrapolation
$\alpha_{\rm max} = 100$	Maximum range for force-free α per iteration (solar radius^{-1})
acc = 0.001	Relative accuracy in α optimization step
$q_{\rm loop} = 0.5$	Relative loop position for starting of field line computation
$q_{\rm zone} = 0.5$	Magnetic zone diminishing factor in subsequent iterations
$q_{\rm v} = 0.0$	Weighting factor of loop data vs. vector magnetograph data
eps = 0.1	Convergence criterion for change in misalignment angle (deg)

(fourth column). While this mostly represents a test of the accuracy of module 5 (forward-fitting of potential-field parameters), the algorithm treats the force-free parameter α_j as a variable too, and thus it represents also a test of the accuracy in determining this parameter in general. Compared with the theoretical value as it was set in the simulation of the input magnetogram ($\alpha_j = 0$), the best-fit values are found to be $\alpha = -0.09 \pm 0.15$ (Table 2, fifth column), which corresponds to $\Delta N_{\rm twist} = bl/2\pi = \alpha l/4\pi = \pm 0.0018$ twist turns over the length $l = 0.05\pi = 0.157$ solar radii of a typical field line (see definitions in Equations (16)–(17) in Paper I). Thus the uncertainty of our forward-fitting corresponds to less than ± 0.2 % of a full twist turn over a loop length. Another measure of the quality of the NLFFF forward-fit is the divergence-freeness, which is found to be $L_{\rm d} = (5 \pm 6) \times 10^{-6}$ (Table 2, sixth column), and the force-freeness, which is found to be $L_{\rm f} = (84 \pm 194) \times 10^{-6}$ (Table 2, seventh column), both being extremely accurate. The average computation time for the NLFFF forward-fitting runs of potential field cases was found to be $t_{\rm CPU} \approx 61$ s (on a Mac OS X with 2 × 3.2 GHz Quad-Core Intel Xeon processor and 32 GB 800 MHz DDR2 FB-DIMM Memory).

We performed also some parametric studies to explore the accuracy of the forward-fitting code as a function of some control parameters that are different from the standard settings given in Table 1. We list the results in Table 3. If we increase the spatial resolution of the field line extrapolation to $\Delta s/\Delta x = 0.5$, the accuracy of the field lines does not change, neither in terms of the mean misalignment angle nor in the divergence-freeness figure of merit (Table 3, second line). Increasing the number of magnetic components in the decomposition of magnetograms does not improve the accuracy for the potential-field cases (*e.g.*, by a factor of two compared with the simulated numbers of $N_{\rm mag} = 1, 2, 4, 10$), but degrades the divergence-freeness and force-freeness and increases the computation time by a factor

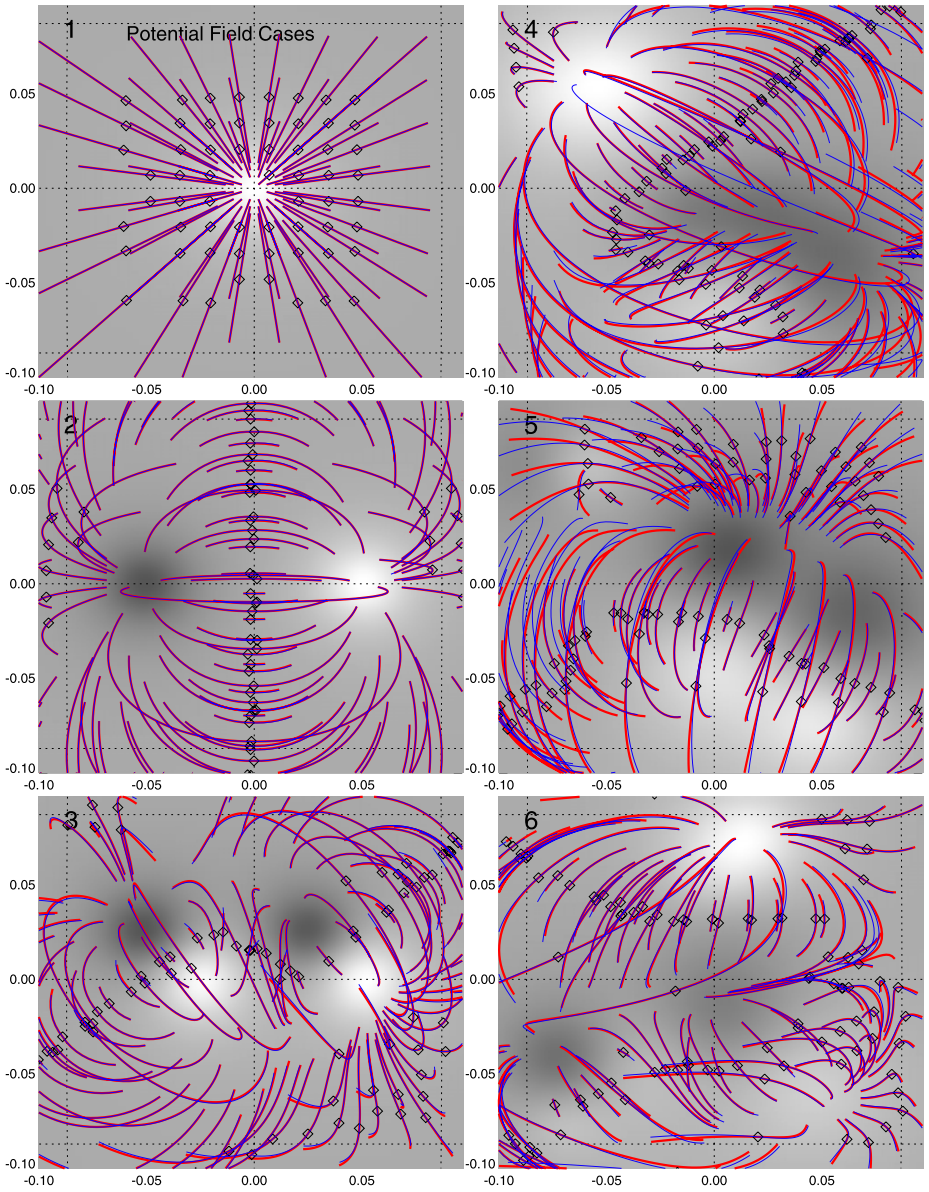


Figure 3 Test cases #1 – 6 are shown, consisting of a unipolar (#1: top left), a dipolar (#2: middle left), a quadrupolar (#3: bottom left), and three decapolar cases (#4 – 6: panels on right side). The displays contain the line-of-sight magnetograms (greyscale), the theoretically simulated loop field lines (red curves), and the overlaid best-fit NLFFF field lines (blue curves). The starting point of the calculated field lines are indicated with diamonds (at midpoint of loops, $q_{\rm loop} = 0.5$). Note the small amount of misalignment, ranging from $\mu = 0.0°$ (#1) to $\mu = 6.6°$ (#5) (the values are given in Table 2).

of ≈ 4 (Table 3, third line). Starting the field line extrapolation at the footpoints ($q_{\rm loop} = 0.0$), rather than from the loop midpoints ($q_{\rm loop} = 0.5$), leads to no significant improvement (Table 3, fourth line). Changing the weighting of coronal loop constraints ($q_{\rm v} = 0$) to using

Table 2 Best-fit parameters of forward-fitting of the NLFFF model to potential field cases (with $\alpha_j = 0$), using standard settings of the forward-fitting code (Table 1). The columns contain the case #= 1 – 6, the number of magnetic charges $N_{\rm mag}$, the number of loop field lines $N_{\rm loop}$, the mean misalignment angle μ, the mean best-fit force-free parameter α per loop, the divergence-freeness figure of merit $L_{\rm d}$, the force-freeness figure of merit $L_{\rm f}$, and the computation time $t_{\rm CPU}$ of the forward-fitting module 6. The last lines of the table contain the means and standard deviations σ of the six cases.

#	$N_{\rm mag}$	$N_{\rm loop}$	μ	α	$L_{\rm d}$	$L_{\rm f}$	$t_{\rm CPU}$
1	1	61	0.0°	0.00	0.000001	0.000001	2 s
2	2	91	2.7°	0.00	0.000002	0.000002	9 s
3	4	91	3.8°	−0.03	0.000007	0.000008	23 s
4	10	107	3.0°	−0.06	0.000001	0.000008	111 s
5	10	95	6.6°	−0.08	0.000017	0.000480	117 s
6	10	98	4.2°	−0.39	0.000003	0.000003	105 s
Mean			3.4°	−0.09	0.000005	0.000084	61 s
$\pm\sigma$			±2.1°	±0.15	±0.000006	±0.000194	±55 s

Table 3 Best-fit parameters of forward-fitting of the NLFFF model to potential field cases (with $\alpha_j = 0$), using some non-standard settings in the spatial resolution $\Delta s/\Delta x$ of calculated field lines, the number of magnetic source components $N_{\rm mag}$, the starting point of field line extrapolation $q_{\rm loop}$, the relative weighting of loop and vector magnetograph data $q_{\rm v}$, but otherwise standard settings as listed in Table 1.

Δs	$N_{\rm mag}$	$q_{\rm loop}$	$q_{\rm v}$	μ	α	$L_{\rm d}$ [10^{-6}]	$t_{\rm CPU}$
×1.0	×1	1.0	0.0	3.4° ± 2.1°	−0.09 ± 0.15	5 ± 6	61 ± 55 s
×0.5	×1	1.0	0.0	3.4° ± 2.1°	−0.08 ± 0.14	5 ± 6	70 ± 66 s
×1.0	×2	1.0	0.0	3.2° ± 1.7°	0.03 ± 0.13	10 ± 19	228 ± 233 s
×1.0	×1	0.0	0.0	3.4° ± 2.1°	−0.07 ± 0.16	5 ± 6	71 ± 66 s
×1.0	×1	1.0	1.0	1.8° ± 2.3°	−0.07 ± 0.10	3 ± 3	314 ± 290 s

only photospheric vector magnetograph data ($q_{\rm v} = 1$) improves the misalignment to $\mu = 1.8° \pm 2.3°$, which represents an improvement in the accuracy by about a factor of two, but requires about five times more computation time. Thus, the accuracy in fitting potential field cases is fairly robust and does not depend the detailed setting of control parameters, except for the weighting of photospheric *versus* coronal constraints.

4. Forward-Fitting to Quasi-NLFFF Models

Now we present the first tests of forward-fitting to non-potential fields (with $\alpha_j \neq 0$), numbered as test cases #7 – 12. These six cases have the same line-of-sight magnetograms $B_z(x, y)$ or magnetic charges (B_j, x_j, y_j, z_j) as the potential-field cases #1 – 6, but have a different twist or force-free parameter α_j. We show the magnetograms and the theoretical field lines of the models in Figure 4 (red curves), and the best-fit field lines of our NLFFF forward-fitting code in Figure 4 (blue curves), using standard control parameter settings (Table 1). The misalignment between these two sets of simulated and forward-fitted field lines amounts to $\mu = 0.7° - 12.8°$, or $\mu = 5.1° \pm 4.3°$ (Table 4, fourth column). These test results are quite satisfactory, first of all since the difference between the theoretical and

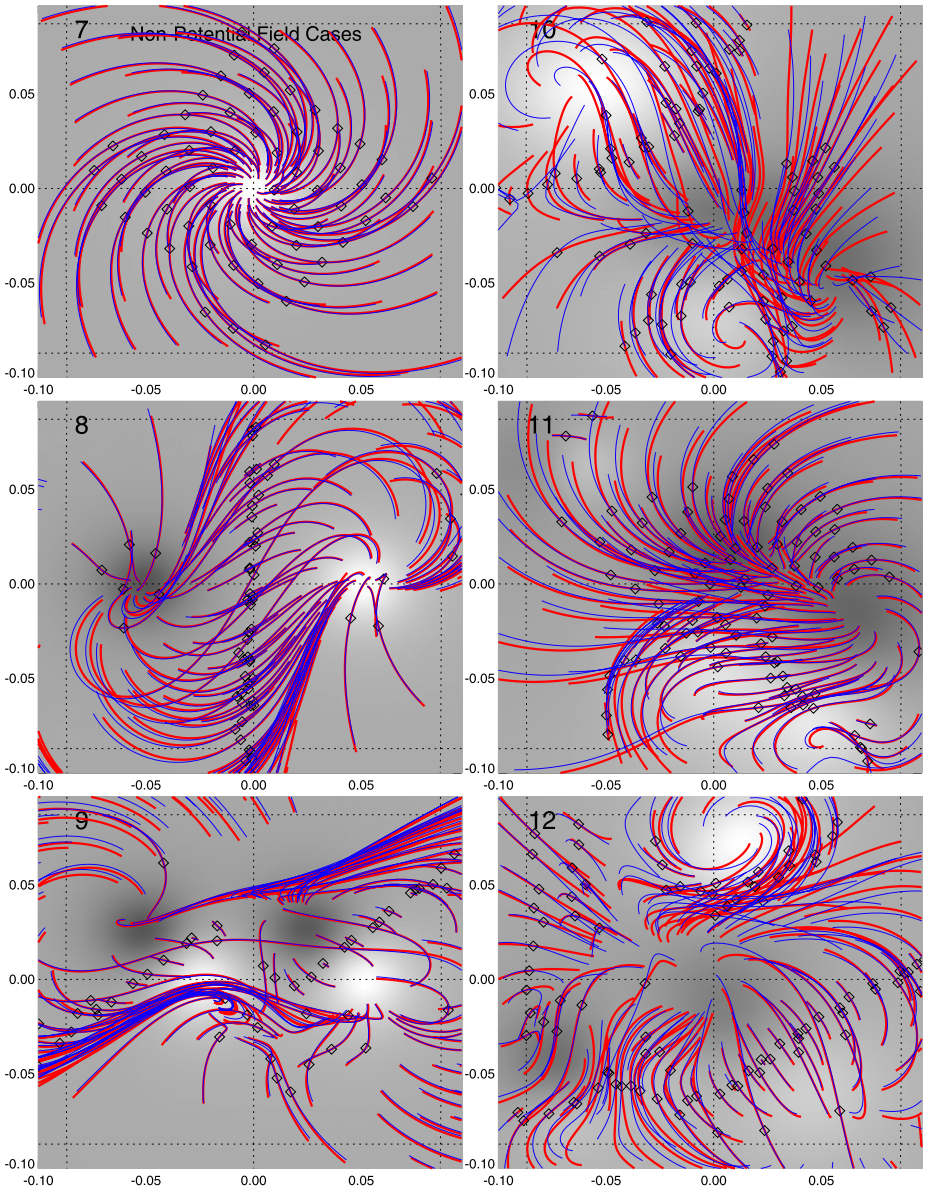


Figure 4 Test cases #7 – 12 are shown, consisting of a unipolar (#7: top left), a dipolar (#8: middle left), a quadrupolar (#9: bottom left), and three decapolar cases (#10 – 12: panels on right side). The displays contain the line-of-sight magnetograms (greyscale), the theoretically simulated loop field lines (red curves), and the overlaid best-fit NLFFF field lines (blue curves). The starting points of the calculated field lines are indicated with diamonds (at midpoint of loops, $q_{\rm loop} = 0.5$). The misalignment angles between the theoretical models and the best fits are listed in Table 4. Note the huge difference of field line topologies compared with the potential-field cases (shown in Figure 3), although the line-of-sight magnetograms are identical.

Table 4 Best-fit parameters of forward-fitting of the NLFFF model to non-potential field cases (with $\alpha_j \neq 0$), using standard settings of the forward-fitting code (Table 1). The columns contain the cases #= 7 – 12, the number of magnetic charges $N_{\rm mag}$, the number of loop field lines $N_{\rm loop}$, the mean misalignment angle μ, the mean input force-free α parameter values, the divergence-freeness figure of merit $L_{\rm d}$, the force-freeness figure of merit $L_{\rm f}$, and the computation time $t_{\rm CPU}$ of the forward-fitting module 6.

#	$N_{\rm mag}$	$N_{\rm loop}$	μ	α	$L_{\rm d}$	$L_{\rm f}$	$t_{\rm CPU}$
7	1	66	0.7°	−20	0.000453	0.000299	2 s
8	2	85	2.2°	-20 ± 1	0.000253	0.000104	8 s
9	4	82	3.6°	-30 ± 12	0.000727	0.000691	21 s
10	10	89	12.8°	29 ± 40	0.001813	0.004672	118 s
11	10	89	4.5°	2 ± 102	0.000784	0.003123	179 s
12	10	99	7.1°	74 ± 62	0.000976	0.005334	271 s
Mean			5.1°	6	0.000834	0.002370	99 s
$\pm\sigma$			±4.3°	±40	±0.000543	±0.002319	±109 s

best-fit field lines in Figure 4 are hardly recognizable by eye, and thus will suffice for all practical purposes, and secondly, the misalignment is about an order of magnitude smaller than found between traditional NLFFF codes and stereoscopically triangulated coronal loops ($\mu \approx 24° - 44°$; DeRosa *et al.*, 2009). We see that the force-free parameters vary substantially, in a range of $\alpha = 6 \pm 40$ (solar radius^{-1}) (Table 4, fifth column), which translates into a number $N_{\rm twist} = \alpha l / 4\pi \approx 0.5$ of (full) twist turns over a typical loop length. The merit of figure for the divergence-freeness is $L_{\rm d} = (0.8 \pm 0.5) \times 10^{-3}$ (Table 4, sixth column), and the merit of figure for the force-freeness is $L_{\rm f} = (2.3 \pm 2.3) \times 10^{-3}$ (Table 4, seventh column). The computation time is ($t_{\rm CPU} \approx 100$ s), less than a factor of two longer than for the potential-field cases (Table 2).

In order to achieve the most accurate performance of our code we explored also other control parameter settings than the standard parameters given in Table 1. Instead of using the hierarchical α-zone subdivision as shown in Figure 2 (Meth = A), we tested also other methods, such as subdivision by magnetically conjugate pairs of magnetic charges (Meth = B), or subdivision by magnetically conjugate loop footpoints (Meth = C). In 90 % of the test cases all three methods converged to the same minimum misalignment angle within ±0.1°, but for the 10 % of discrepant cases method A performed always best, so we conclude that method A is the most robust one.

Increasing the resolution of calculating field lines to $\Delta s = 0.5 \Delta x$ does not improve the misalignment ($\mu = 5.1° \pm 4.3°$; Table 5, second case); Increasing the number of magnetic sources by a factor of two does not improve the misalignment significantly either (Table 5; third case). Starting the field line extrapolation at the footpoints ($q_{\rm loop} = 0.0$) rather than from the loop midpoints, has no effect either (Table 5; fourth case). However, the change of replacing the coronal ($q_{\rm v} = 0$) to photospheric constraints, using 3D vector magnetograph data $q_{\rm v} = 1$ does improve the best fits substantially, but is more costly in computation time (Table 5, fifth case). The reason for this improvement is probably that photospheric field vectors are more uniformly distributed than coronal loops, but coronal constraints are more important when the photospheric magnetic field is not force-free.

The agreement between the best forward-fitting solutions of the magnetic field components (B_x, B_y, B_z) and the model are shown in Figure 5. Note that only the line-of-sight magnetogram $B_z(x, y, z_{\rm ph})$ was used as input to the forward-fitting code, for standard control parameter settings ($q_{\rm v} = 0$). For these tests, the code predicts the transverse component

Table 5 Best-fit parameters of forward-fitting of the NLFFF model to potential field cases (with $\alpha_j = 0$), using some non-standard settings in the spatial resolution $\Delta s/\Delta x$ of calculated field lines, the number of magnetic source components $N_{\rm mag}$, the starting point of field line extrapolation $q_{\rm loop}$, the relative weighting of loop and vector magnetograph data $q_{\rm v}$, but otherwise standard settings as listed in Table 1.

Δs	$N_{\rm mag}$	$q_{\rm loop}$	$q_{\rm v}$	μ	$L_{\rm d}$ $[10^{-3}]$	$L_{\rm f}$ $[10^{-3}]$	$t_{\rm CPU}$
$\times 1.0$	$\times 1$	1.0	0.0	$5.1° \pm 4.3°$	0.8 ± 0.5	2.3 ± 2.3	99 ± 109 s
$\times 0.5$	$\times 1$	1.0	0.0	$5.1° \pm 4.3°$	0.8 ± 0.6	2.3 ± 2.3	100 ± 110 s
$\times 1.0$	$\times 2$	1.0	0.0	$4.5° \pm 2.9°$	0.7 ± 0.3	2.1 ± 2.1	303 ± 307 s
$\times 1.0$	$\times 1$	0.0	0.0	$5.1° \pm 4.3°$	0.8 ± 0.5	2.4 ± 2.3	100 ± 110 s
$\times 1.0$	$\times 1$	1.0	1.0	$3.4° \pm 3.6°$	0.7 ± 0.3	2.2 ± 2.1	459 ± 342 s

maps $B_x(x, y)$ and $B_y(x, y)$, which is quite satisfactory for this set of tests, as Figure 5 demonstrates. The mean ratios of the absolute magnetic field strengths are accurate within a few percents (indicated in each panel of Figure 5).

The force-free parameter α is shown as a photospheric map $|\alpha(x, y, z_{\rm ph})|$ for the model (Figure 6, top and third row) and for the forward-fitting solution (Figure 6, second and fourth row). The comparison can be quantified by the ratio of the two values, which agrees within a few percents. A sensible test is also to display a scatterplot of the best-fit α-values versus the model α-values for each pixel of a photospheric map (Figure 7), or averaged along each of the fitted coronal loops (Figure 8). The ratios of the two quantities ranges from $\alpha_{\rm fit}/\alpha_{\rm loop} = 0.99 \pm 0.00$ for the best case (#7, Figure 8 top left) to $\alpha_{\rm fit}/\alpha_{\rm loop} = 0.88 \pm 0.29$ for the worst case (#12, Figure 6, bottom right). Our forward-fitting code retrieves the correct sign of the α-parameter in all cases, and their absolute values agree within a few percents with the theoretical model. Thus we conclude that the convergence behavior of our forward-fitting code is quite satisfactory, because it retrieves the force-free α-parameters with high accuracy, at least for the given parameterization.

5. Forward-Fitting to Low and Lou (1990) Model

The foregoing tests were necessary to verify how accurately the forward-fitting code can retrieve the solution with many free parameters (from $n_{\rm free} = 5, \ldots, 50$), which represents a numerical convergence test. Of course, because the same parameterization is used in simulating the input data as in the model that is forward-fitted to the simulated data, this represents the most favorable condition where the model parameterization is adequate for the input data. Moreover, the simulated data were only force-free to second order, so we cannot use the force-freeness figure of merit calculated from the solution as an absolute criterion to evaluate how accurate the forward-fitting solution fulfills Maxwell's equations. So, the foregoing tests do not tell us whether the model parameterization of the forward-fitting code is adequate for arbitrary data, and how physical the solution is.

We conduct now a test that generates the input data with a completely different parameterization than our model and fit a non-potential field case that is exactly force-free, which is provided by analytical NLFFF solutions of the Low and Lou (1990) model, described and used also in Malanushenko, Longcope, and McKenzie (2009). The particular solution we are using is defined by the parameters $(a = 0.6, n = 2.0)$, where a is a Grad–Shafranov constant and n is the harmonic number of the Legendre polynomial.

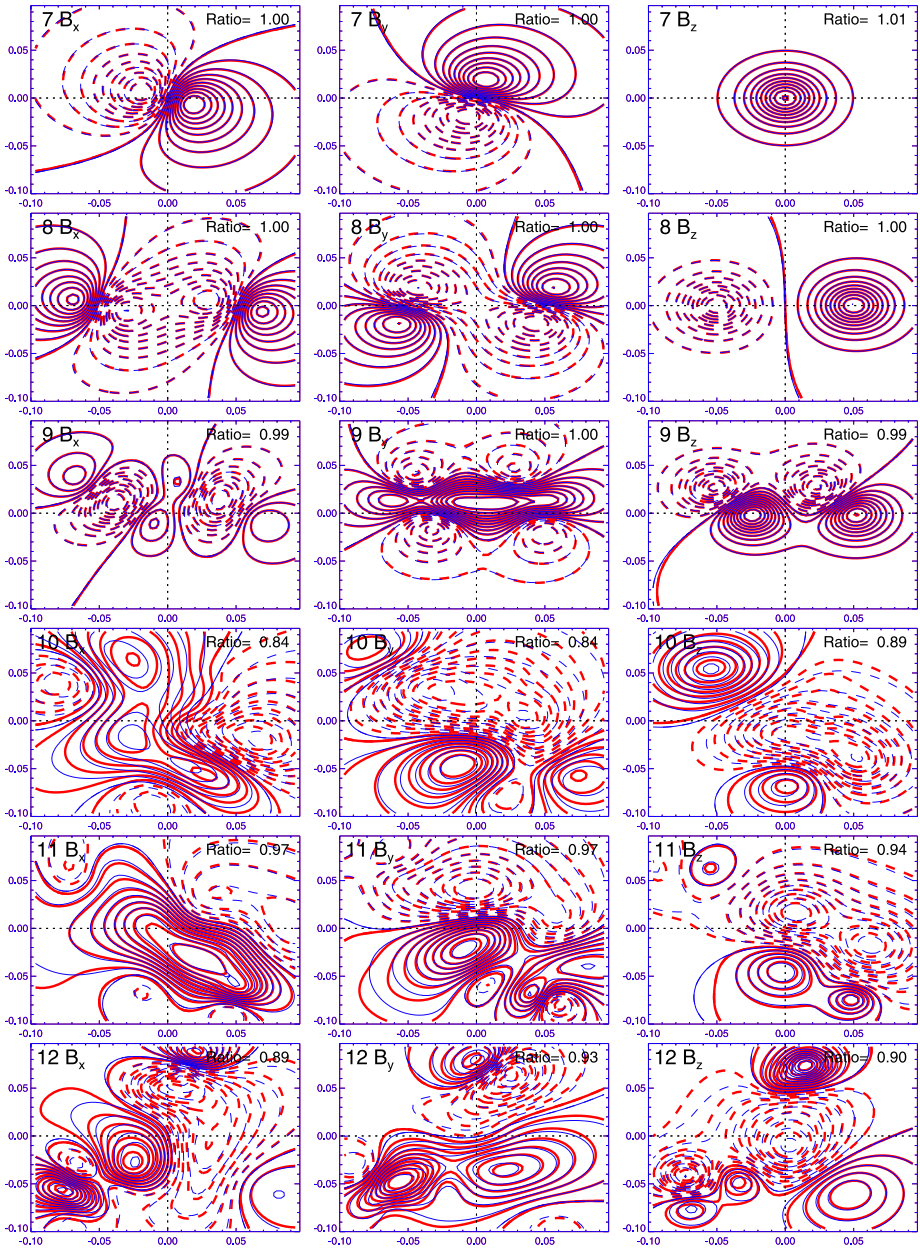


Figure 5 Contour maps of magnetic field component maps $B_x(x, y)$ (left column), $B_y(x, y)$ (middle column), and line-of-sight component $B_z(x, y)$ at the photospheric level for cases #7 – 12 (rows), shown with red contours (solid for positive and dashed for negative magnetic polarity). The best fits that result from the decomposition of the line-of-sight component are shown with blue curves, and the mean ratio of the absolute magnetic field strengths between the best fit and the model are indicated in each frame.

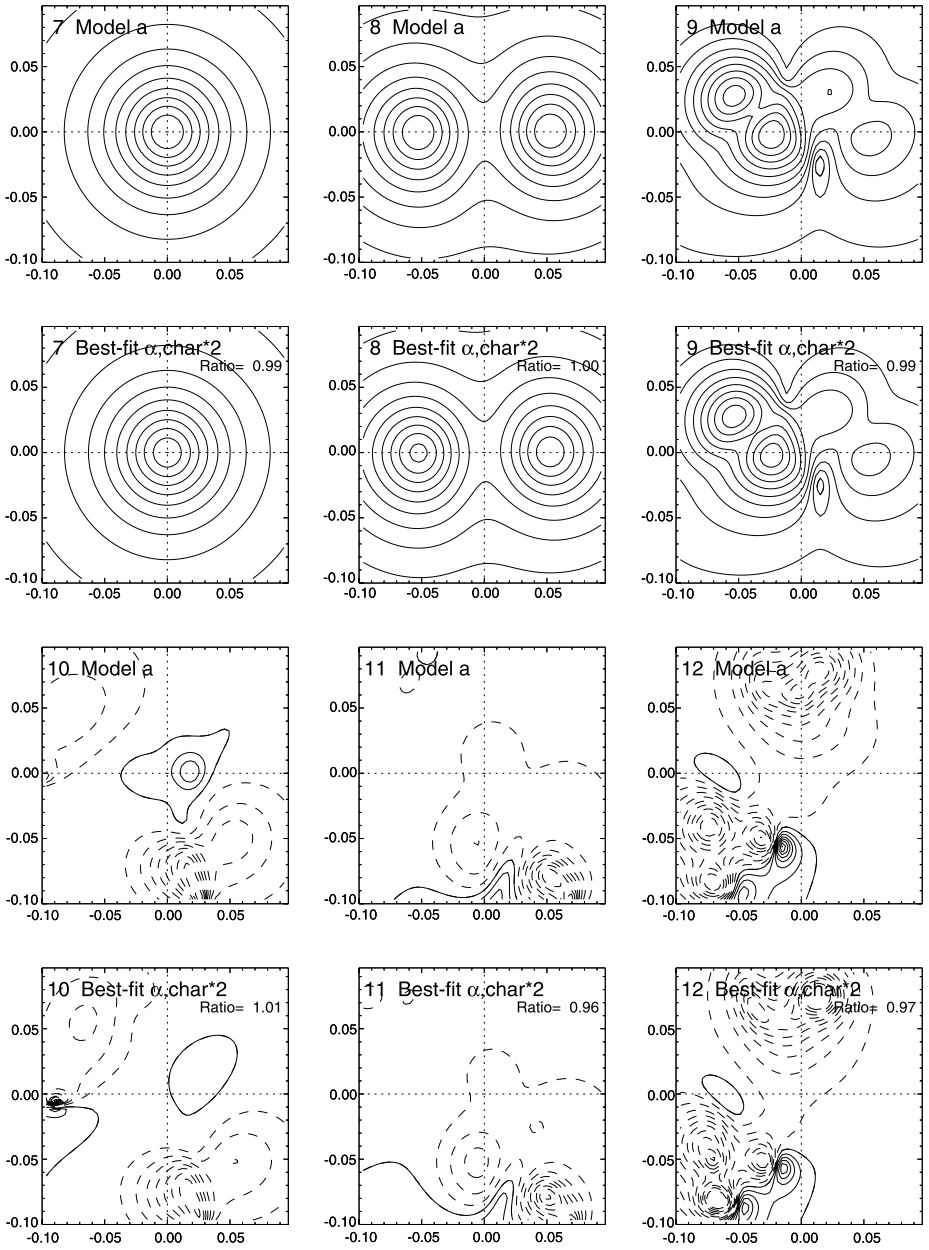


Figure 6 Contour maps of the force-free parameter $|\alpha(x, y)|$ of the simulated models (top and third row) and best-fit solutions (second and fourth row), for the six cases #7 – 12.

The line-of-sight magnetogram $B_z(x, y)$ of the Low and Lou case consists of three smooth patches with an elliptical geometry, where the central patch has a positive magnetic polarity, and the eastern and western patch a negative polarity (see greyscale image in Figure 9 in left panel). The ideal number of decomposed features in the magnetogram is not

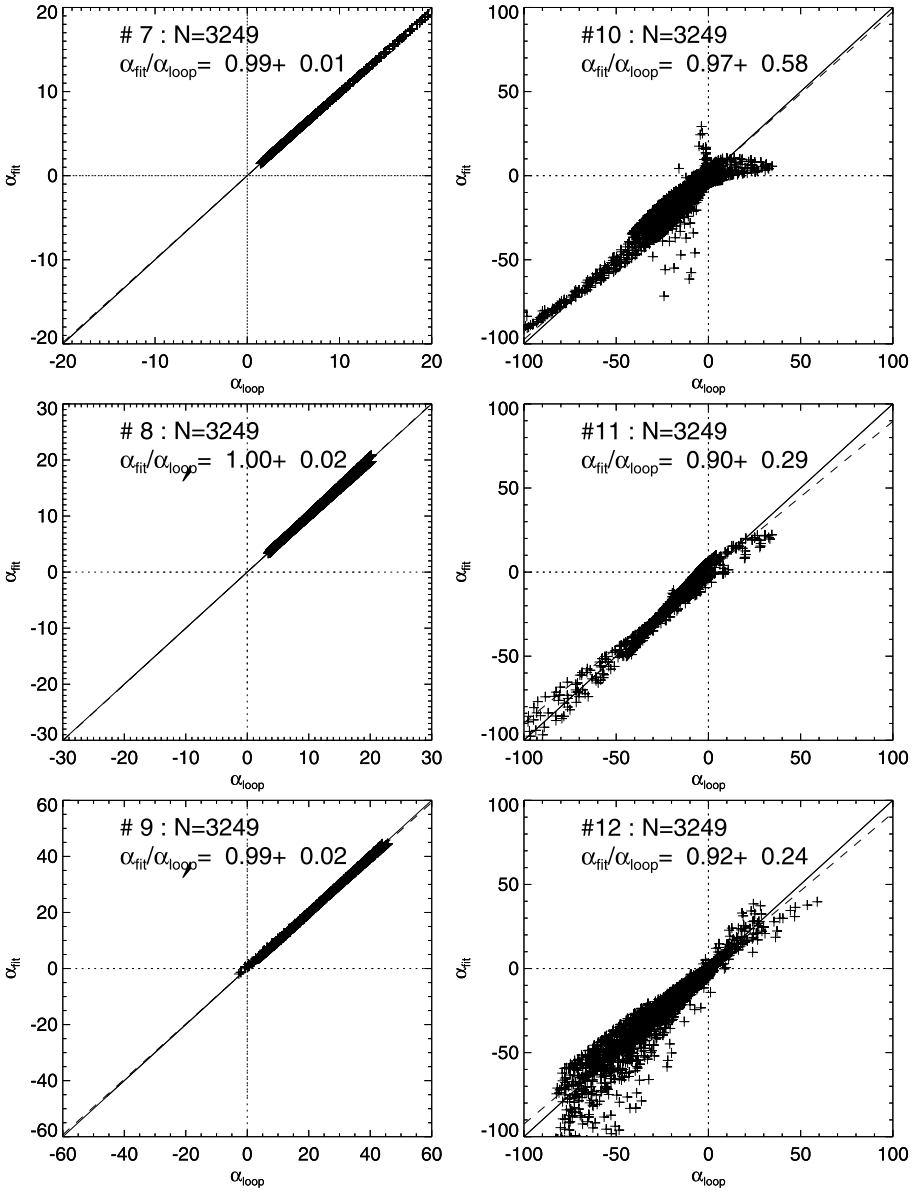


Figure 7 Scatter plot of the best-fit force-free parameters $\alpha_{\rm fit}(x, y, z_{\rm ph})$ of every map pixel (x, y) *versus* the corresponding value $\alpha_{\rm sim}(x, y, z_{\rm ph})$ of the simulated models for the six cases #7 – 12. The mean and standard deviation of the ratio $\alpha_{\rm fit}/\alpha_{\rm sim}$ is indicated in each panel.

known a priori, because a too small number leaves too large residuals of magnetic flux that is not accounted for in the forward-fit, while a too large number leads to overlapping magnetic field components and force-free α-parameter zones, which may jeopardize the quality of forward-fitting (which works best for spatially non-overlapping and independent zones). We show three different trials with $N_{\rm mag} = 4, 10, 50$ in Figure 9. The forward-fitted magne-

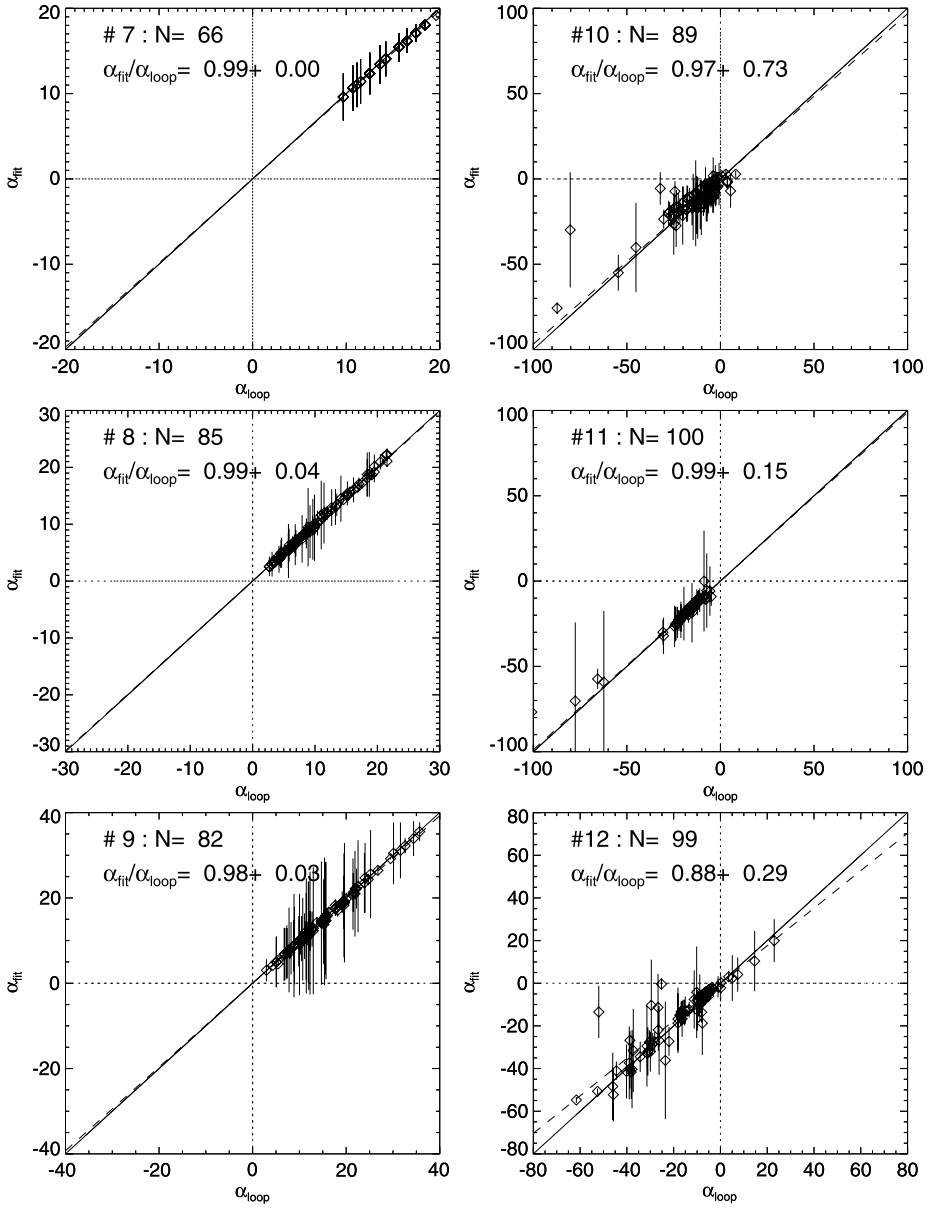


Figure 8 Scatter plot of the best-fit force-free parameters $\alpha_{\rm fit}$ averaged from each fitted coronal loop *versus* the corresponding value $\alpha_{\rm sim}$ of the simulated model loops for the six cases #7 – 12. The vertical error bars indicate the standard deviation of the spatial variation of $\alpha_{\rm fit}(s)$ along each loop. The mean and standard deviation of the ratio $\alpha_{\rm fit}/\alpha_{\rm sim}$ is indicated in each panel.

tograms and the difference images with respect to the input magnetogram are also shown in Figure 9. The residuals in the difference images have a mean and standard deviation of $(B_{\rm fit} - B_{\rm model})/B_{\rm max} = 0.0022 \pm 0.0243$ for $N_{\rm mag} = 4$; -0.0005 ± 0.0082 for $N_{\rm mag} = 10$; and -0.0016 ± 0.0043 for $N_{\rm mag} = 50$, respectively. Thus, the forward-fitted magnetograms

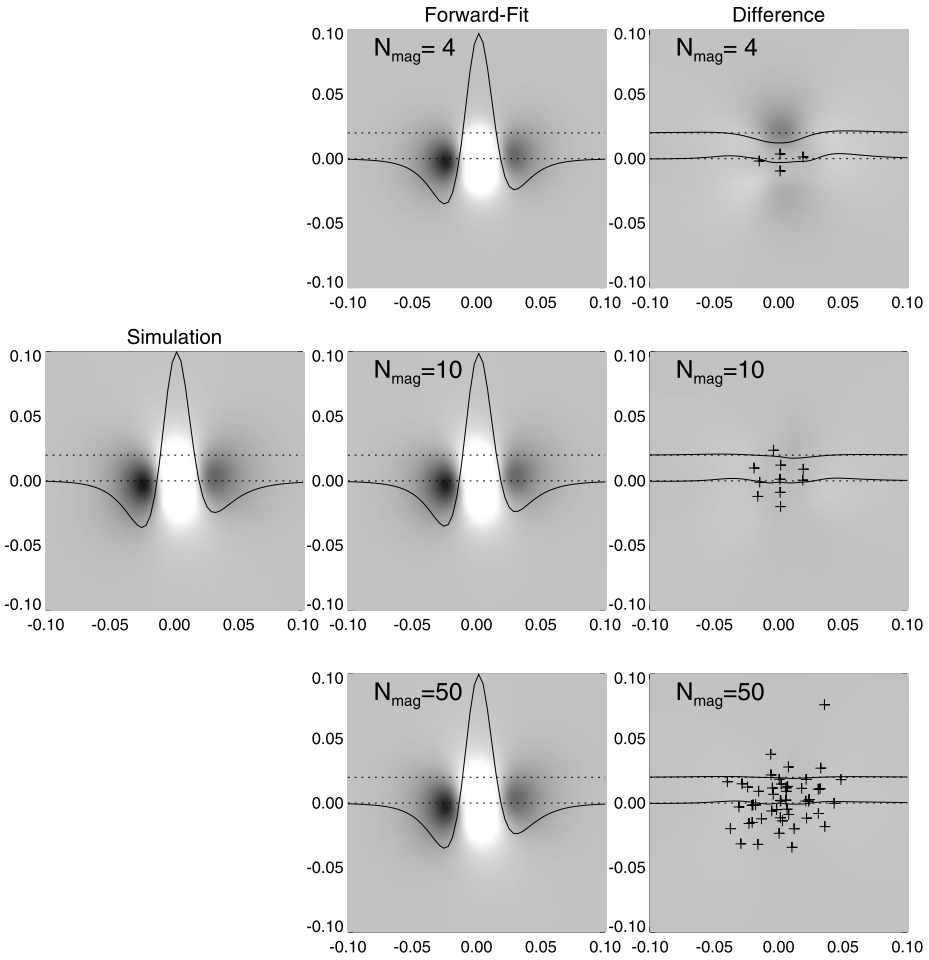


Figure 9 The decomposition of line-of-sight magnetogram $B_z(x, y)$ (simulation in left middle frame) of the Low and Lou (1990) model is shown for three trials with different numbers of magnetic components ($N_{\rm mag} = 4, 10, 50$, first, second, and third row). The locations of the center positions of the magnetic components is shown with crosses in the difference images (right-hand panels). Two profiles across the middle of the magnetogram are also shown (solid curves).

agree with the Low and Lou (1990) model within $\lesssim 1$ % of the magnetic flux. Note that the parameters that decompose the line-of-sight magnetogram make up 80 % of the free parameters in our forward-fitting model, fully determine the potential field extrapolation, but ignore the force-free α-parameters so far. The potential field solution for the Low and Lou (1990) model is shown in Figure 10 (top panel), for a decomposition of $N_{\rm mag} = 50$ magnetic components, for a set of $N_{\rm loop} = 60$ loops. The resulting mean misalignment between the model and the potential field is $\mu = 21.9°$ (Table 6, first line), and $\mu = 30.8°$ for $N_{\rm mag} = 10$, respectively.

We forward-fitted several hundred runs to the Low and Lou (1990) model with different parameter settings (Table 1) and list the results of a selection of four cases in Table 6, and two cases thereof in Figure 10. For $N_{\rm mag} = 50$ and a threshold of Thresh = 0 G we find a so-

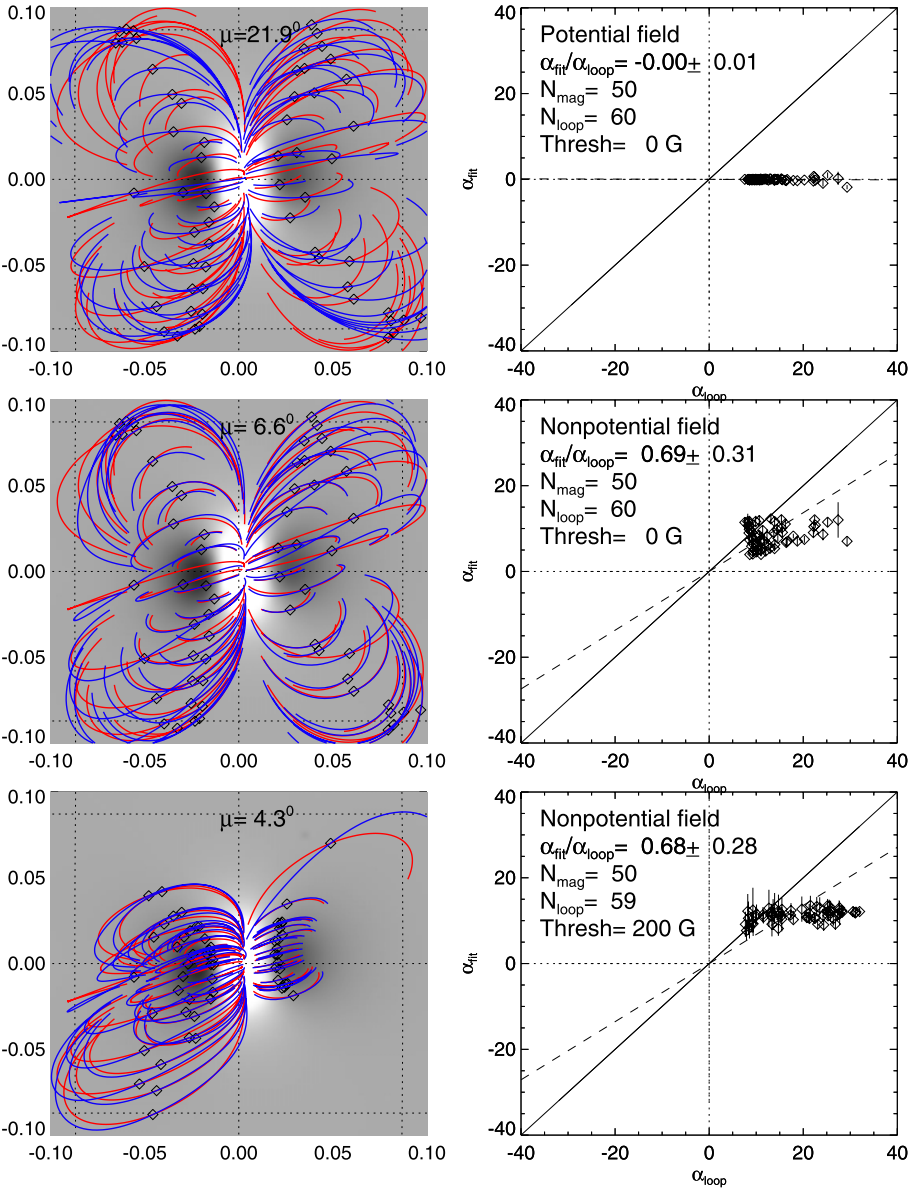


Figure 10 Potential field calculation (top) and forward-fitting of a nonpotential (quasi-NLFFF) model (middle and bottom) to different sets (Thresh = 0, 200 G) of $N_{\rm loop} \approx 60$ coronal loops, which represent an accurate nonlinear force-free field solution of the Low and Lou (1990) model. The model loops are outlined in red color, and the best-fit field lines in blue color. The average misalignment angle μ is indicated in each panel. The photospheric magnetogram is rendered with a greyscale. A scatterplot of the best-fit $\alpha_{\rm fit}$-parameters averaged along each loop versus the model parameters $\alpha_{\rm model}$ are shown in the right-hand panels.

Table 6 Best-fit results of forward-fitting to the Low and Lou (1990) model, using the following parameter settings: the number of magnetic source components $N_{\rm mag} = 10, 50$, the threshold of the magnetic field for selected loops Thresh = 0, 200 G, but otherwise standard settings as listed in Table 1. The results are quantified by the number of fitted loops $n_{\rm loop}$, the mean misalignment angle μ (degrees), the ratio of the fitted to the model input force-free parameter, $\alpha_{\rm fit}/\alpha_{\rm model}$, the divergence-freeness $L_{\rm d}$, the force-freeness $L_{\rm f}$, and the computation time $t_{\rm CPU}$ (s).

$N_{\rm mag}$	Thresh [G]	$n_{\rm loop}$	μ [deg]	$\alpha_{\rm fit}/\alpha_{\rm model}$	$L_{\rm d}$	$L_{\rm f}$	$t_{\rm CPU}$ [s]
50	0	60	21.9°	0.00 ± 0.02	0.000021	0.000023	0
10	0	60	12.7°	0.66 ± 0.43	0.000083	0.000751	257
50	0	60	6.6°	0.69 ± 0.31	0.000045	0.000082	1359
10	200	59	6.1°	0.63 ± 0.22	0.000121	0.000617	123
50	200	59	4.3°	0.68 ± 0.28	0.000084	0.000174	1338

lution that has only a misalignment of $\mu = 6.6°$ (Figure 10, middle panel, and Table 6, third line). This case retrieves the force-free parameter α with an average ratio of $\alpha_{\rm fit}/\alpha_{\rm model} = 0.69 \pm 0.31$ (Figure 10, middle right panel) for the 60 loops shown. The divergence-freeness and force-freeness amount to $L_{\rm d} = 4.5 \times 10^{-5}$ and $L_{\rm f} = 8.2 \times 10^{-5}$. If we select a set of coronal loops with only strong magnetic field strengths at the footpoints (Thresh = 200 G), the misalignment improves to $\mu = 4.3°$ (Figure 10, bottom left panel), while the accuracy of the retrieved α-values remains about the same ($\alpha_{\rm fit}/\alpha_{\rm model} = 0.68 \pm 0.28$ (Figure 10, bottom right panel). It appears that our forward-fitting code always underestimates the values in loops with the highest α-parameter, which was not the case in all of our previous simulations (Figure 8). It appears that the elliptical shape of magnetic patches could be responsible for this underestimate, while it did not occur for spherical shapes of magnetic patches (Simulation runs #7 – 12) described in Section 4. Nevertheless, the achieved small amount of misalignment down to $\mu = 4.3°$ yields a good approximation to a nonlinear force-free field that is sufficiently accurate for most practical purposes of coronal field modeling and can be obtained in a relatively short computation time. The computation times for the five runs listed in Table 6 amounted to $t_{\rm CPU} \approx 2 - 20$ min. We obtained even higher accuracies down to misalignments of $\mu \lesssim 1°$ for smaller subgroups of coronal loops that were localized in partial domains of the active region.

6. Discussion and Conclusions

In this study we developed a numeric code that accomplishes (second-order) nonlinear force-free field fast forward-fitting of combined photospheric magnetogram and coronal loop data. The goal of this code is to compute a realistic magnetic field of a solar active region. Previously developed magnetic field extrapolation codes used either photospheric data only, such as potential-field source surface (PFSS) codes (*e.g.*, Altschuler and Newkirk, 1969) and nonlinear force-free field (NLFFF) codes (*e.g.*, Yang, Sturrock, and Antiochos, 1986; Wheatland, Sturrock, and Roumeliotis, 2000; Wheatland, 2006; Wiegelmann, 2004; Schrijver *et al.*, 2006, 2008; Amari, Boulmezaoud, and Aly, 2006; Valori, Kliem, and Fuhrmann, 2007; Metcalf *et al.*, 2008; DeRosa *et al.*, 2009; Malanushenko, Longcope, and McKenzie, 2009), or (stereoscopically triangulated) coronal loop data only (Sandman *et al.*, 2009; Sandman and Aschwanden, 2011). There are only very few attempts where both photospheric and coronal data constraints were used together to obtain a magnetic field solution,

using either a potential field model with unipolar buried charges that could be forward-fitted to the observed loops (Aschwanden and Sandman, 2010), a linear force-free field (Feng *et al.*, 2007a, 2007b), or a NLFFF code (Malanushenko, Longcope, and McKenzie, 2009, 2011). For special geometries, potential field stretching methods (Gary and Alexander, 1999) or a minimum dissipative rate method for non-force-free fields have also been explored (Gary, 2009).

The new approach of including coronal magnetic field data, in form of stereoscopically triangulated loop 3D coordinates, requires a true forward-fitting approach, while the traditional use of photospheric magnetogram (or vector magnetograph) data represents an extrapolation method from given boundary constraints. Both methods require numerous iterations, and thus are computing-intensive, but the classical forward-fitting method requires a suitable parameterization of a magnetic field model, while extrapolation methods put no constraints on the functional form of the solutions (such as the 3D geometry of magnetic field lines). Thus, the new approach developed here makes use of a parameterization of the 3D magnetic field model in terms of analytical functions that can be fitted relatively fast to the given coronal constraints, but may lack the absolute generality of nonlinear force-free field solutions that NLFFF codes are providing. However, our analytical NLFFF model, which is accurate to second-order (Paper I), probably represents one of the most general parameterizations that is possible with a minimum of free parameters, adapted to uniformly twisted field lines. The parameter space given by this model represents a particular class of quasi-force-free solutions, which is supposed to be most suitable for a superposition of twisted field line structures, but only fitting to real data can reveal how useful and suitable our model is for applications to solar data.

In this study we described the numeric code, which is based on the analytical second-order solutions derived in Paper I, and performed test with 12 simulated cases (six potential and six non-potential), as well as with an analytical NLFFF solution of the Low and Lou (1990) model. The forward-fitting to the 12 simulated cases demonstrated

i) the satisfactory convergence behavior of the forward-fitting code (with mean misalignment angles of $\mu = 3.4^\circ \pm 2.1^\circ$ for potential field cases (see Table 2), and $\mu = 5.1^\circ \pm 4.3^\circ$ for non-potential field cases (see Table 4),
ii) the relatively fast computation speed (from $\lesssim 1$ s to $\lesssim 10$ min), and
iii) the high fidelity of retrieved force-free α-parameters ($\alpha_{\rm fit}/\alpha_{\rm model} \approx 0.9-1.0$; see Figure 8).

The additional test of forward-fitting to the analytical solution of Low and Lou (1990) data yielded similar results, *i.e.*, satisfactory convergence behavior (with mean misalignment angles of $\mu = 4.3^\circ - 6.6^\circ$ for two subsets of loops, see Figure 10),

ii) relatively fast computation speed ($t_{\rm CPU} \approx 2-20$ min), and
iii) the fidelity of retrieved force-free α-parameters ($\alpha_{\rm fit}/\alpha_{\rm model} \approx 0.7 \pm 0.3$; see Figure 10).

The only significant difference of the second test is the trend of underestimating the α-parameter for those loops with the highest α-values, by a factor of $\gtrsim 0.5$. However, if the loops with the highest α-values are fitted individually, the code retrieves the correct α-value. It is not clear whether this feature of the code is related to the geometrical shape of the magnetic concentrations in the magnetogram, which is spherical in our simulation and forward-fitting model, but elliptical in the Low and Lou (1990) case. We simulated the elliptical magnetic sources of the Low and Lou (1990) model by a superposition of spherical sources and found that the code retrieves the correct α-values for each loop (within a few percent accuracy). It is possible that the geometric shape of the Low and Lou (1990)

model, which represents a special class of nonlinear force-free solutions anyway (in terms of Legendre polynomials) cannot efficiently be parameterized with a small number of spherical components, which is the intrinsic parameterization of our code. Anyway, since the Low and Lou (1990) model represents also a very special subclass of nonlinear force-free solutions that may or may not be adequate to model real solar active regions, it may not matter much for the performance of our code with real solar data.

After having tested our numeric code we can proceed to apply it to solar data, such as active regions observed with STEREO since 2007, for which stereoscopic triangulation of coronal loops is available (Feng *et al.*, 2007a, 2007b; Aschwanden and Sandman, 2010; Aschwanden *et al.*, 2012a, 2012b). The second-order NLFFF approximations of our code may be used as an initial guess for other more accurate NLFFF codes, resulting into a significantly shorter computation time. Other future developments may involve the reduction of coronal constraints from 3D to 2D coordinates, which can be furnished by automatic loop tracing codes (*e.g.*, Aschwanden *et al.*, 2008; Aschwanden, 2010; and references therein) and does not require the availability of STEREO data. However, non-STEREO data provide less rigorous constraints for coronal loop modeling, and thus increase the ambiguity of force-free field solutions. Nevertheless, more realistic coronal magnetic field models seem now in the grasp of our computation methods, which has countless benefits for many research problems in solar physics.

Acknowledgements Part of the work was supported by NASA contract NNG 04EA00C of the SDO/AIA instrument and the NASA STEREO mission under NRL contract N00173-02-C-2035.

References

Altschuler, M.D., Newkirk, G. Jr.: 1969, *Solar Phys.* **9**, 131.
Amari, T., Boulmezaoud, T.Z., Aly, J.J.: 2006, *Astron. Astrophys.* **446**, 691.
Aschwanden, M.J.: 2004, *Physics of the Solar Corona. An Introduction*, Praxis Publishing Co., Chichester. Section 5.3.
Aschwanden, M.J.: 2010, *Solar Phys.* **262**, 399.
Aschwanden, M.J.: 2012, *Solar Phys.* doi:10.1007/s11207-012-0069-7; Paper I.
Aschwanden, M.J., Sandman, A.W.: 2010, *Astron. J.* **140**, 723.
Aschwanden, M.J., Lee, J.K., Gary, G.A., Smith, M., Inhester, B.: 2008, *Solar Phys.* **248**, 359.
Aschwanden, M.J., Wülser, J.P., Nitta, N.V., Lemen, J.R., DeRosa, M., Malanushenko, A.: 2012a, *Astrophys. J.*, submitted.
Aschwanden, M.J., Wülser, J.P., Nitta, N.V., Lemen, J.R.: 2012b, *Solar Phys.*, in press.
DeRosa, M.L., Schrijver, C.J., Barnes, G., Leka, K.D., Lites, B.W., Aschwanden, M.J., *et al.*: 2009, *Astrophys. J.* **696**, 1780.
Feng, L., Inhester, B., Solanki, S., Wiegelmann, T., Podlipnik, B., Howard, R.A., Wülser, J.P.: 2007a, *Astrophys. J. Lett.* **671**, L205.
Feng, L., Wiegelmann, T., Inhester, B., Solanki, S., Gan, W.Q., Ruan, P.: 2007b, *Solar Phys.* **241**, 235.
Gary, G.A.: 2009, *Solar Phys.* **257**, 271.
Gary, A., Alexander, D.: 1999, *Solar Phys.* **186**, 123.
Grad, H., Rubin, H.: 1958, In: *Proc. 2nd UN Int. Conf. Peaceful Uses of Atomic Energy* **31**, 190.
Low, B.C., Lou, Y.Q.: 1990, *Astrophys. J.* **352**, 343.
Malanushenko, A., Longcope, D.W., McKenzie, D.E.: 2009, *Astrophys. J.* **707**, 1044.
Malanushenko, A., Yusuf, M.H., Longcope, D.W.: 2011, *Astrophys. J.* **736**, 97.
Metcalf, T.R., Jiao, L., Uitenbroek, H., McClymont, A.N., Canfield, R.C.: 1995, *Astrophys. J.* **439**, 474.
Metcalf, T.R., DeRosa, M.L., Schrijver, C.J., Barnes, G., van Ballegooijen, A.A., Wiegelmann, T., Wheatland, M.S., Valori, G., McTiernan, J.M.: 2008, *Solar Phys.* **247**, 269.
Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T.: 1986, *Numerical Recipes, The Art of Scientific Computing*, Cambridge University Press, Cambridge, 294.
Sandman, A.W., Aschwanden, M.J.: 2011, *Solar Phys.* **270**, 503.
Sandman, A., Aschwanden, M.J., DeRosa, M., Wülser, J.P., Alexander, D.: 2009, *Solar Phys.* **259**, 1.

Schrijver, C.J., DeRosa, M.L., Metcalf, T.R., Liu, Y., McTiernan, J., Regnier, S., Valori, G., Wheatland, M.S., Wiegelmann, T.: 2006, *Solar Phys.* **235**, 161.
Schrijver, C.J., DeRosa, M.L., Metcalf, T., Barnes, G., Lites, B., Tarbell, T., *et al.*: 2008, *Astrophys. J.* **675**, 1637.
Valori, G., Kliem, B., Fuhrmann, M.: 2007, *Solar Phys.* **245**, 263.
Wheatland, M.S.: 2006, *Solar Phys.* **238**, 29.
Wheatland, M.S., Sturrock, P.A., Roumeliotis, G.: 2000, *Astrophys. J.* **540**, 1150.
Wiegelmann, T.: 2004, *Solar Phys.* **219**, 87.
Wiegelmann, T., Inhester, B.: 2003, *Solar Phys.* **214**, 287.
Wiegelmann, T., Inhester, B.: 2006, *Solar Phys.* **236**, 25.
Wiegelmann, T., Neukirch, T.: 2002, *Solar Phys.* **208**, 233.
Wiegelmann, T., Lagg, A., Solanki, S.K., Inhester, B., Woch, J.: 2005, *Astron. Astrophys.* **433**, 701.
Yang, W.H., Sturrock, P.A., Antiochos, S.K.: 1986, *Astrophys. J.* **309**, 383.

DOI 10.1007/978-1-4899-8005-2_23
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0203-6

A Nonlinear Force-Free Magnetic Field Approximation Suitable for Fast Forward-Fitting to Coronal Loops. III. The Free Energy

Markus J. Aschwanden

Received: 11 September 2012 / Accepted: 24 November 2012 / Published online: 21 December 2012

Abstract An analytical approximation of a nonlinear force-free magnetic field (NLFFF) solution was developed in Paper I, while a numerical code that performs fast forward-fitting of this NLFFF approximation to a line-of-sight magnetogram and coronal 3D loops has been described and tested in Paper II. Here we calculate the free magnetic energy $E_{\rm free} = E_{\rm N} - E_{\rm P}$, *i.e.*, the difference of the magnetic energies between the non-potential field and the potential field. A second method to estimate the free energy is obtained from the mean misalignment angle change $\Delta\mu = \mu_{\rm P} - \mu_{\rm N}$ between the potential and non-potential field, which scales as $E_{\rm free}/E_{\rm P} \approx \tan^2(\Delta\mu)$. For the four active regions observed with STEREO in 2007 we find free energies in the range of $q_{\rm free} = (E_{\rm free}/E_{\rm P}) \approx 1\,\%-10\,\%$, with an uncertainty of less than $\pm 2\,\%$ between the two methods, while the free energies obtained from 11 other NLFFF codes exhibit a larger scatter of about $\pm 10\,\%$. We also find a correlation between the free magnetic energy and the GOES flux of the largest flare that occurred during the observing period, which can be quantified by an exponential relationship, $F_{\rm GOES} \propto \exp(q_{\rm free}/0.015)$, implying an exponentiation of the dissipated currents.

Keywords Sun: corona · Sun: magnetic fields

1. Introduction

The free magnetic energy is the maximum amount of energy that can be released in an active region of the solar corona, such as during a solar flare, a filament eruption, or a coronal mass ejection (CME). Therefore, it is important to design reliable methods and tools that can calculate the amount of free energy to quantify the energy budget in a catastrophic energy-release event, as well as for estimating upper limits in forecasting individual events in real-time. Traditionally, the free energy is calculated by computing the non-potential field $\mathbf{B}_{\rm N}(\mathbf{x})$ with a numerical nonlinear force-free field (NLFFF) code and a potential field $\mathbf{B}_{\rm P}(\mathbf{x})$ for the

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

M.J. Aschwanden (✉)
Solar and Astrophysics Laboratory, Dept. ADBS, Lockheed Martin Advanced Technology Center, Bldg. 252, 3251 Hanover St., Palo Alto, CA 94304, USA
e-mail: aschwanden@lmsal.com

same photospheric boundary data $\mathbf{B}(x, y, z_{\rm phot})$, so that the difference of the magnetic field energy density integrated over a volume V that encompasses the active region of interest can be quantified as

$$E_{\rm free} = \frac{1}{8\pi}\left(\int \mathbf{B}_{\rm N}^2(\mathbf{x})\,{\rm d}V - \int \mathbf{B}_{\rm P}^2(\mathbf{x})\,{\rm d}V\right). \tag{1}$$

This standard method may not necessarily reflect the correct amount of maximum free energy released during a solar flare, since the magnetic field in the photospheric boundary $\mathbf{B}(x, y, z_{\rm phot})$ may change during a flare (*e.g.*, see measurements by Wang, Ewell, and Zirin, 1994; Wang *et al.*, 2002, 2004; Wang, Liu, and Wang, 2013; Wang, 1997, 2006; Wang and Liu, 2010). Another problem with NLFFF codes using the photospheric vector field is the non-force-freeness of the lower chromosphere (Metcalf *et al.*, 1995; DeRosa *et al.*, 2009), which can be ameliorated, however, by preprocessing the magnetic boundary data with chromospheric field measurements (*e.g.*, Metcalf, Leka, and Mickey, 2005; Jing *et al.*, 2010), or by a multi-grid optimization that minimizes a joint measure of the normalized Lorentz force and the divergence of the magnetic field, as proposed by Wiegelmann (2004) and applied by Jing *et al.* (2009). A quantitative comparison of NLFFF computation methods, however, revealed a substantial scatter of free energies in the order of $\approx \pm 10$ % (of the potential energy), depending on the numeric code, the boundary specifications, and the spatial resolution (Schrijver *et al.*, 2006; DeRosa *et al.*, 2009). Alternatively, some studies showed that the free energy is better estimated by the minimum-energy state above the linear force-free field with the same magnetic helicity (Woltjer, 1958; Régnier and Priest, 2007). Since NLFFF calculations take very long computing times for forward-fitting tasks to coronal constraints, which requires many iterations (for an overview and discussion of different numerical methods see recent reviews by Aschwanden (2004) or Wiegelmann and Sakurai (2012)), faster non-numerical methods are desirable. Some proxy of the active region's free magnetic energy has been defined based on the twist and magnetic field orientation near the neutral line (*e.g.*, Falconer, Moore, and Gary, 2006, 2011).

Free magnetic energies have been calculated for a variety of solar phenomena, for instance for the evolution of the free magnetic energy during flux emergence and cancellation, using NLFFF codes (*e.g.*, Fang *et al.*, 2012), for the evolution of active regions (*e.g.*, Kusano *et al.*, 2002), for helmet-shaped streamer configurations (Choe and Cheng, 2002), or for breakout CMEs, using MHD simulations (*e.g.*, DeVore and Antiochos, 2005). Theoretical studies quantify the evolution of free magnetic energy for dipolar (Régnier, 2009) and quadrupolar magnetic configurations with a null-point, using different force-free models (Régnier, 2012).

An analytical approximation of a nonlinear force-free magnetic field (NLFFF) solution was developed in Paper I (Aschwanden, 2012), while a numerical code that performs fast forward-fitting of this NLFFF approximation to coronal 3D loops and a line-of-sight magnetogram has been described and tested in Paper II (Aschwanden and Malanushenko, 2012). In this Paper III of the series we calculate the free energy, which is generally defined by the difference of the magnetic energies between the non-potential field and the potential field (Equation (1)). We calculate free energies from simulated data (from Paper II), from the analytical NLFFF solution of Low and Lou (1990), from active region NOAA 10930 during an X3.4 flare modeled by Schrijver *et al.* (2008) and Malanushenko *et al.* (2012), and from stereoscopically triangulated loops observed with STEREO (Aschwanden *et al.*, 2012). Section 2 describes the analytical treatment, Section 3 the application to simulated datasets and observations, Section 4 contains a discussion, and Section 5 the conclusions.

2. Analytical Formulation

2.1. The NLFFF Approximation

In Paper I (Aschwanden, 2012) we derived an analytical approximation of a nonlinear force-free field (NLFFF) solution, which fulfills Maxwell's divergence-free equation ($\nabla \cdot \mathbf{B} = 0$) and the force-free equation ($\nabla \times \mathbf{B}) = \alpha(\mathbf{x})\mathbf{B}$ with second-order accuracy (of the force-free parameter α). The analytical approximation can be specified by a radial field B_r and an azimuthal field component B_φ,

$$B_r(r,\theta) = B_j \left(\frac{d^2}{r^2}\right) \frac{1}{(1 + b^2 r^2 \sin^2\theta)}, \tag{2}$$

$$B_\varphi(r,\theta) = B_j \left(\frac{d^2}{r^2}\right) \frac{br\sin\theta}{(1 + b^2 r^2 \sin^2\theta)}, \tag{3}$$

$$B_\theta(r,\theta) \approx 0, \tag{4}$$

$$\alpha(r,\theta) \approx \frac{2b\cos\theta}{(1 + b^2 r^2 \sin^2\theta)}, \tag{5}$$

where (r, φ, θ) are the spherical coordinates of a single magnetic field component (B_j, x_j, y_j, z_j, b_j) with a unipolar magnetic charge B_j that is buried at position (x_j, y_j, z_j), has a depth $d = 1 - [x_j^2 + y_j^2 + z_j^2]^{1/2}$, a vertical twist $\alpha = 2b_j$, and $r = [(x - x_j)^2 + (y - y_j)^2 + (z - z_j)^2]^{1/2}$ is the distance of an arbitrary coronal position (x, y, z) to the subphotospheric location (x_j, y_j, z_j) of the buried magnetic charge. The force-free parameter α is expressed in terms of the parameter b, which quantifies the number $N_{\rm twist}$ of full-twist turns over a (loop) length L,

$$b = \frac{2\pi N_{\rm twist}}{L}. \tag{6}$$

This analytical approximation is divergence-free and force-free to second-order accuracy in the parameter ($br\sin\theta$), which is approximately proportional to the force-free parameter α as defined by Equation (5).

A general magnetic field configuration can be composed by superposing $N_{\rm m}$ twisted magnetic field components,

$$\mathbf{B}_{\rm N}(\mathbf{x}) = \sum_{j=1}^{N_{\rm m}} \mathbf{B}_j(\mathbf{x}), \tag{7}$$

which also fulfills the divergence-free and force-free condition with second-order accuracy in α (or b).

2.2. The Free Magnetic Energy of a Single Twisted Component

We now calculate the free magnetic energy $dE_{\rm free}(\mathbf{x})$ at location $(\mathbf{x})$ for the field that results from a single twisted (buried) magnetic charge as defined by Equations (1) to (6). Since the radial B_r and azimuthal components B_φ are always orthogonal to each other (Figure 1), we can calculate the total non-potential magnetic field strength $B_{\rm N}$ at every given point (r, φ) simply from the sum of the squared components B_r and B_φ,

$$B_{\rm N} = \left(B_r^2 + B_\varphi^2\right)^{1/2} = B_j \left(\frac{d^2}{r^2}\right) \frac{1}{\sqrt{(1 + b^2 r^2 \sin^2\theta)}}, \tag{8}$$

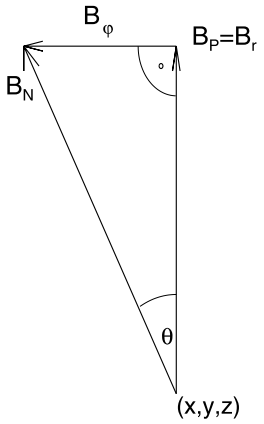


Figure 1 Diagram of the non-potential field vector $\mathbf{B}_{\rm N}$, which is composed of the two orthogonal components of the azimuthal field vector $\mathbf{B}_{\varphi}$ and the radial field vector $\mathbf{B}_r$, subtending an angle θ.

while the field strength $B_{\rm P}$ of a potential field corresponds to the radial component B_r (of a single buried magnetic charge, Equation (2)),

$$B_{\rm P} = B_r = B_j \left(\frac{d^2}{r^2}\right) \frac{1}{(1 + b^2 r^2 \sin^2\theta)}, \tag{9}$$

and thus the free energy $dE_{\rm free}(\mathbf{x})$ is just the magnetic energy associated with the azimuthal field component B_{φ}, with Equations (8) and (9),

$$dE_{\rm free}(\mathbf{x}) = dE_{\rm N}(\mathbf{x}) - dE_{\rm P}(\mathbf{x}) = \frac{1}{8\pi}\left[B_r(\mathbf{x})^2 + B_{\varphi}(\mathbf{x})^2\right] - \frac{1}{8\pi} B_r(\mathbf{x})^2 = \frac{1}{8\pi} B_{\varphi}(\mathbf{x})^2. \tag{10}$$

This definition of the free magnetic energy $dE_{\rm free}(\mathbf{x})$ fulfills the following conditions:

i) *Positivity constraint*: The free energy is positive for every nonpotential field at any location $\mathbf{x}$, *i.e.*, $dE_{\rm free}(\mathbf{x}) > 0$, since $B_{\varphi}^2(\mathbf{x}) \geq 0$ and $dE_{\rm free}(\mathbf{x}) \propto B_{\varphi}^2(\mathbf{x})$ according to Equation (10).

ii) *Additivity of energies and orthogonality of magnetic field components*: The nonpotential energy corresponds to the sum of the potential energy and the free energy $dE_{\rm N}(\mathbf{x}) = dE_{\rm P}(\mathbf{x}) + dE_{\rm free}(\mathbf{x})$ according to Equation (1) and Figure 1. Since the energies scale with the square of the magnetic field components, *i.e.*, $dE_{\rm N}(\mathbf{x}) \propto B_{\rm N}(\mathbf{x})^2$, $dE_{\rm P}(\mathbf{x}) \propto B_r(\mathbf{x})^2$, and $dE_{\rm free}(\mathbf{x}) \propto B_{\varphi}(\mathbf{x})^2$, the Pythagoras theorem can be applied and it follows that B_r, B_{φ}, and $B_{\rm N}$ form a rectangular triangle, with B_{φ} perpendicular to the potential field component $B_{\rm P} = B_r$.

iii) *Positive scaling with force-free parameter*: The free energy quantifies an excess of nonpotential magnetic field energy compared with the potential field energy, and thus should have a positive scaling with the force-free parameter α. Since the free energy scales proportional to the square of the azimuthal magnetic field component (Equation (10)), we have a positive scaling, which is approximately $dE_{\rm free}(\mathbf{x}) \propto b^2 \propto \alpha^2$ (Equations (3) and (5)).

iv) *Potential field limit*: The free energy vanishes asymptotically ($dE_{\rm free}(\mathbf{x}) \to 0$) with vanishing force-free parameter $|\alpha| \to 0$ or $|b| \to 0$, with the potential field being the limit, $dE_{\rm N}(\mathbf{x}) \to dE_{\rm P}(\mathbf{x})$.

v) *Finiteness of nonpotential energy*: The total nonpotential magnetic energy integrated over an arbitrary large height r converges to a finite value, $E(r \to \infty) = E_{\rm max}$. We can

prove the finiteness of the potential energy for a single magnetic charge, which has a square-dependence of the magnetic field, $B(r) \propto r^{-2}$, yielding a 4th-power dependence of the magnetic energy $dE_{\rm P}(r) \propto B(r)^2 \propto r^{-4}$, and thus a 3th-power dependence for the integrated magnetic energy, $E_{\rm P,tot} \propto \int dE_{\rm P}(r)\,{\rm d}r \propto r^{-3}$. For a finite amount of twist and a finite number of magnetic sources, it can be shown that the integral of the resulting nonpotential energy is also finite.

2.3. The Free Magnetic Energy of Multiple Twisted Components

While the foregoing definition of the free energy is calculated for a single twisted (buried) magnetic charge, how can it be generalized for a superposition of an arbitrary number of magnetic charges, as defined in Equation (7)? The sum of the magnetic field contributions from each buried magnetic charge component $\mathbf{B}_j(\mathbf{x})$ add up to the non-potential field vector $\mathbf{B}_{\rm N}(\mathbf{x})$, which can be decomposed into two orthogonal components $\mathbf{B}_{\parallel}(\mathbf{x})$ and $\mathbf{B}_{\perp}(\mathbf{x})$ in every point of space $(\mathbf{x})$,

$$\mathbf{B}_{\rm N}(\mathbf{x}) = \sum_{j=1}^{N_{\rm m}} \mathbf{B}_j(\mathbf{x}) = \mathbf{B}_{\parallel}(\mathbf{x}) + \mathbf{B}_{\perp}(\mathbf{x}), \tag{11}$$

where the parallel component is aligned with the potential field direction, $\mathbf{B}_{\parallel} \parallel \mathbf{B}_{\rm P}$, and the perpendicular component is orthogonal to the potential field direction, $\mathbf{B}_{\perp} \perp \mathbf{B}_{\rm P}$. For a single (twisted) magnetic component the parallel component $\mathbf{B}_{\parallel}$ is identical with the radial component $\mathbf{B}_r$, and the perpendicular component $\mathbf{B}_{\perp}$ is identical with the azimuthal component $\mathbf{B}_{\varphi}$ (Figure 1). The three magnetic field components $B_{\parallel} = B_{\rm P}$, $B_{\perp}$, and $B_{\rm N}$ are then associated each with one of the three energy components,

$$\begin{aligned} dE_{\rm free}(\mathbf{x}) &= (1/8\pi) B_{\perp}^2(\mathbf{x}), \\ dE_{\rm P}(\mathbf{x}) &= (1/8\pi) B_{\parallel}^2(\mathbf{x}), \\ dE_{\rm N}(\mathbf{x}) &= (1/8\pi) \left[B_{\parallel}^2(\mathbf{x}) + B_{\perp}^2(\mathbf{x}) \right]. \end{aligned} \tag{12}$$

Alternatively, the magnetic energy $dE_{\rm P}(\mathbf{x})$ of the potential field can be computed by using current-free magnetic field components $(B_j, x_j, y_j, z_j, b_j = 0)$ straightforward with Equation (1), and the magnetic energy $dE_{\rm N}(\mathbf{x})$ of the non-potential field with the current components $(B_j, x_j, y_j, z_j, b_j \neq 0)$ with Equation (1) as well, which yields the free energies $(E_{\rm free} = E_{\rm N} - E_{\rm P})$, after volume integration. Both methods are fitting the same line-of-sight component of the photospheric boundary $B_z(x, y, z_{\rm phot})$ given by the magnetogram, while the non-potential magnetic field affects the transverse field components $B_x(x, y, z_{\rm phot})$ and $B_y(x, y, z_{\rm phot})$, which are not used as a boundary condition in our forward-fitting method.

2.4. Free Energy Estimated from Misalignment Angle with Loops

An alternative method is to estimate the free energy in an active region from the misalignment angle μ between the potential field and the observed coronal loops, or the best-fit non-potential field. Using the coronal loops as a proxy for the non-potential field, this would provide a very fast method that requires only the computation of a potential field, supposing we have the 3D coordinates of coronal loops, *e.g.*, from stereoscopic triangulation. Since the mean azimuthal field component is $B_{\perp} = B_{\rm N} \sin(\Delta\mu)$ and the mean radial (potential) field component is $B_{\parallel} = B_{\rm N} \cos(\Delta\mu)$, the non-potential energy ratio $q_{\mu} = E_{\rm N}/E_{\rm P}$ follows then

directly, using the definition of Equation (11), for the free energy,

$$q_\mu = \frac{E_N}{E_P} = \frac{B_\perp^2 + B_\parallel^2}{B_\parallel^2} = 1 + \tan^2(\Delta\mu), \tag{13}$$

where the relative misalignment angle $\Delta\mu$ is defined as the difference between the (median) potential μ_P and (median) non-potential field directions μ_N,

$$\Delta\mu = \mu_P - \mu_N. \tag{14}$$

Ideally, if the NLFFF forward-fitting code perfectly matches the coronal loops (with $\mu_N \approx 0$), the relative misalignment angle of the potential field to the loops can be used, $\Delta\mu \approx \mu_P$. However, in reality there is always a significant difference between the best-fit NLFFF solution and the observed loop data, either due to stereoscopic measurement errors or to an additional field misalignment that cannot be described with the particular parameterization of our NLFFF approximation.

Ideally, if the NLFFF forward-fitting code matches the coronal loops perfectly (with $\mu_N \approx 0$), and thus the relative misalignment angle of the potential field to the loops can be used, $\Delta\mu \approx \mu_P$. However, in reality there is always a significant difference between the best-fit NLFFF solution and the observed loop data, either due to stereoscopic measurement errors or due to an additional field misalignment that cannot be described with the particular parameterization of our NLFFF approximation.

3. Numerical Tests and Results

We now test the calculation of the free magnetic energy with the definitions given in Equations (12) and (13) for four different datasets, using i) simulated data produced by our analytical NLFFF approximation, ii) the NLFFF solution of Low and Lou (1990) with a known exact analytical solution, iii) active region NOAA 10930 during an X3.4 flare modeled Schrijver *et al.* (2008) and Malanushenko *et al.* (2012), and iv) stereoscopic data from four observed active regions.

3.1. Tests with Simulated Data (P1 – N12)

We simulated six cases of potential fields (shown in Figure 3 of Paper II), and six cases with non-potential fields (Figures 2 and 3), which are similar to the cases N7 to N12 in Paper II, except that we reduced the amount of twist by a factor of five to better compare them with observations of real active regions, which have free energies of about < 10 %. We used the parameterization of our analytical NLFFF approximation described in Paper I, which contains 1, 2, 4, and 10 magnetic source components with variable vertical twist, labeled as cases P1 – P6 and N7 – N12. We integrated the non-potential magnetic energy E_N in a box that covers the displayed field-of-view centered at the center of the solar disk, which has a height range of $h = 0.15$ solar radii above the photosphere. The values of the potential energies E_P and the ratio of the non-potential to the potential energy of the model ($q_{\rm model}$), or of the fit ($q_{\rm fit} = E_N/E_P$), are listed in Table 1. The free energy is $E_{\rm free} = E_N - E_P = E_P(q_N - 1)$.

Table 1 demonstrates several results. First, the misalignment angle between the forward-fitted non-potential field and the simulated loops is significantly smaller ($\mu_N = 1.4° \pm 0.8°$)

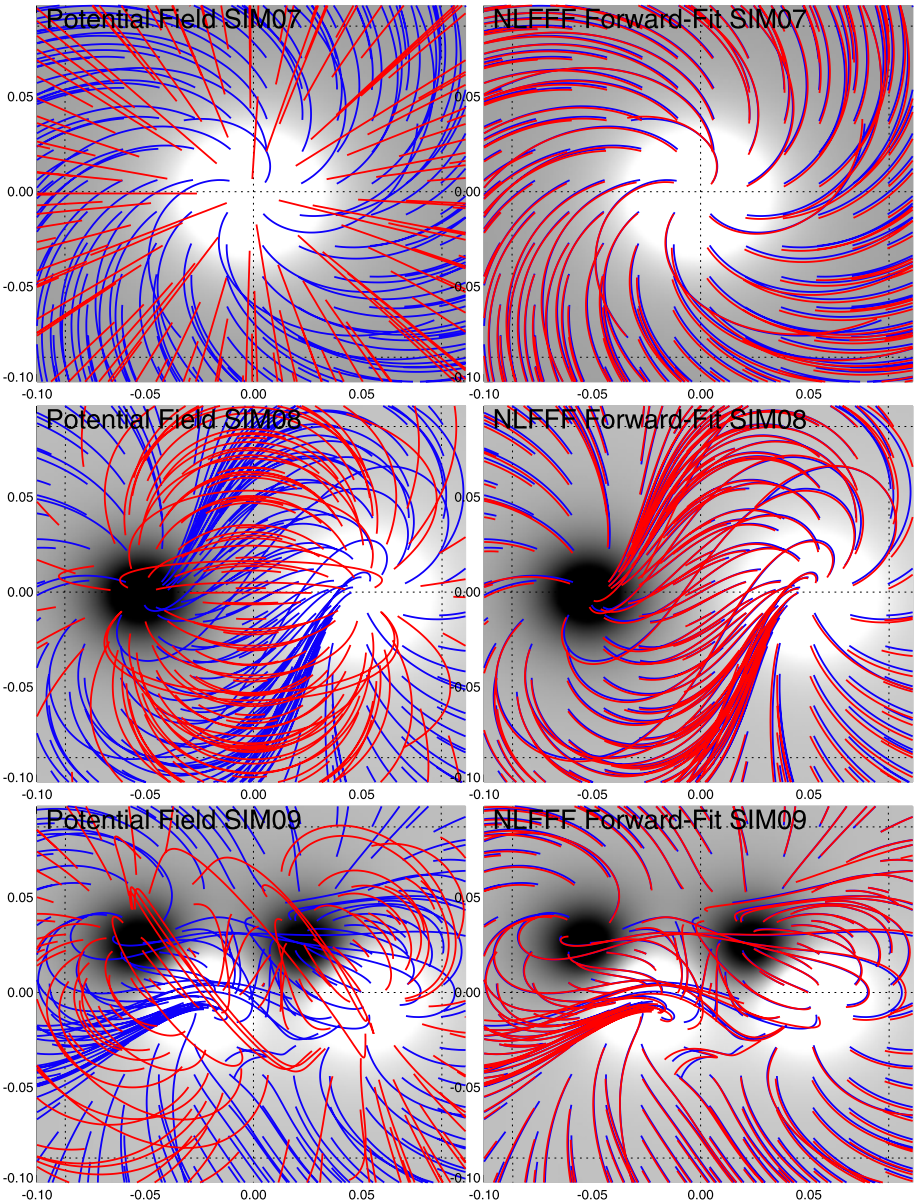


Figure 2 Forward-fitting of simulated non-potential field data (cases N7 – N9). Each panel shows the line-of-sight magnetogram (gray), simulated coronal loops (the targets of the forward fit) (blue), and magnetic field lines of a theoretical model (red), either the potential field constrained by the line-of-sight magnetogram (left panels), or the forward fit of the NLFFF approximation (right panels).

in all 12 cases (P1, . . . , N12) than the initial potential field misalignment angle $\mu_P = 9.6° \pm 8.5°$, which confirms a satisfactory convergence of the forward-fit to the simulated target loops. Second, all ratios of the non-potential to potential energies E_N/E_P are higher than one, which means that the free energy $E_{free} = E_N - E_P$ is always positive

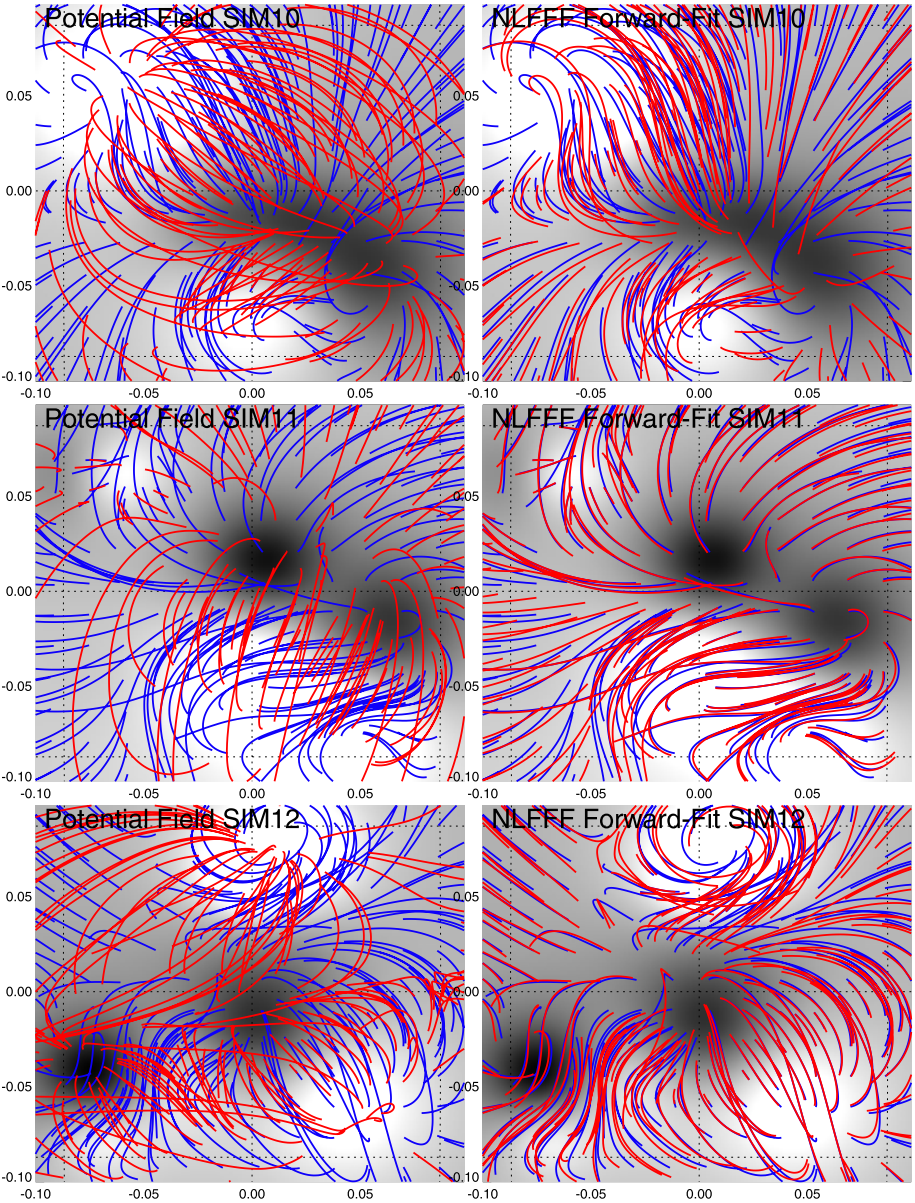


Figure 3 Forward-fitting of cases N10 – N12. Representation similar to Figure 2.

with the definition given in Equation (12). Third, all 12 simulated cases agree in the non-potential energy ratio with the simulated input model, *i.e.*, $q_{\rm fit} = q_{\rm model} = E_{\rm N}/E_{\rm P}$ with an accuracy better than 10^{-5}, which also confirms the perfect convergence of the forward-fit algorithm. The non-potential-field cases (N7 – N12) have free energies in the range of $q_{\rm fit} = E_{\rm N}/E_{\rm P} = 1.010 - 1.163$, or up to 16 % of the potential-field energy, which are also retrieved with an accuracy better than 10^{-5}. We also compare the non-potential energy ra-

Table 1 Magnetic energy calculations are listed for 20 cases, including six potential-field cases (P1 – P6), six non-potential-field cases (N7 – N12), the Low and Lou (1990) cases (L1, L2), the Schrijver *et al.* (2008) cases (S1, S2), and four stereoscopically observed active regions (A, B, C, and D), specified by the number of magnetic charges ($N_{\rm m}$), the fraction of magnetic energy captured by the model ($q_{\rm E}$), the number of simulated loops $N_{\rm loop}$, the potential energy $E_{\rm P}$ (corrected by the factor q_E), the median misalignment angle $\mu_{\rm P}$ of the potential field, the median misalignment angle $\mu_{\rm N}$ after forward-fitting of the NLFFF model, the predicted energy ratio $q_\mu = (1 + \tan^2(\Delta\mu))$ based on the misalignment-angle change $\Delta\mu = \mu_{\rm P} - \mu_{\rm N}$, the forward-fitted non-potential energy ratio $q_{\rm fit} = E_{\rm N}/E_{\rm P}$, and the volume-integrated non-potential magnetic energy ratio of the model $q_{\rm model} = E_{\rm model}/E_{\rm P}$, with values computed by Anna Malanushenko (private communication, 2012) using the Low and Lou (1990) data ([a]), the value of 3D-fits labeled as II.b in Tables 3 and 4 of Malanushenko *et al.* (2012) ([b]), and the value of the Wh^+ code with the smallest misalignment angle $\mu = 24^\circ$ in Table 1 of DeRosa *et al.* (2009) ([c]).

Case	$N_{\rm m}(q_{\rm E})$	$N_{\rm loop}$	$\mu_{\rm P}$	$\mu_{\rm N}$	$E_{\rm P}$ (10^{32} erg)	q_μ	$q_{\rm fit}$	$q_{\rm model}$
P1	1(1.000)	49	0.5	0.0	4.65	1.000	1.000	1.000
P2	2(1.000)	121	1.8	1.2	5.84	1.000	1.000	1.000
P3	4(1.000)	121	2.1	1.4	4.27	1.000	1.000	1.000
P4	10(1.000)	121	2.5	1.3	14.6	1.000	1.000	1.000
P5	10(1.001)	121	3.7	1.4	12.9	1.002	1.000	1.000
P6	10(0.998)	121	3.3	1.9	6.80	1.001	1.000	1.000
N7	1(1.000)	49	18.8	0.2	4.65	1.114	1.010	1.010
N8	2(1.000)	121	10.4	1.2	5.84	1.026	1.009	1.009
N9	4(1.000)	121	13.0	1.8	4.27	1.039	1.016	1.016
N10	10(1.000)	121	17.4	3.1	14.6	1.065	1.082	1.082
N11	10(1.001)	121	17.2	1.1	12.9	1.084	1.113	1.113
N12	10(0.998)	121	25.3	1.9	6.80	1.188	1.163	1.163
L1	100(0.913)	133	14.7	4.8	0.246	1.030	1.023	1.129[a]
L2	100(0.910)	35	3.9	1.6	0.028	1.002	1.001	1.091[a]
S1: 2006/12/12	200(0.860)	331	37.3	14.4	18.3	1.179	1.112	1.21 ± 0.05[b]
S2: 2006/12/13	200(0.878)	98	27.3	13.4	17.8	1.061	1.104	1.08 ± 0.01[b]
A: 2007/04/30	100(0.854)	200	23.8	21.4	12.7	1.002	1.006	1.030[c]
B: 2007/05/09	100(0.807)	70	21.2	17.9	1.08	1.003	1.023	
C: 2007/05/19	100(0.784)	100	41.3	22.1	3.47	1.121	1.085	
D: 2007/12/11	100(0.810)	87	22.1	14.7	7.83	1.017	1.044	

tios calculated from the forward-fit ($q_{\rm fit}$) that were estimated from the median misalignment angles (q_μ; Equation (13)) and find an agreement of $q_\mu/q_{\rm fit} = 1.02 \pm 0.05$ between the two methods.

We also tested the finiteness of the non-potential energy. In Figure 4 we plot the height-integrated total non-potential energies $E_{\rm N}(r)$ as a function of the height limit r for the six cases N7 – N12 (Table 1) and find that each one approximately follows an exponential height dependence (dashed curves in Figure 4),

$$E_{\rm P}(r) \approx E_{\rm max}\left[1 - \exp\left(-\frac{r-1}{\lambda}\right)\right], \tag{15}$$

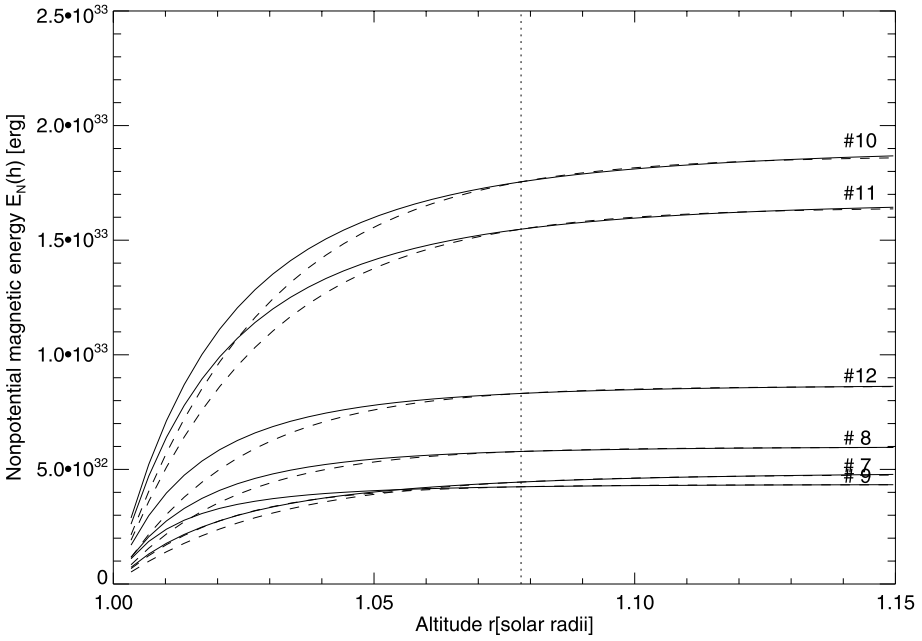


Figure 4 Height dependence of the non-potential energy $E_{\rm N}(h)$ integrated over a volume from height $r = 1.0$ to $r = r_{\rm max}$ with $r_{\rm max} = 1.15$ solar radii for the six non-potential cases (N7 – N12 in Table 1). An exponential function is fitted in the upper half height range $r > 1.075$ (dotted line).

with the finite limit $E(r \rightarrow \infty) = E_{\rm max}$, which confirms the convergence of the code. In our calculations we generally used a height limit of $h_{\rm max} = 0.15$ solar radii, which corresponds to about two density scale heights of $T_{\rm e} = 1.0$ MK plasma, or four emission-measure scale heights.

3.2. Tests of the Low and Lou (1990) NLFFF Solution

An analytically exact solution of an NLFFF model was derived by Low and Lou (1990), described also in Malanushenko, Longcope, and McKenzie (2009). The particular solutions we used are defined by the parameters ($a = 0.6$, $n = 2.0$) for L1, and ($a = 0.01$, $n = 1.0$) for L2, where a is the Grad–Shafranov constant, n is the harmonic number of the Legendre polynomial, and additional parameters are the depth l of the source below the photosphere and the inclination angle ϕ of the axis of symmetry). Forward-fits of our analytical NLFFF approximation to the exact NLFFF solution of Low and Lou (1990) are shown in Figure 5 (top and middle), where we used a computation box of ($100 \times 100 \times 75$), with a pixel size of $\Delta x = 0.002$ solar radii and a height range of $h_{\rm max} = 75 \Delta x \approx 0.15$ solar radii. The forward-fit was accomplished by using the line-of-sight magnetogram at a planar surface ($B_z(x, y, z) = 1$) and a set of $N_{\rm loop} = 133$ (L1) and $N_{\rm loop} = 35$ (L2) target field lines that mimic coronal loops. The misalignment angle was reduced from $\mu_{\rm P} = 14.7^\circ$ to $\mu_{\rm N} = 4.8^\circ$ for L1, and from $\mu_{\rm P} = 3.9^\circ$ to $\mu_{\rm N} = 1.6^\circ$ for L2, which means that the forward-fitting reduces the misalignment by about a factor of 3. The resulting non-potential energy ratios are listed in Table 1, yielding free energies of 2.3 % for L1, and 0.1 % for L2. These free energies are significantly below the theoretical values calculated in Malanushenko, Longcope,

and McKenzie (2009), where values of 13 % and 9 % are quoted (listed under $q_{\rm model}$ in Table 1), although the misalignment of the forward-fitted field is quite satisfactory. The reasons for the mismatch in the free energy for this particular case is not fully understood, since the convergence behavior of our code seems to be no problem for unique solutions (such as the simulated cases N7 – N12). We suspect that the parameterization of our NLFFF approximation, which consists of a number of buried point-like magnetic sources, is not adequate or suitable to represent the analytical Low and Lou (1990) solution, which consists of extended, smooth magnetic distributions with elliptical shapes, parameterized in terms of Legendre polynomials. It is conceivable that the representation of Legendre polynomials by (spherically symmetric) point sources leads to clustering of closely spaced point sources, which cancels the non-potential (azimuthally twisted) field components.

3.3. Tests of the Schrijver *et al.* (2008) Case

The magnetic field of active region NOAA 10930, observed with TRACE, *Hinode*/XRT, and *Hinode*/SOT on 12 December 2006, 20:30 UT before a GOES-class X3.4 flare (case S1), and on 13 December 2006, 03:40 UT after the flare (case S2), has been extensively modeled with NLFFF codes (Schrijver *et al.*, 2008; Malanushenko *et al.*, 2012). We show a forward-fit of our NLFFF approximation in Figure 5 (bottom panels) to a set of loops (*i.e.*, closed field lines that are randomly chosen from an NLFFF solution computed by Malanushenko *et al.*, 2012). The accuracy of the forward-fitting mostly depends on the number of the magnetic field components $N_{\rm m}$, but generally reaches asymptotically a flat plateau for $N_{\rm m} \gtrsim 100$ (Aschwanden *et al.*, 2012; Figure 10). For this magnetically very complex active region we needed $N_{\rm m} = 200$ magnetic source components to reach the plateau, while $N_{\rm m} \lesssim 100$ was sufficient for all other cases. Here, the median misalignment angle of $\mu_{\rm P} = 37.3°$ for the potential field was reduced by about a factor of 2.6 to $\mu_{\rm N} = 14.4°$ for the best-fit non-potential NLFFF solution (S1 in Table 1). We measure a potential-field energy of $E_{\rm P} = 1.83 \times 10^{33}$ erg before the flare (S1), and $E_{\rm P} = 1.78 \times 10^{33}$ erg after the flare (S2), which means a small difference of ≈ 2.6 %. For the non-potential magnetic energy ratio we measure $q_{\rm fit} = E_{\rm N}/E_{\rm P} = 1.112$ before the flare and $q_{\rm fit} = 1.104$ after the flare, and similar values with the misalignment method, *i.e.*, $q_{\mu} = 1.179$ and $q_{\mu} = 1.061$. Thus the total non-potential energy decreases by $\Delta E_{\rm N} = 0.7 \times 10^{32}$ erg (≈ 5 % according to the forward-fit method), or by $\Delta E_{\rm N} = 2.7 \times 10^{32}$ erg (≈ 13% according to the misalignment method).

From the same observing times the free-energy ratio was measured with 14 different NLFFF codes in Schrijver *et al.* (2008; Table 1 therein), which yield energy ratios of $E_{\rm N}/E_{\rm P,free} = 1.05 \pm 0.05$ before the flare, and $E_{\rm N}/E_{\rm P,free} = 1.16 \pm 0.14$ after the flare, if we average all methods with equal weight. However, the most reliable NLFFF method among them, according to a quality assessment by visual inspection of five magnetic features seems to be the Wh^{+}_{pp} NLFFF code, which yields an energy ratio of $E_{\rm N}/E_{\rm P,pre} = 1.32$ before the flare and $E_{\rm N}/E_{\rm P,pre} = 1.19$ after the flare, that is, a decrease of 13 % in the free energy, corresponding to a drop of $\Delta E_{\rm free} \approx 3 \times 10^{32}$ erg in free energy, similar to our measurement with the misalignment method ($E_{\rm free} = 2.7 \times 10^{32}$ erg or 13 %). Similarly, Malanushenko *et al.* (2012; cases with 3D fits labeled II.b in Table 1 therein) calculated energy ratios of $E_{\rm N}/E_{\rm P} = 1.21 \pm 0.05$ before the flare and $E_{\rm N}/E_{\rm P} = 1.08 \pm 0.01$ after the flare, corresponding to a drop of 13 % in the free energy.

Thus all three studies agree with a drop in free energy, by an amount of 5 % – 13 % according to our two methods, *versus* ≈ 13 % for the most reliable NLFFF codes, while lower free energy values and larger misalignment angles result for the other NLFFF codes (Table 1 in Schrijver *et al.*, 2008). Thus our two methods appear to be commensurable with the most reliable NLFFF codes.

3.4. Tests with Stereoscopic Observations

We now calculate the free magnetic energy for four active regions (Table 1) that have all been observed with STEREO and were subjected to previous magnetic modeling, which we label as active regions A, B, C, and D. All four active regions have been subject of potential-field modeling using stereoscopic data, including potential-field stretching (Sandman *et al.*, 2009), buried magnetic charges (Aschwanden and Sandman, 2010), or buried dipoles (Sandman and Aschwanden, 2011), and nonlinear force-free field modeling (Aschwanden *et al.*, 2012).

Active region NOAA 10953, 30 April 2007 (A) has been observed during a flare-energy build-up phase, and a flare trigger by magnetic reconnection at a 3-D null point of a separatrix surface was identified (Su *et al.*, 2009). The magnetic modeling of this active region has been scrutinized with 11 different NLFFF codes, using SoHO/MDI, *Hinode*/SOT-SP, XRT, and STEREO/EUVI data (DeRosa *et al.*, 2009; Su *et al.*, 2009). For this active region (shown in Figure 6, top panels) we measure a potential-field energy of $E_{\rm P} = 1.27 \times 10^{33}$ erg on 30 April 2007, while Su *et al.* (2009) measure $E_{\rm P} = 0.96 \times 10^{33}$ erg on 2 May 2007, which agrees within $\approx$30%, over a time difference of two days. For the non-potential energy ratio we find $E_{\rm N}/E_{\rm P} = 1.006$ on 30 April 2007, while Su *et al.* (2009) find ≈ 1.1 on 2 May 2007, in the central core of the active region, a few hours before a GOES-class B3.8 and C8.5 flare. Their higher value could thus be attributed to flaring activity. Extensive NLFFF modeling was carried out for 30 April 2007 at 22:24 UT, using *Hinode* vector magnetograph data (DeRosa *et al.*, 2009). The free energies obtained from 11 different NLFFF codes scatter in the range of $E_{\rm N}/E_{\rm P} = 0.87 - 1.24$, with a value of 1.03 for the Wh^{+} code with the smallest misalignment angle of ($\mu = 24^{\circ}$) with respect to the stereoscopically triangulated loops. Including the uncertainties of the boundary conditions in the NLFFF code, a self-consistent NLFFF solution with a non-potential energy ratio of $E_{\rm N}/E_{\rm P} \approx 1.08$ was obtained, with a potential-field energy of $E_{\rm P} = 0.84 \times 10^{33}$ erg (Wheatland and Leka, 2011).

Part of the discrepancy can likely be explained by the different field-of-view, because the *Hinode* field-of-view used in DeRosa *et al.* (2009) only covers the central part of the active region, while our field-of-view encompasses the entire active region. Accordingly, we calculated the magnetic energy in the same *Hinode* field-of-view as used in DeRosa *et al.* (2009), but we found almost identical energy ratios. That our value is lower than that obtained with the other NLFFF codes could also be attributed to an underestimation of the twist (and thus non-potential energy) in the core of the active region, where stereoscopic loop triangulation is very sparse due to confusion of loops with "moss" background. If this is true, we generally expect that the avoidance of twisted core structures leads to a stereoscopic undersampling bias, resulting in lower estimates of the free magnetic energy. We also have to keep in mind that the free-energy ratio for this active region is the lowest among the four active regions and thus has the largest relative uncertainty.

Active region NOAA 10953, 9 May 2007 (B) was subject to the first 3D reconstruction with STEREO (Aschwanden *et al.*, 2008a), stereoscopic electron density and temperature measurements (Aschwanden *et al.*, 2008b), and instant stereoscopic tomography and DEM modeling (Aschwanden *et al.*, 2009). This active region exhibits the simplest bipolar magnetic configuration among all four analyzed active regions and we find a non-potential energy ratio of $E_{\rm N}/E_{\rm P} = 1.023$ (Table 3), *i.e.*, a free-energy ratio of ≈ 2.3 %.

Active region NOAA 10953, 19 May 2007 (C) has exhibited multiple filament eruptions in the complex and highly non-potential magnetic configuration during 19 May 2007 (Li *et al.*, 2008; Liewer *et al.*, 2009; Hara *et al.*, 2011). Some 22 GOES B-class and 2 GOES C-class flares were detected during the observing period (Li *et al.*, 2008). A filament eruption, accompanied by a B9.5 flare, coronal dimming, and an EUV wave was observed and traced

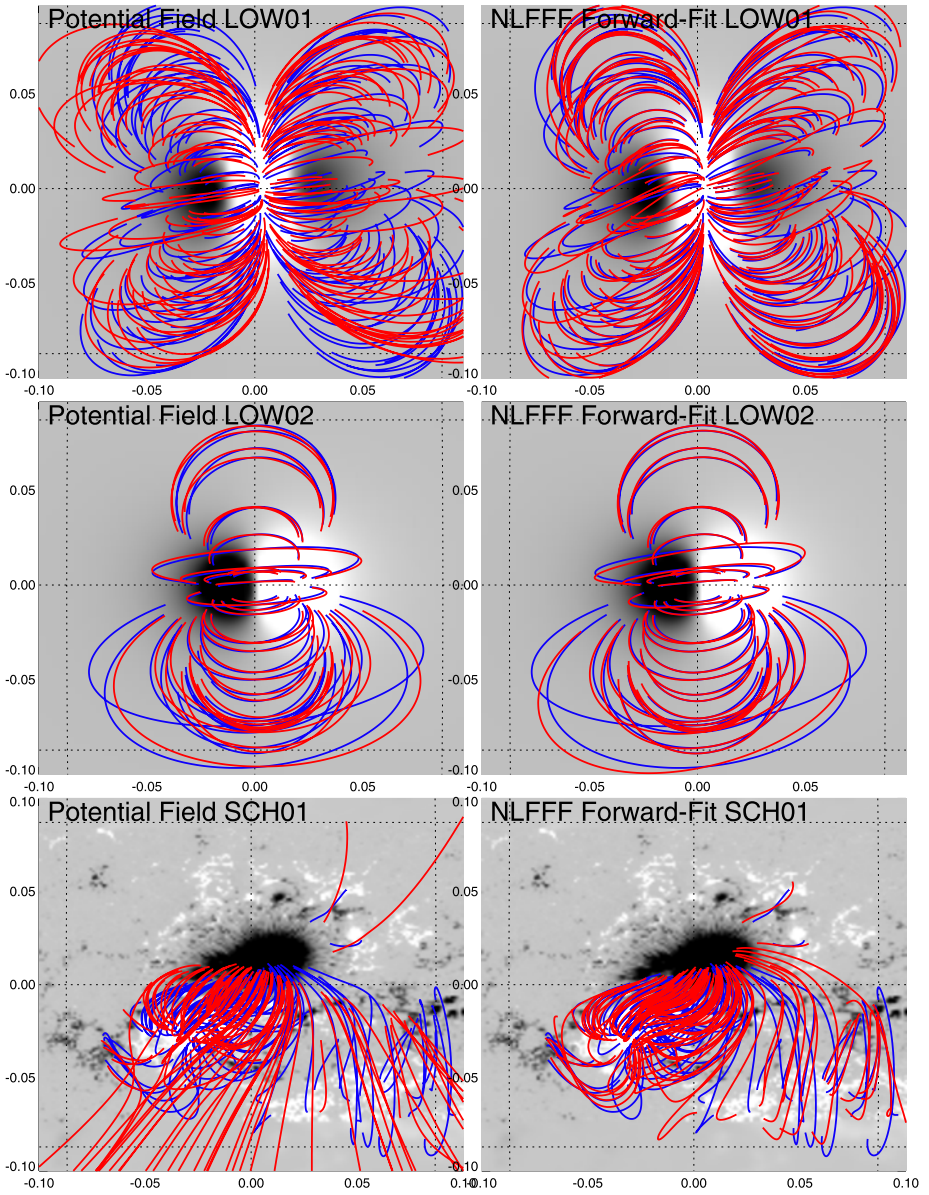


Figure 5 Forward-fitting of L1, L2 (Low and Lou, 1990), and S1 (Schrijver *et al.*, 2008). Representation similar to Figure 2.

with 3D reconstruction after 19 May 2007, 13:00 UT (Liewer *et al.*, 2009). The associated EUV dimming and EUV wave caused by the filament eruption was also analyzed (Attrill *et al.*, 2009). Plasma motion and heating up to $T_e = 9$ MK was observed for the same flare around 13:00 UT (Hara *et al.*, 2011). For this active region, which we analyzed at 12:47 UT shortly before the flare (shown in Figure 6, middle panels), we found the largest amount

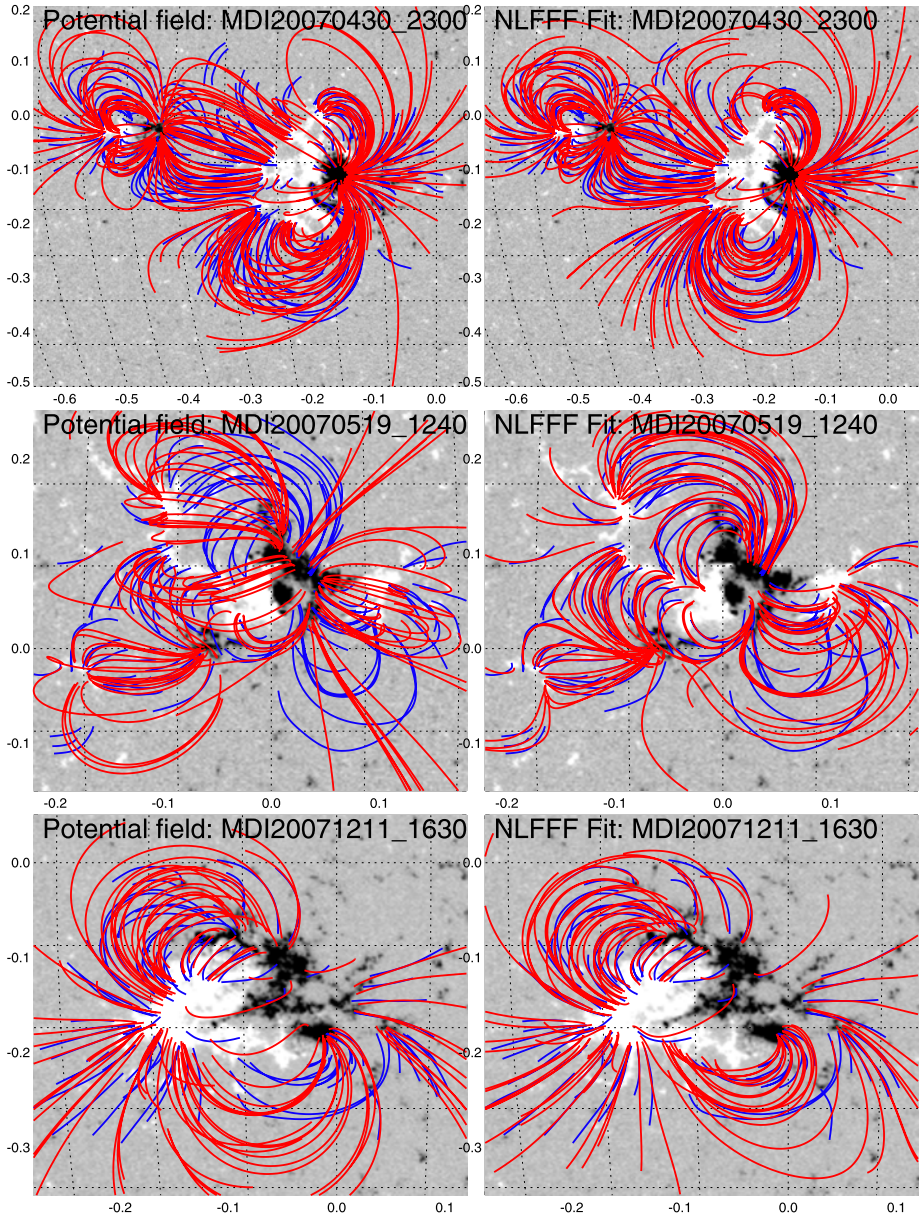


Figure 6 Forward-fitting of active region A (30 April 2007), C (19 May 2007), and D (11 December 2007). Representation similar to Figure 2.

of free energy (9 %), *i.e.*, a non-potential energy ratio of $E_N/E_P = 1.085$ (Table 1), which clearly is associated with the filament eruption and flaring activity after 13:00 UT.

Active region NOAA 10978, 11 December 2007 (D) also appears to have a dominant bipolar structure (shown in Figure 6, bottom panels), but some apparent currents along the central neutral line have been modeled with a flux-insertion method (Alex Engell and van

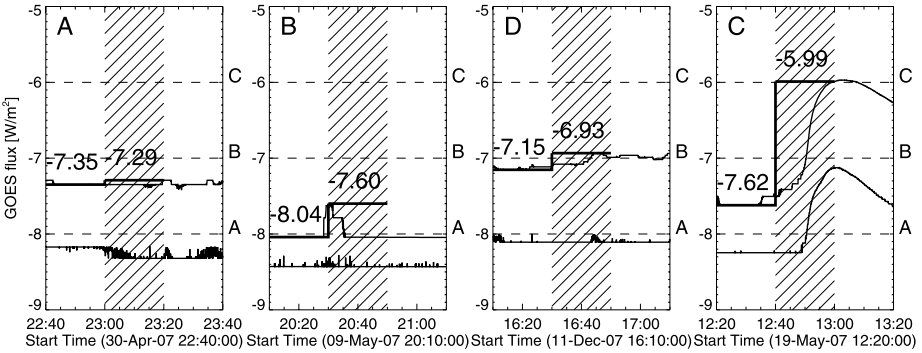


Figure 7 GOES soft X-ray light curves of the 0.5 – 4 Å (upper curve) and 1 – 8 Å channel (lower curve) during the time of stereoscopic triangulation and magnetic modeling of the four analyzed active regions. The flare peak and preflare background levels are indicated with a step function (thick solid lines). The four active regions are in order of increasing GOES flux (adapted from Aschwanden and Sandman, 2010).

Ballegooijen; private communication 2012). For this active region we found a moderate amount of free energy (4 %), *i.e.*, a non-potential energy ratio of $E_N/E_P = 1.044$ (Table 1), which is likely to be stored in the flux rope or filament above the central neutral line.

3.5. Correlation of Free Energies with Flare Fluxes

It was noted previously that the level of flaring activity in these four active regions is positively correlated with the non-potentiality of the magnetic field, evaluated by the average misalignment angle between a magnetic potential-field model and the observed 3D loop coordinates. This correlation was interpreted in terms of a relationship between electric currents and plasma heating (Aschwanden and Sandman, 2010). Since the misalignment angle is a measure of the magnetic non-potentiality, we expect that there should also be a correlation between the flaring activity level and the free magnetic energy in a flaring active region. The GOES 0.5 – 4 and 1 – 8 Å light curves are shown in Figure 7. We plot the (preflare background-subtracted) GOES 0.5 – 4 Å fluxes f_{GOES} of the largest flare that occurred during the observing period of an active region *versus* the free-energy ratio $q_{B,free} = E_{B,free}/E_P$ in Figure 8 (top panel). The preflare background-subtracted GOES fluxes F_{GOES} show a correlation that can be fitted with an exponential function,

$$\left(\frac{f_{GOES}}{f_0}\right) \approx \exp\left(\frac{q_{free}}{q_0}\right) \tag{16}$$

with the constants $q_0 = 0.015$ and $f_0 = 10^{-8.4}$ (W m^{-2}). This result implies that the magnitude of the flare (measured here with the GOES soft X-ray flux) is directly related to the free magnetic energy stored in the active region before the flare. If we include the X3.4 GOES flare of AR 10930 (Figure 8, case S1), we see that the observed free-energy ratio $q_{fit} = E_N/E_P - 1 \approx 0.112$ lies along the same trend of the extrapolated exponential function.

The free energy E_{free}, an absolute measure of the non-potential energy, also shows a correlation with the observed GOES flux, approximately following a power-law (Figure 8, bottom),

$$\left(\frac{f_{GOES}}{f_0}\right) \approx \left(\frac{E_{free}}{E_0}\right)^3 \tag{17}$$

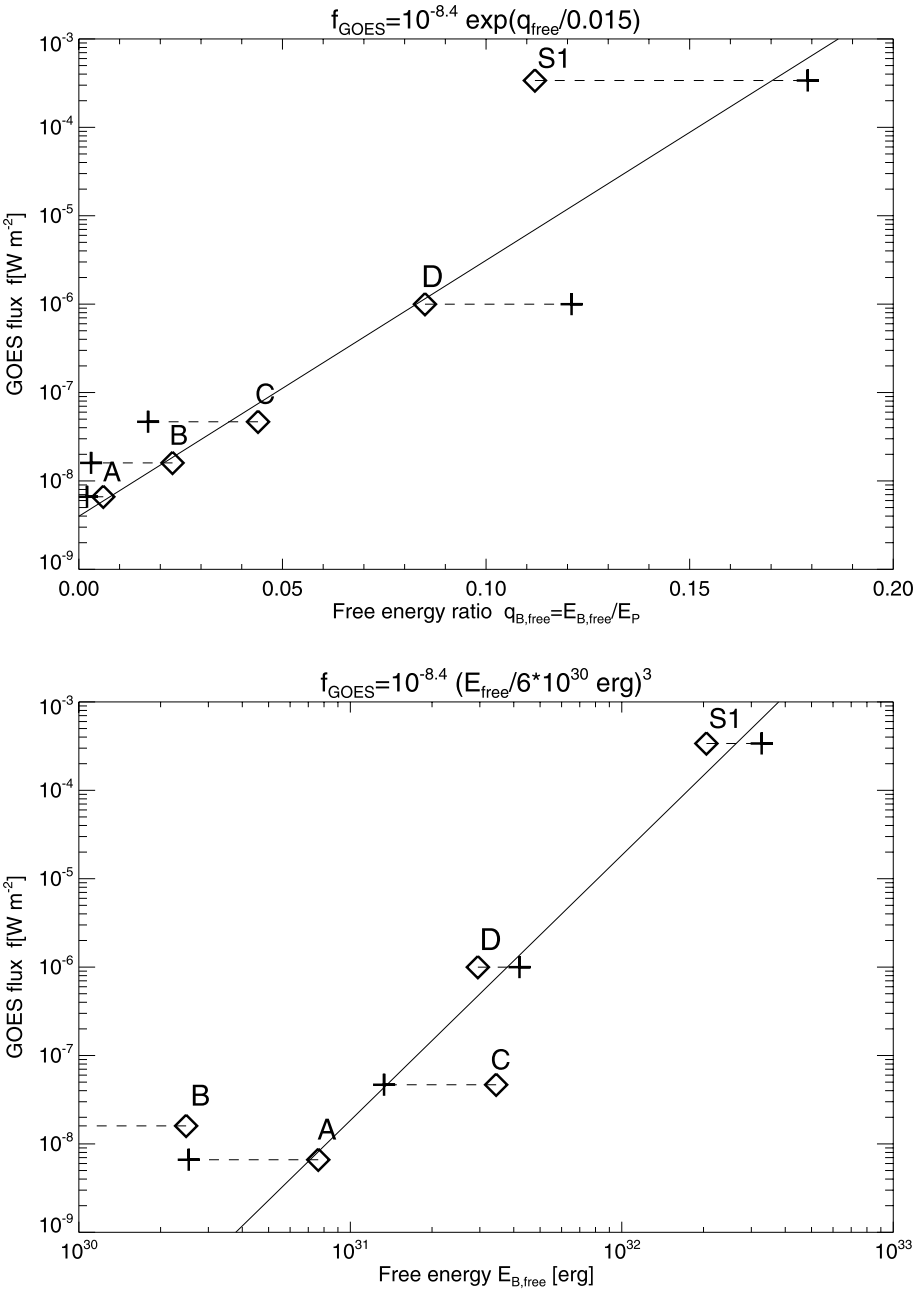


Figure 8 Correlation between the free-energy ratio $q_{\rm free} = E_{\rm free}/E_{\rm P}$ (top panel) and the free energy $E_{\rm free}$ *versus* the logarithmic (preflare background-subtracted) GOES 0.5 – 4 Å flux $f_{\rm GOES}$ for active regions A, B, C, and D, determined with the forward-fit method (diamonds) and with the misalignment method (crosses). Linear regression fits are indicated between linear and logarithmic values, yielding an exponential function (top panel) or a power-law function (bottom panel).

with the constants $f_0 = 10^{-8.4}$ (W m^{-2}) and $E_0 = 6 \times 10^{30}$ erg. Only the active region (B) with the lowest free energy is an outlier to this power-law relationship.

4. Discussion

4.1. Accuracy of the Free Energy

Compared with the other (numerical) NLFFF codes, our analytical NLFFF code has the following advantages: i) computational speed that allows fast forward-fitting to observed coronal data; ii) simplicity of explicit analytical formulation; iii) the spherical geometry of solar surface is fully implemented; and iv) the free energy meets the five criteria of positivity, additivity of energy (or orthogonality of magnetic field components), positive scaling with the force-free parameter α, the potential-field limit for small α's, and finiteness of the non-potential energy. In this study we calculated the free magnetic energy for four different test datasets and aim to validate the accuracy.

The first test dataset consists of simulated loops, which are analytically defined by the same parameterization as the magnetic field lines that are forward-fitted to the simulated data. We find that the total non-potential magnetic energies E_N are retrieved with an accuracy of $\approx 10^{-5}$ for these cases, and the free energy ratios E_N/E_P are retrieved with an accuracy of $\approx 10^{-3}$. The high accuracy just confirms the convergence behavior of our code for a unique solution, as it is the case when the forward-fitting model has the same parameterization as the fitted simulation dataset.

The second test dataset consists of an analytical NLFFF solution, which has a completely different parameterization in terms of Green's function applied to constant-α point sources (Chiu and Hilton, 1977; Low and Lou, 1990; Lothian and Browning, 1995; Malanushenko, Longcope, and McKenzie, 2009; Malanushenko, Yusuf, and Longcope, 2011, 2012). The particular case described in Low and Lou (1990) consists of three smooth elliptical magnetic field concentrations and contains an amount of 13 % free energy, while we recovered only 2 % with the forward-fit. The poor performance is possibly due to the particular topology of the Low and Lou (1990) model, which has extended (elliptical) magnetic field regions parameterized with smooth Legendre polynomials, which cannot be fitted uniquely with spherical point sources. However, because this case is untypical for solar observations, which appear to consist of many small-scale flux concentrations that can be easier fitted with point sources, the accuracy of free energies may be much better for real solar data.

The third data set consists of an observed flaring active region with a large X3.4 flare, for which an NLFFF solution was calculated. For the free energy we obtain a value of 11 % – 18 % before the flare, which approximately agrees with the most reliable NLFFF solution (16 % – 32 %) computed in Schrijver *et al.* (2008) and Malanushenko *et al.* (2012). For the sign of the change in non-potential magnetic energy during this X3.4 flare we also obtain the same sign and a similar value for the decrease of free energy (5 % – 13 %) as the values computed in Schrijver *et al.* (2008) and Malanushenko *et al.* (2012), $\approx$ 13 %.

The fourth test dataset consists of four solar active regions, for which the 3D coordinates of coronal loops could be stereoscopically triangulated. For these four cases we obtained free energies of 0.6 %, 2.3 %, 4.4 %, and 8.5 %. Only for the first case (active region A, 30 April 2007) we have comparisons with other NLFFF codes. The most reliable NLFFF code that exhibits the smallest misalignment angle with stereoscopic loops, yield 3 % free energy, which is close to our result of $E_{free}/E_P = 0.6$ %. The free energy obtained from all 11 NLFFF solutions yields a much larger scatter, *i.e.*, $E_{free}/E_P = 10\ \% \pm 12\ \%$ (Table 1 in DeRosa *et al.*, 2009).

4.2. Scaling Law Between Magnetic Energy and Flare Soft X-Ray Flux

We found a correlation between the free magnetic energy and the GOES flux of the largest flare that occurred during the observing period (Equation (16)). Since our non-potential field solution is parameterized by a vertical twist of magnetic charges, the free energy is directly proportional to the magnetic energy associated with the azimuthal field component B_φ, and thus approximately proportional to the squared force-free α parameter or the number $N_{\rm twist}$ of twists per length unit (Equations (10) and (6)),

$$q_{\rm free}(\mathbf{x}) = \frac{{\rm d}E_{\rm N}(\mathbf{x})}{{\rm d}E_{\rm P}(\mathbf{x})} - 1 = \frac{B_\varphi^2}{B_r^2} = (br\sin\theta)^2 \propto N_{\rm twist}^2 \propto \alpha^2. \tag{18}$$

On the other side, the stress-induced heating rate $E_{\rm H}$ by Ohmic dissipation (or Joule dissipation) is proportional to the square of the currents (*e.g.*, van Ballegooijen, 1986), which is also proportional to the square of α (with $\mathbf{j}/4\pi = \nabla \times \mathbf{B} = \alpha\mathbf{B}$), and thus to the free energy,

$$E_{\rm H} = \frac{j^2}{\sigma} \propto \frac{\alpha^2 B^2}{4\pi\sigma}, \tag{19}$$

where σ is the classical conductivity. Thus we can express Equation (16) in terms of twists, currents, or heating rates,

$$F_{\rm SXR} \propto \exp\left(q_{\rm free}\right) \propto \exp\left(N_{\rm twist}^2\right) \propto \exp\left(\alpha^2\right) \propto \exp\left(j^2/B^2\right) \propto \exp\left(E_{\rm H}/B^2\right). \tag{20}$$

Our empirical finding suggests that the energy radiated from the heated plasma is not just proportional to the heating energy, but "exponentiated" by the current density j, the force-free parameter α, the number of twists $N_{\rm twist}$, or the heating rate $E_{\rm H}$ of the Joule dissipation. This implies a highly nonlinear mechanism that converts vertical twist into thermal radiation via dissipated currents, as envisioned in stress-induced reconnection (Sturrock and Uchida, 1981; Parker, 1983). Our linear regression fit of the free energy *versus* the GOES flux yields a lower limit of $F_{\rm GOES} = 10^{-8.4}$, which corresponds to a sub-A flux GOES class event, about the magnitude of the smallest detectable nanoflare. The largest flare among our analyzed cases is an X3.4 GOES-class flare. Apparently, the exponential relationship between soft X-ray flux and the free energy discovered here approximately holds even for large GOES X-class flares (Figure 8). The nonlinearity is also reflected by the fact that the thermal energy of the soft X-ray flux is not just proportional to the free magnetic energy, but rather exhibits a highly nonlinear power-law function with a power-law index of ≈ 3 (Equation (17)).

5. Conclusions

We calculated the total free magnetic energy contained in a coronal volume that encompasses an active region for four different kinds of datasets: i) simulated data, ii) data created from an analytical NLFFF solution by Low and Lou (1990), iii) a flaring active region, and iv) four active regions observed with STEREO and SOHO/MDI. The free magnetic energy $E_{\rm free} = E_{\rm N} - E_{\rm P}$ is defined by the difference of the non-potential ($E_{\rm N}$) and the potential magnetic field energy ($E_{\rm P}$). The non-potential magnetic field $\mathbf{B}_{\rm N}(\mathbf{x})$ is defined by an analytical approximation of an NLFFF solution that is parameterized by buried magnetic charges with vertical twists (derived in Paper I). The numerical code that performs fast

forward-fitting of magnetic field lines to coronal 3D constraints, such as stereoscopically triangulated loops, is described in Paper II, along with the simulated data. Our findings are as follows:

i) A first method to calculate the free energy results from forward-fitting of our analytical NLFFF approximation by associating the perpendicular magnetic field component $B_\perp$ with the free energy, $dE_{\rm free} = B_\perp^2/8\pi$, while the parallel component $B_\parallel$ is associated with the potential field energy $dE_{\rm P} = B_\parallel^2/8\pi$. This definition of the free energy fulfills the conditions of i) positivity of free energy, ii) additivity of energies, $E_{\rm N} = E_{\rm P} + E_{\rm free}$ and orthogonality $B_\perp(\mathbf{x}) \perp B_{\rm P}(\mathbf{x})$, iii) a positive scaling with the force-free parameter, $dE_{\rm free} \propto \alpha^2$, iv) the potential-field limit, $E_{\rm N}(\alpha \rightarrow 0) = E_{\rm P}$, and v) the finiteness of the nonpotential energy with height h, $E_{\rm N}(h \rightarrow \infty) = E_{\rm max}$.

ii) A second method to estimate the free energy can be obtained from the mean misalignment angle $\Delta\mu = \mu_{\rm P} - \mu_{\rm N}$ between a potential and a nonpotential field (or stereoscopically triangulated coronal loops). The free energy ratio is then $E_{\rm free}/E_{\rm P} \approx \tan^2(\Delta\mu)$. We find that the uncertainty of this method amounts to $\approx \pm 2$ % for the nonpotential magnetic energy.

iii) Calculating the free energies for the simulated data we find a high fidelity of order 10^{-5} in retrieving the free energy, which is due to the fact that the simulated data have the same parameterization as the forward-fitting method, constraining a single best-fit solution.

iv) Calculating the free energy for the Low and Lou (1990) analytical case, our NLFFF code finds a significantly lower value than theoretically calculated, probably because of the special morphopology (parameterized with smooth Legendre functions), which cannot adequately be fitted with our NLFFF code that is designed for spherical magnetic point sources, as found in solar magnetograms.

v) Calculating the free energy for observed active regions constrained by the 3D coordinates of stereoscopically triangulated coronal loops, we find free energy ratios of $q_{\rm free} = E_{\rm free}/E_{\rm P} \approx 1\,\% - 10\,\%$. The uncertainty of the free energy determined with our forward-fitted NLFFF approximation appears to be at least as good as the uncertainty among other (standard NLFFF extrapolation) codes.

vi) We find also a correlation between the free magnetic energy $E_{\rm free}$ and the GOES flux of the largest flare that occurred during the observing period, which can be quantified by an exponential relationship, $F_{\rm GOES} \propto \exp(q_{\rm free})$, implying an exponentiation of the dissipated currents.

In summary, this study demonstrates that the free energy in active regions can be calculated and predicted with our analytical NLFFF approximation with an accuracy that is commensurable with other standard NLFFF codes. Our code has the additional advantages of computational speed for forward-fitting of coronal data, correct treatment of the curved solar surface, positivity, and finiteness of free energy. In addition, forward-fitting of our NLFFF approximation achieves a significantly smaller misalignment angle with respect to the observed coronal loops ($\mu \approx 2^\circ - 22^\circ$), compared with the results of other NLFFF codes ($\mu = 20^\circ - 44^\circ$; DeRosa *et al.*, 2009). The most limiting drawback of our method is the availability of stereoscopic data with suitable spacecraft separation angle (which was most favorable in 2007, the first year of the STEREO mission). In future work we attempt to circumvent the 3D geometry of coronal loops by using only the 2D projections of coronal loops, which can (manually or automatically) be traced from loop-rich EUV or soft X-ray images and do not require stereoscopic data at all.

Acknowledgements The author appreciates the provided data and helpful discussions with Anna Malanushenko and Marc L. DeRosa. Part of the work was supported by NASA contract NNG 04EA00C of the SDO/AIA instrument and the NASA STEREO mission under NRL contract N00173-02-C-2035.

References

Aschwanden, M.J.: 2004, *Physics of the Solar Corona. An Introduction*, Praxis Publishing Co., Chichester, and Springer, Berlin, Section 5.3.
Aschwanden, M.J.: 2012, *Solar Phys.* doi:10.1007/s11207-012-0069-7 (Paper I).
Aschwanden, M.J., Malanushenko, A.: 2012, *Solar Phys.* doi:10.1007/s11207-012-0070-1 (Paper II).
Aschwanden, M.J., Sandman, A.W.: 2010, *Astron. J.* **140**, 723.
Aschwanden, M.J., Wülser, J.P., Nitta, N.V., Lemen, J.R.: 2008a, *Astrophys. J.* **679**, 827.
Aschwanden, M.J., Nitta, N.V., Wülser, J.P., Lemen, J.R.: 2008b, *Astrophys. J.* **680**, 1477.
Aschwanden, M.J., Wülser, J.P., Nitta, N.V., Lemen, J.R., Sandman, A.W.: 2009, *Astrophys. J.* **695**, 12.
Aschwanden, M.J., Wülser, J.P., Nitta, N.V., Lemen, J.R., DeRosa, M.L., Malanushenko, A.: 2012, *Astrophys. J.* **756**, 124.
Attrill, G.D.R., Engell, A.J., Wills-Davey, M.J., Grigis, P., Testa, P.: 2009, *Astrophys. J.* **704**, 1296.
Chiu, Y.T., Hilton, H.H.: 1977, *Astrophys. J.* **212**, 873.
Choe, G.S., Cheng, C.Z.: 2002, *Astrophys. J. Lett.* **574**, L179.
DeRosa, M.L., Schrijver, C.J., Barnes, G., Leka, K.D., Lites, B.W., Aschwanden, M.J., *et al.*: 2009, *Astrophys. J.* **696**, 1780.
DeVore, C.R., Antiochos, S.K.: 2005, *Astrophys. J.* **628**, 1031.
Falconer, D., Moore, R., Gary, G.A.: 2006, *Astrophys. J.* **644**, 1258.
Falconer, D., Barghouty, A.F., Khazanov, I., Moore, R.: 2011, *Space Weather* **9**, S04003.
Fang, F., Manchester, W., Abbett, W.P., van der Holst, B.: 2012, *Astrophys. J.* **754**, 15.
Hara, H., Watanabe, T., Harra, L.K., Culhane, J.L., Young, P.R.: 2011, *Astrophys. J.* **741**, 107.
Jing, J., Chen, P.F., Wiegelmann, T., Xu, Y., Park, S.H., Wang, H.: 2009, *Astrophys. J.* **696**, 84.
Jing, J., Tan, C., Yuan, Y., Wang, B., Wiegelmann, T., Xu, Y., Wang, H.: 2010, *Astrophys. J.* **713**, 440.
Kusano, K., Maeshiro, T., Yokoyama, T., Sakurai, T.: 2002, *Astrophys. J.* **577**, 501.
Li, Y., Lynch, B.J., Stenborg, G., Luhmann, J.G., Huttunen, K.E.J., Welsch, B.T., Liewer, P.C., Vourlidas, V.: 2008, *Astrophys. J. Lett.* **681**, L37.
Liewer, P.C., De Jong, E.M., Hall, J.R., Howard, R.A., Thompson, W.T., Culhane, J.L., Bone, L., van Driel-Gesztelyi, L.: 2009, *Solar Phys.* **256**, 57.
Lothian, R.M., Browning, P.K.: 1995, *Solar Phys.* **161**, 289.
Low, B.C., Lou, Y.Q.: 1990, *Astrophys. J.* **408**, 689.
Malanushenko, A., Longcope, D.W., McKenzie, D.E.: 2009, *Astrophys. J.* **707**, 1044.
Malanushenko, A., Yusuf, M.H., Longcope, D.W.: 2011, *Astrophys. J.* **736**, 97.
Malanushenko, A., Schrijver, C.J., DeRosa, M.L., Wheatland, M.S., Gilchrist, S.A.: 2012, *Astrophys. J.* **756**, 153.
Metcalf, T.R., Leka, K.D., Mickey, D.L.: 2005, *Astrophys. J. Lett.* **623**, L53.
Metcalf, T.R., Jiao, L., Uitenbroek, H., McClymont, A.N., Canfield, R.C.: 1995, *Astrophys. J.* **439**, 474.
Parker, E.N.: 1983, *Astrophys. J.* **264**, 642.
Régnier, S.: 2009, *Astron. Astrophys.* **497**, 17.
Régnier, S.: 2012, *Solar Phys.* **277**, 131.
Régnier, S., Priest, E.R.: 2007, *Astron. Astrophys.* **669**, L53.
Sandman, A.W., Aschwanden, M.J.: 2011, *Solar Phys.* **270**, 503.
Sandman, A.W., Aschwanden, M.J., DeRosa, M.L., Wülser, J.P., Alexander, D.: 2009, *Solar Phys.* **259**, 1.
Schrijver, C.J., DeRosa, M., Metcalf, T.R., Liu, Y., McTiernan, J., Regnier, S., Valori, G., Wheatland, M.S., Wiegelmann, T.: 2006, *Solar Phys.* **235**, 161.
Schrijver, C.J., DeRosa, M., Metcalf, T.R., Barnes, G., Lites, B., Tarbell, T., *et al.*: 2008, *Astrophys. J.* **675**, 1637.
Sturrock, P.A., Uchida, Y.: 1981, *Astrophys. J.* **246**, 331.
Su, Y., Van Ballegooijen, A., Lites, B.W., DeLuca, E.E., Golub, L., Grigis, P.C., Huang, G., Ji, H.S.: 2009, *Astrophys. J.* **691**, 105.
Van Ballegooijen, A.A.: 1986, *Astrophys. J.* **311**, 1001.
Wang, H.: 1997, *Solar Phys.* **174**, 163.
Wang, H.: 2006, *Astrophys. J.* **649**, 490.
Wang, H., Liu, C.: 2010, *Astrophys. J. Lett.* **716**, L195.
Wang, H., Ewell, M.W. Jr., Zirin, H.: 1994, *Astrophys. J.* **424**, 436.

Wang, S., Liu, C., Wang, H.: 2013, *Astrophys. J. Lett.* **757**, L5.
Wang, H., Spirock, T.J., Qiu, J., Ji, H., Yurchyshyn, V., Moon, Y.J., Denker, C., Goode, P.R.: 2002, *Astrophys. J.* **576**, 497.
Wang, H., Qiu, J., Jing, J., Spirock, T.J., Yurchyshyn, V., Abramenko, V., Ji, H., Goode, P.R.: 2004, *Astrophys. J.* **605**, 931.
Wheatland, M.S., Leka, K.D.: 2011, *Astrophys. J.* **728**, 112.
Wiegelmann, T.: 2004, *Solar Phys.* **219**, 87.
Wiegelmann, T., Sakurai, T.: 2012, *Living Rev. Solar Phys.* **9**(5). http://solarphysics.livingreviews.org/Articles/lrsp-2012-5.
Woltjer, L.: 1958, *Proc. Natl. Acad. Sci.* **44**, 489.

DOI 10.1007/978-1-4899-8005-2_24
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0194-3

Origins of Rolling, Twisting, and Non-radial Propagation of Eruptive Solar Events

Olga Panasenco · Sara F. Martin · Marco Velli · Angelos Vourlidas

Received: 29 February 2012 / Accepted: 5 November 2012 / Published online: 7 December 2012

Abstract We demonstrate that major asymmetries in erupting filaments and CMEs, namely major twists and non-radial motions are typically related to the larger-scale ambient environment around eruptive events. Our analysis of prominence eruptions observed by the STEREO, SDO, and SOHO spacecraft shows that prominence spines retain, during the initial phases, the thin ribbon-like topology they had prior to the eruption. This topology allows bending, rolling, and twisting during the early phase of the eruption, but not before. The combined ascent and initial bending of the filament ribbon is non-radial in the same general direction as for the enveloping CME. However, the non-radial motion of the filament is greater than that of the CME. In considering the global magnetic environment around CMEs, as approximated by the Potential Field Source Surface (PFSS) model, we find that the non-radial propagation of both erupting filaments and associated CMEs is correlated with the presence of nearby coronal holes, which deflect the erupting plasma and embedded fields. In addition, CME and filament motions, respectively, are guided towards weaker field regions, namely null points existing at different heights in the overlying configuration. Due to the presence of the coronal hole, the large-scale forces acting on the CME may be asymmetric. We find that the CME propagates usually non-radially in the direction of least resistance, which is always away from the coronal hole. We demonstrate these results using both low- and high-latitude examples.

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

Electronic supplementary material The online version of this article (doi:10.1007/s11207-012-0194-3) contains supplementary material, which is available to authorized users.

O. Panasenco (✉) · S.F. Martin
Helio Research, La Crescenta, CA 91214, USA
e-mail: panasenco.olga@gmail.com

M. Velli
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

A. Vourlidas
Space Sciences Division, Naval Research Laboratory, Washington, DC, USA

Keywords Coronal mass ejections, low coronal signatures · Coronal mass ejections, initiation and propagation · Magnetic fields, corona · Coronal holes, prominences, formation and evolution · Filaments

1. Introduction

Eruptive solar events, which include a coronal mass ejection (CME), an erupting filament with surrounding cavity and a flare, are distinguished from *simple confined* flares which lack most of these eruption signatures. The term CME is increasingly used in reference to the whole eruptive sequence and not just the ejection of coronal mass. Regardless of definition, there has been widespread recognition, for over three decades, of the close association among CMEs, erupting filaments, coronal cavities and flares.

What do CMEs and erupting filaments have in common? All of these events originate above and around regions where the magnetic field changes polarity, seen in photospheric magnetograms and variously named a neutral line or polarity inversion line or a polarity reversal boundary. We know that erupting filaments, in particular, originate from specific polarity reversal boundaries that are also *inside filament channels* (Martin, 1990; Gaizauskas, 1998). We also already know they only come from filament channels that have reached maximum development; *i.e*. have reached maximum magnetic shear along the polarity boundary (Martin, 1998). Therefore, not every polarity reversal boundary is a *filament channel*.

A filament channel is a volume of space around (and encompassing) a polarity reversal boundary with maximum magnetic shear. A filament is not necessarily present but the channel is a necessary condition for the formation of a filament. The formation of filament channels in the chromosphere has been described by Smith (1968), Foukal (1971), Martin (1990), Gaizauskas *et al.* (1997), Wang and Muglach (2007), Martin, Lin, and Engvold (2008) and Martin *et al.* (2008). In and around active regions or decaying active regions observed in Hα a filament channel may be recognized by the presence of fibrils aligned along a polarity reversal boundary and the local magnetic field. Because no fibrils cross the polarity reversal boundary in a fully developed filament channel and fibrils are field-aligned, one may conclude that no magnetic field lines from active region or network magnetic fields cross this boundary in the chromosphere, above or within filaments. This is also true for the low corona. The initial lack of coronal connectivity across polarity reversal boundary inside the filament channel volume, under, across, and along the filament is confirmed by Wood and Martens (2003). We find that the coronal cells observed and modeled by Sheeley and Warren (2012), do not cross the polarity reversal boundary within a filament channel at the heights below the filament spine top (Figure 1). Coronal cells originate from the network field concentrations and show the same patterns as chromospheric fibrils because they follow the same filament channel magnetic topology, with a (presumably) strong horizontal component of the field.

As deduced from chromospheric fibrils and coronal cells, the direction of the network magnetic fields associated with chromospheric fibrils and coronal cells changes from upward on one side of a polarity reversal boundary to horizontal along the polarity reversal boundary and to vertically downward on the other side of the polarity boundary. A vital precondition for the formation of a filament channel is a magnetic field with a strong horizontal component (Gaizauskas, 1998) and the chromospheric fibrils in a filament channel trace such a horizontal pattern along its polarity reversal boundary. So far no observational evidence of the existence of an ascending twisted flux rope at the chromospheric level within

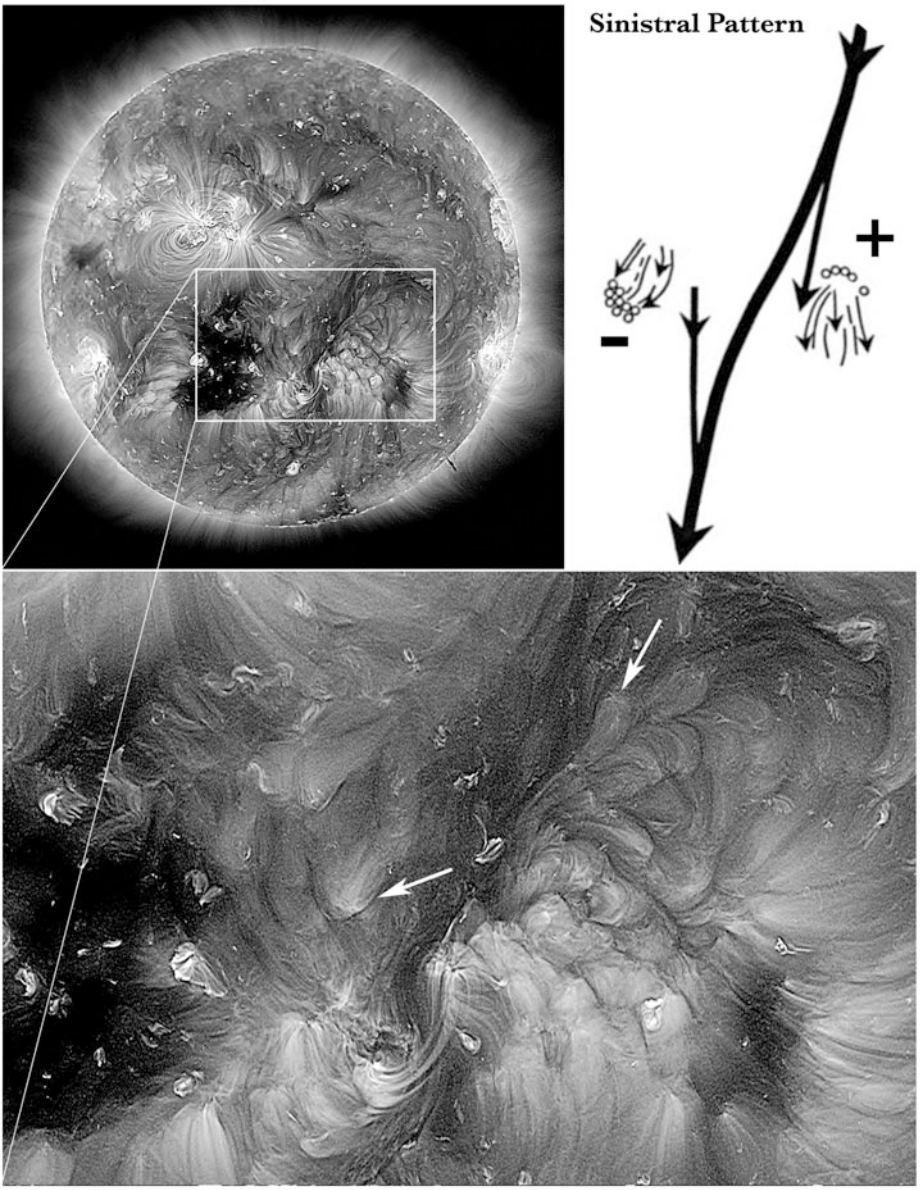


Figure 1 Shown here are coronal cells on both sides of the filament channel (two representative cells are indicated by the white arrows), note that the cells on opposite sides of the channel have cusp-shaped tops that are in opposite directions: observe that they do not cross the channel but follow the same pattern as chromospheric fibrils within a sinistral filament channel (a left-handed filament channel) as depicted in the schematic representation in the upper right where a sinistral filament and adjacent chromospheric fibrils with a sinistral pattern are shown (Martin, Lin, and Engvold, 2008). Coronal cells form at heights $\sim 6-10$ Mm (Sheeley and Warren, 2012), much lower than average filament heights $\sim 50-70$ Mm. The SDO/AIA 193 Å image is from 19 June 2011 05:39 UT.

a filament channel above a polarity reversal boundary has been found; this would require fibrils to systematically change their orientation and cross the polarity reversal boundary. Conversely, were the reverse true, emerging, twisted flux tubes playing a major role in the physics of filaments and filament channels, one would expect there to be ubiquitous evidence of polarity crossing fibrils, which would necessarily appear earlier, if there was a flux rope axis submerged in the photosphere (Wood and Martens, 2003).

This is a fundamental observation which allows one to deduce at least one significant moment in the formation of CMEs: the most likely time of initiation of twist in CMEs. There is evidence that once an eruption is under way, a twisted flux rope does indeed form during magnetic reconnection of the stretched, sheared overlying coronal loop system and concurrently with the initiation of flare ribbons in the chromosphere (Wang, Muglach, and Kliem, 2009). In other words, the observational evidence suggests that the formation of the flux rope of CMEs occurs after the filament and coronal loop system are already ascending.

Similar to the overlying coronal loops within a coronal loop system, before eruption, filaments rise slowly. Unlike coronal loops and if a filament erupts non-radially, as frequently happens, the top of its spine first bends to one side and evolves into a sideways rolling motion, known as the roll effect. As shown by 304 Å observations from SOHO and STEREO and earlier Hα Doppler observations, the rolling motion unambiguously propagates down the legs of erupting filaments resulting in the large-scale twists which are commonly observed in them. The rolling of the spine initiates structural twist with opposite chirality in the two legs (Figure 2). Concurrently, rotational mass motions with the opposite signs of chirality (and helicity) in the two legs are observed to flow downward while the body of the filament ascends. However, in asymmetric eruptions, both legs are not always observed (Martin, 2003; Panasenco and Martin, 2008).

In addition to the observed absence of twist in the pre-eruptive state, further evidence that the energy creating the twist comes from above was found in Doppler shifts (Martin, 2003); the rotational motions in the legs of erupting filaments are not only opposite in sign to each other but the twists in both legs are opposite in sign to that required if the observed

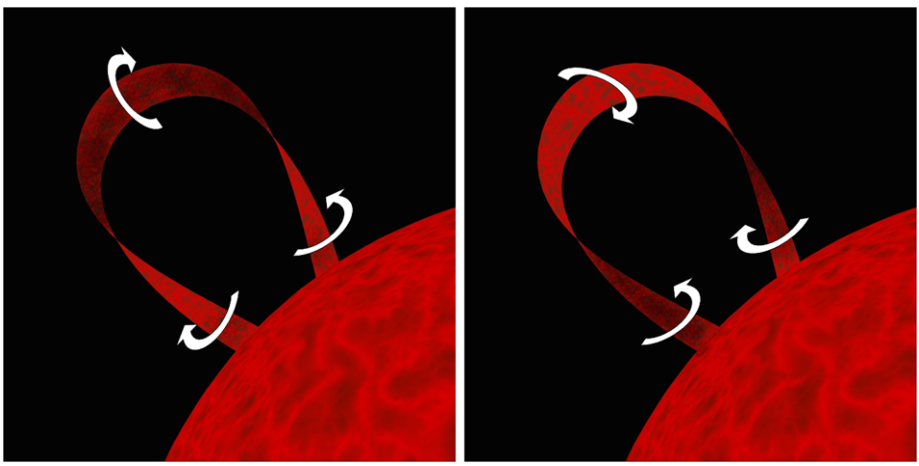


Figure 2 Schematic of the two forms of the roll effect in erupting prominences. The roll effect initiates at the top of a filament and propagates downward into the legs. The legs respond by twisting. Rotational motions of the legs depend on the roll direction. There is no photospheric motion fast enough to drive the observed rotation of the legs.

sense of twist were generated at the feet or in the legs of the erupting filament (Martin, 2003; Panasenco and Martin, 2008).

Here we will examine events showing that the combined ascent and initial bending of the filament is non-radial in the same general direction as for the surrounding CME. Also, the non-radial motion of the filament is *greater* than that of the CME. By considering the global magnetic environment around CMEs, as can be approximated by the Potential Field Source Surface (PFSS) model, we will see that both erupting filaments and their surrounding CMEs are non-radial moving away from nearby coronal holes and toward null points. Due to the presence of the coronal hole, both the local forces on an erupting filament and the global forces on the CME are asymmetric. That the CME propagates non-radially always away from coronal holes was demonstrated in many studies (e.g., Plunkett *et al.*, 2001; Gopalswamy *et al.*, 2003, 2009; Cremades, Bothmer, and Tripathi, 2006; Cremades and Bothmer, 2004; Kilpua *et al.*, 2009; Zuccarello *et al.*, 2012). Prominence deflection and rolling motion during eruptions received less attention, though there has been some work in the same general direction as advocated here (Filippov, Gopalswamy, and Lozhechkin, 2001; Martin, 2003; Panasenco and Martin, 2008; Bemporad, 2009; Panasenco *et al.*, 2011; Pevtsov, Panasenco, and Martin, 2012; Liewer, Panasenco, and Hall, 2012). The formation of the CME flux rope appears to occur in the early phase of filament eruption, when the rolling motion of the filament is already in progress (Liewer, Panasenco, and Hall, 2012).

The roll effect appears very early during the beginning phase of prominence eruptions. We will show that it is caused by an asymmetry in magnetic structure and magnetic flux density on the two sides of the filament. The non-radial direction of the eruption of the whole filament system is caused directly either by the open coronal hole magnetic field near the filament channel (cases studied in this paper), or by other strong magnetic field configurations which might be in the neighborhood of the eruption (strong in reference to the overall energy of the eruption).

The paper is organized as follows: we begin with a general introduction to the coronal structures which lead to eruptive solar events. We then discuss a series of recent eruptions within such structures, which may be referred to generically as either streamers or pseudostreamers, exhibiting non-radial expansion and propagation. We then correlate such behavior with the properties of the ambient corona. We also discuss the three-dimensional structure of such events, focusing on one particular case where the two well separated STEREO viewpoints show that what appears to be a highly kinked, flux rope-like erupting prominence and cavity is a projection effect.

2. Magnetic Topologies Overlying Filament Channels

Filament channels and filaments develop along and above polarity reversal boundaries in the photosphere and lie under coronal loop systems also known as coronal arcades. In the instances most commonly studied, these arcades correspond to a helmet-streamer type structures in the outer corona. The outer, weaker fields do not always close over the arcade but may remain open, oppositely directed adjacent fields (what we will refer to as a globally dipolar structure). Recently, filament channels and filaments lying below arcades and belonging to more complex magnetic structures have been shown to be very important for multiple and global solar eruptive events, such as that of 01 August 2010 (Schrijver and Title, 2011; Török *et al.*, 2011). The prominences in that event had a magnetic structure corresponding to a hierarchy of higher order multipolar structures resulting in a pseudostreamer (Wang, Sheeley, and Rich, 2007), which was found to contain *twin* filaments at its base

(Panasenco and Velli, 2010). Such twin filaments are topologically connected, sharing a neutral point and a separatrix dome. This was a case in which two polarity reversal boundaries contain between them fields with a polarity opposite to that of the global unipolar configuration surrounding them (*tripolar* pseudostreamer).

The magnetic field, based on the potential field source surface (PFSS) reconstruction, shows that field lines of the arcades above filaments are often asymmetrically placed with respect to the polarity reversal boundaries, because the overarching magnetic field is stronger at one side. In the PFSS this is shown by the greater density of field lines on one side of the arcade with respect to the other. As we shall see, this is a tell-tale sign for strong rolling and nonradiality of the eruption. The question naturally arises as to the fidelity of the PFSS model in the neighborhood of filaments and filament channels, where the strong shears naturally lead to the formations of currents in the corona, currents which are also required to get to the instability involved in the filament eruption and CME formation. In reality, the PFSS does a good job because at least a component of the axial field along the filament spine is potential in nature. This may be understood as having to do with the finite longitudinal extent of the channel so that the polarities on opposite side of the inversion line are not translationally invariant along it. Indeed, this component is recovered also by photospheric shear and flux transport models which correctly describe filament channel formation statistically based on photospheric transport (Yeates, Mackay, and van Ballegooijen, 2007). In the following sections we will make use of the PFSS model of Schrijver and De Rosa (2003) to study the coronal magnetic field and its role in determining the trajectory of filaments and CMEs for a whole series of events including even higher order multipoles below streamers and pseudostreamers.

3. Non-radial Eruptions with the Roll Effect

We consider here a set of prominence and CME eruptions which display both the roll effect and significant non-radiality. The different examples paint a picture of the rich variability in motions which the prominence and CME may display, depending on the detailed magnetic field configuration surrounding the eruption and the global coronal context. We use images from the *Sun–Earth Connection Coronal and Heliospheric Investigation* (SECCHI: Howard *et al.*, 2008) aboard the *Solar-Terrestrial Relations Observatory* (STEREO: Kaiser *et al.*, 2008) and from the *Atmospheric Imaging Assembly* (AIA: Lemen *et al.*, 2012) aboard the *Solar Dynamics Observatory* (SDO: Pesnell, Thompson, and Chamberlin, 2012). We also use information from the *Large Angle Spectrometric Coronagraph* (LASCO: Brueckner *et al.*, 1995) on the *Solar and Heliospheric Observatory* (SOHO: Domingo, Fleck, and Poland, 1995).

3.1. 2 November 2008 Filament Eruption

This strongly non-radial eruption (Figure 3) was described in detail by Kilpua *et al.* (2009). The right panel in Figure 3 shows that the filament channel was located under the northernmost arcade of a double arcade system underneath a pseudostreamer, a large-scale coronal structure which differs from the classical helmet streamer because the magnetic field is unipolar above the dome-like closed field regions. The pseudostreamer itself, adjacent to a strong northern coronal hole, expanded outward with considerable equatorward deflection. At the early stage of the eruption, the motion of the filament was directed towards the local null point, O, below the pseudostreamer separatrix (Figure 3).

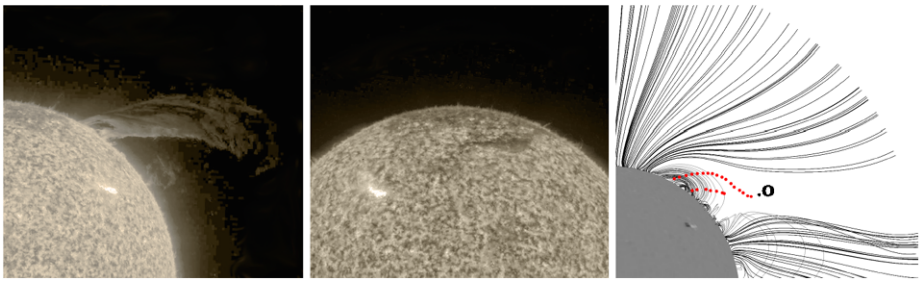


Figure 3 Non-radial eruption with the roll on 2 November 2008. The large curvature results from sideways rolling of the top of the prominence. Left and middle are STEREO-B and STEREO-A/EUVI 304 Å images of the prominence during the eruption at 01:57:04 and 01:56:15 UT, respectively. The right image is a PFSS extrapolation of the open magnetic field lines (pseudostreamer) and the coronal loops overlying the filament before the eruption on 1 November 2008 12:04 UT (rotated to the limb view); red dotted lines represent the erupting filament directed toward the null point, O.

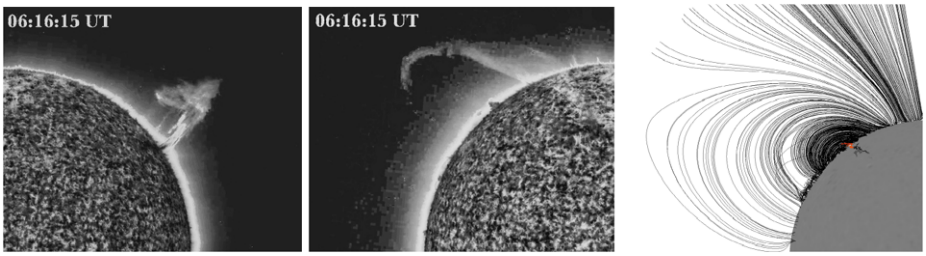


Figure 4 Non-radial eruption on 12 December 2008. Left and middle are STEREO-B and STEREO-A/EUVI 304 Å images of the prominence during the eruption. The right image is a PFSS extrapolation of the open solar magnetic field and the coronal loops overlying the filament (STEREO-A/EUVI 304 Å) before its eruption. The position of the filament under the arcade is very asymmetric.

3.2. 12 December 2008 Filament Eruption

This event, illustrated in Figure 4, is described in detail in the paper by Panasenco *et al.* (2011). It was observed by both STEREO-A, at the north east limb (middle panel), and STEREO-B, at the north-west limb (left panel). Though the spacecraft were separated at the time by 86.7°, both imagers show the rolling motion of the erupting prominence. The erupting prominence and CME were deviated from radial trajectories by different degrees. To illustrate this, in Figure 5 we superimpose outlines of the outer boundaries of the CME and the bright core of the prominence plasma for different consecutive moments during the eruption (as observed by the *Extreme Ultraviolet Imager* (EUVI-A) 304 Å and COR1-A, respectively). The deviation of the CME from radial propagation is about 40°, for the prominence about 60° and the difference between the prominence and CME central direction of the eruption is about 20° for this STEREO-A view. Since the prominence erupts approximately toward STEREO-B, and the separation angle between A and B is close to 90°, we may safely assume that the projection on the sky in the STEREO-A view shows a separation of trajectories which is smaller than the real separation between the propagation directions for prominence and CME, shown in Figure 5.

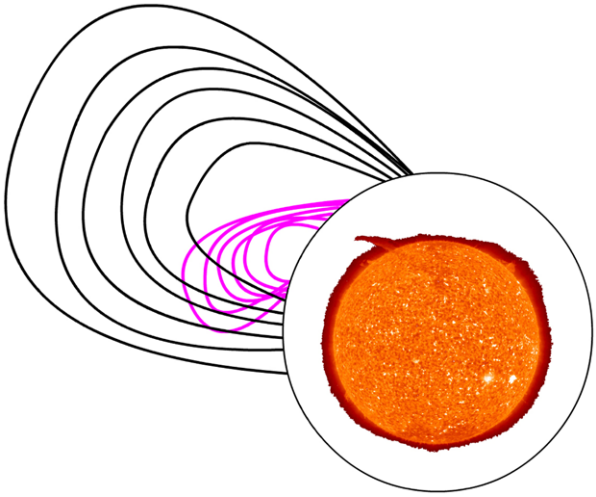


Figure 5 Non-radial eruption on 12 December 2008. Superposition of the outer boundaries of CME (black lines) and erupting prominence inside (magenta lines) as observed by STEREO-A/COR1 during different moments of the eruption and the solar disk with the prominence before eruption (STEREO-A/EUVI 304 Å).

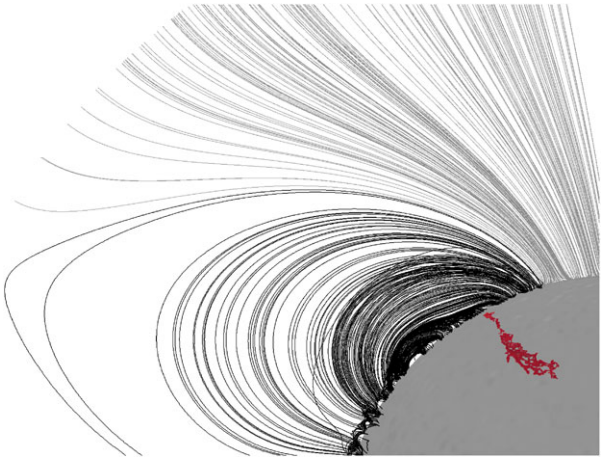


Figure 6 Magnetic configuration above the filament before its eruption on 26 September 2009 (STEREO-B). The asymmetry in the position of the filament under the arcade is very strong.

3.3. 26 September 2009 Filament Eruption

This eruption was observed by both STEREO, at the north east limb by A, and across the disk by B as shown in Figure 7. The separation angle between the spacecraft was 117.4°. EUVI-B 304 Å images show the prominence crossing the disk during the eruption. STEREO-A observed this eruption above the limb. The position of the filament relative to the coronal arcade and open field of the coronal hole before the eruption is shown in Figure 6. The asymmetric location of the filament is very pronounced. Figure 7 shows the eruption on 26 September 2009 at 22:45 UT as observed simultaneously by STEREO-A/COR1 and EUVI instruments. The non-radial motion of the erupting filament is much greater than that of the CME. The CME envelope expands, containing the prominence within it, but the two bodies are not rigidly linked to one another. Plasma of the prominence moves approximately along the line which can be described as a tangent line to the limb from where the prominence erupted. Apparently the imbalance in magnetic pressure under the coronal arcade and around the filament during the early stage of the eruption was such that the stronger field on

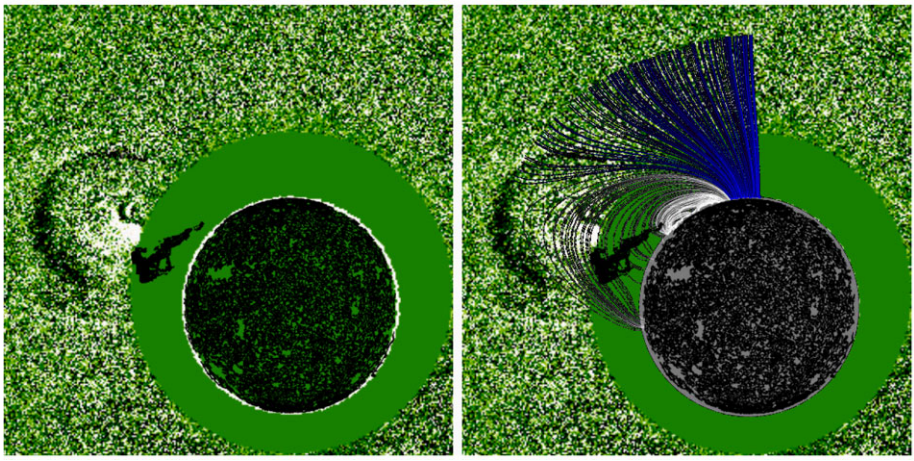


Figure 7 Non-radial eruption on 26 September 2009 at 22:45 UT. Left: A superposition of the COR1-A and EUVI-A 304 Å images. Right: The same with PFSS reconstruction added. The non-radial motion of the filament is much greater than that of the CME.

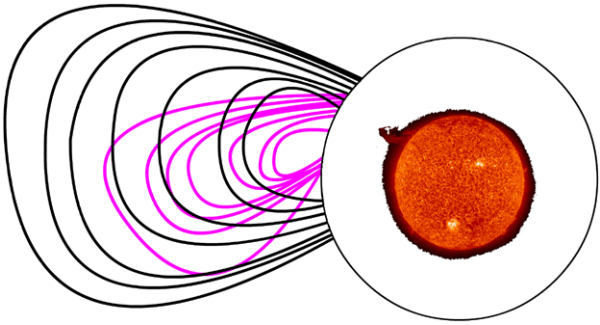


Figure 8 Non-radial eruption on 26 September 2009. Superposition of the outer boundaries of the CME (black lines) and erupting prominence inside (magenta lines) as observed by SOHO/LASCO C2 during different moments of the eruption (26 – 27 September) and the solar disk with the prominence before eruption (SOHO/EIT 304 Å).

the coronal hole side of the arcade caused such an extensive lateral motion of the filament. The corresponding CME also deviates from radial propagation. Analyzing the projection on the sky as viewed by STEREO-A we found that the difference between the prominence and CME deviations amounted to about 30°. Figure 8 shows an intermediate projection of the eruption on the plane of the sky as observed by the SOHO/EIT and LASCO C2 instruments. This projection viewpoint is midway between those of STEREO-A and B, therefore the apparent difference in the deviation angle for the prominence and CME is smaller, about 20°.

3.4. 30 April 2010 Filament Eruption

This non-radial filament eruption with the roll effect was observed against the solar disk by the Mauna Loa Solar Observatory (MLSO) on 30 April 2010. The movie of the first 40 min of the eruption (23:25 – 00:05 UT) in the He I spectral line shows that the filament crossed the disk and the neighboring active region in the south-eastern direction. This direction of the eruption was away from the coronal hole north-west of the filament channel (movie available in the supplementary material). The distance which the filament covered over these 40 min was about 160 000 km, corresponding to a projected speed of ~ 65 km s^{-1}. The first four

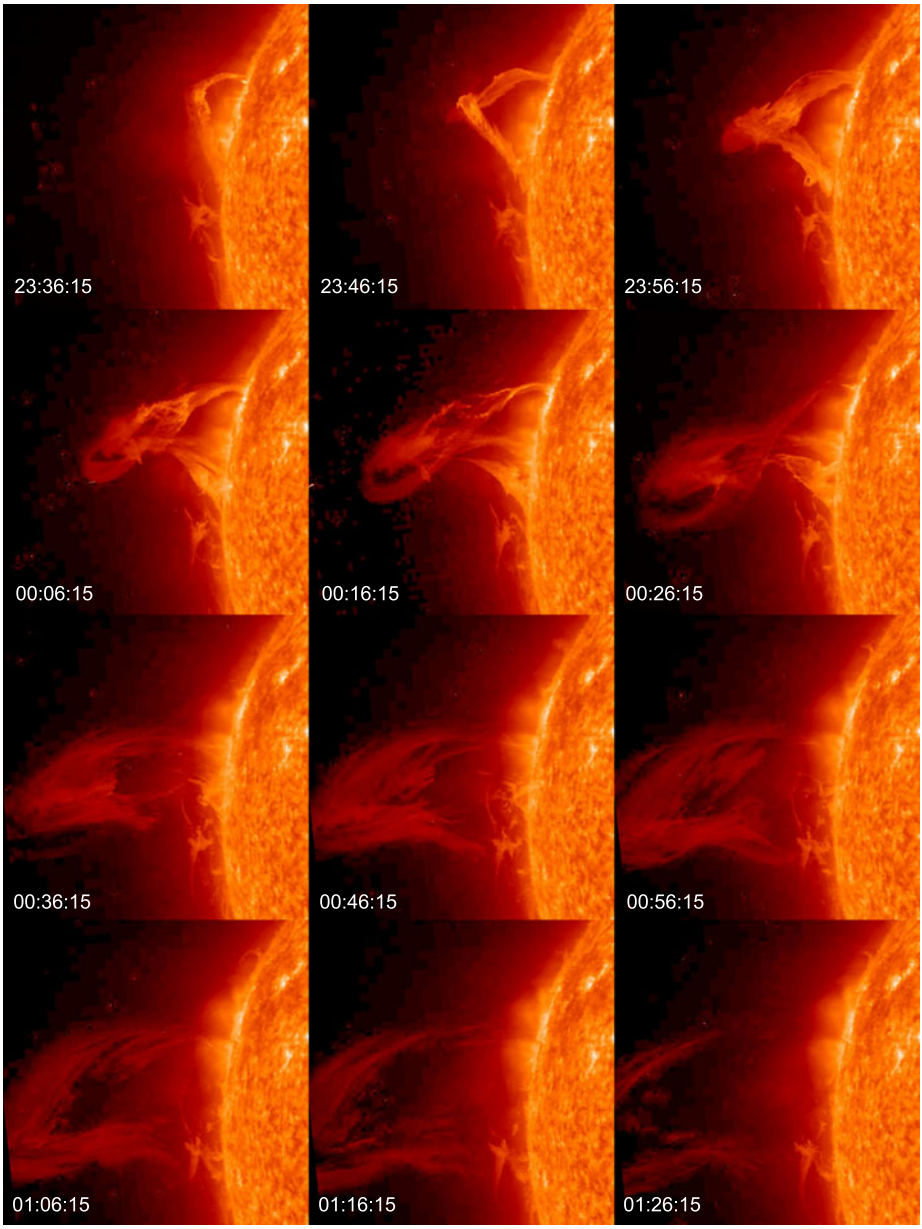


Figure 9 Non-radial eruption on 30 April – 1 May 2010 as observed by STEREO-A/EUVI 304 Å (the simultaneous observation of the eruption from the ground-based observatory (MLSO) is in the supplementary material).

frames in Figure 9 correspond to this period but observed from the different point of view of STEREO-A. The top of the prominence is bending away and to the left of the viewing direction of STEREO-A (Figure 9) creating an apparent kink and twist in the prominence ribbon. However, the apparent kink crossing may be an illusion from viewing the warped

ribbon of erupting filament mass seen in a two-dimensional image. MLSO and STEREO-B observations show the filament spine being straight, without any twisting, crossing the disk. The apparent kink disappears as seen in the final six frames in Figure 9. The images are suggestive of a kink structure until 00:26 UT. With the next 304 Å image (at 00:36 UT) the apparent kink disappears as the eruption proceeds along the southeast extension of the filament channel (final six frames in Figure 9).

Török, Berger, and Kliem (2010) have pointed out examples of kink unstable flux ropes where there is no crossing as seen from angles. While this may be the case, additional considerations may come from the properties of the line-tied kink instability. The kink instability requires, for a flux rope, a certain minimum number of turns that the magnetic field must make from end to end, not on the axis of the flux rope configuration, but in its immediate surroundings. Oftentimes this is described in terms of the ratio of the average poloidal flux (*i.e.* the flux circulating around the axis of the flux rope) to the axial flux (the integrated field along the flux rope axis or spine). The required twist for kink instability is at least 2.5π and depending on the configurations considerably more. This is more than one complete turn of the field in the domain (Hood and Priest, 1979; Velli, Hood, and Einaudi, 1990) and is a prerequisite, independent of observed kinking of the spine. If the spine is bent, and there is no tracer of plasma making one complete turn around the spine, the presumed axis of the prominence, the evidence for kinking becomes weak. The introduction of toroidal geometry (and the hoop force) does not change these constraints on the kink instability, *i.e.* the flux ropes of Török, Berger, and Kliem (2010) also have field lines wrapping around the axis for more than one complete turn. Because individual field lines are the only real relevant concept once line tying is taken into account (because the concept of a flux surface is lost if lines end in different regions in the photosphere), rather than flux surfaces, plasma tracers along such turning field lines should never disappear no matter what the point of view.

The position of the filament under its coronal arcade before its eruption was very asymmetric. Figure 10 shows the PFSS extrapolation of the magnetic field lines above the filament on 30 April 2010 18:04 UT. This asymmetry is thought to lead to asymmetric forces on the filament which cause the rolling motion of the filament at the early stages of the eruption,

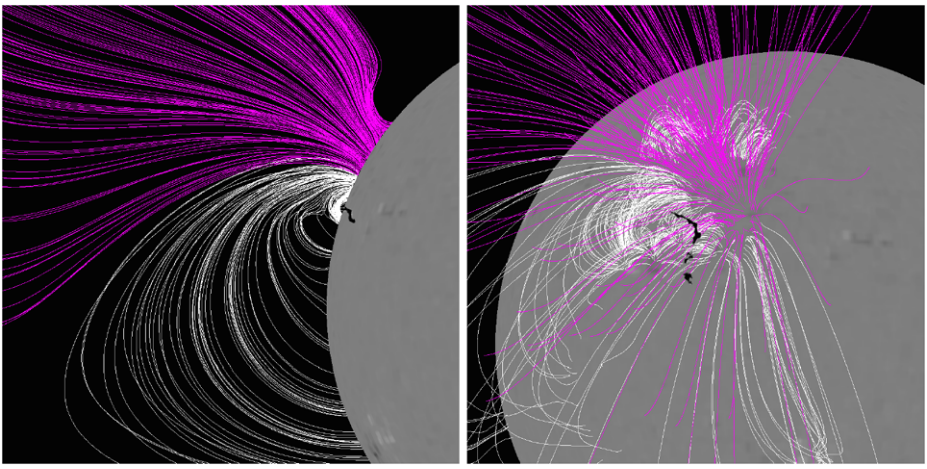


Figure 10 Superposition of the Hα filament and the PFSS reconstruction of the filament arcade and the open field of the neighboring coronal hole on 30 April 2010 18:04 UT. Left: Limb view. Right: Normal view.

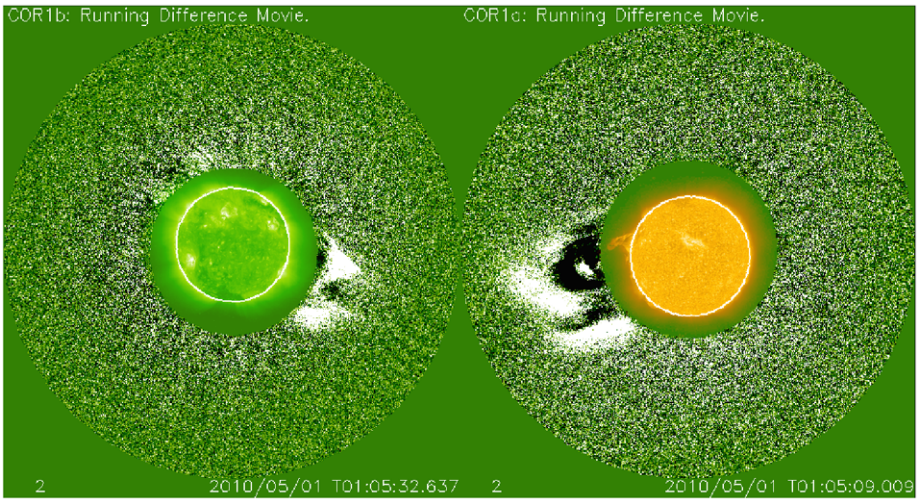


Figure 11 Difference images of the CME observed on 1 May 2010. Left: STEREO-B/COR1 superimposed with EUVI 195 Å. Right: Superposition of the STEREO-A/EUVI 304 Å prominence at 00:06 UT and the STEREO-A/COR1 image of the CME at 01:05 UT.

and create a condition for the non-radial propagation of the corresponding CME (Figure 11). The right image in Figure 11 shows the CME front and position of the prominence at the limb from STEREO-A; the CME non-radial propagation relatively to its point of origin is clearly visible.

4. Radial Eruptions with the Roll Effect

Here we consider cases where there is still significant rolling of the prominence spine before the CME is formed, but subsequently the filament and CME propagate outwards radially. Such cases are usually observed when the pre-eruptive filament is lying at the base of a pseudostreamer magnetic configuration where the arcades below the pseudostreamer have a much greater lateral extent relative to the pseudostreamer stalk which is essentially radial (an example of a non-radial pseudostreamer is shown in Figure 3). The asymmetry in the filament position relative to the overlying coronal arcade might lead to additional force on the filament that could enhance the rolling motion of the erupting filament during the early stage of the eruption. This lateral motion of the filament is directed towards the null point in the magnetic configuration. The radial geometry of the pseudostreamer branches (defined as bundles of open magnetic field lines on opposite sides of the spine-fan projection on the sky) in this case favors the radial propagation for the CME flux rope formed in the later stages of the filament eruption.

4.1. 1 August 2010 Twin Filaments Sympathetic Eruptions

Pseudostreamers appear in unipolar regions above multiple polarity reversal boundaries. Some of these polarity reversal boundaries can be filament channels, and when this is the case they can occur as twin filament channels often containing twin filaments. Figure 12 shows twin filaments and the coronal magnetic field configuration above them on 01 August

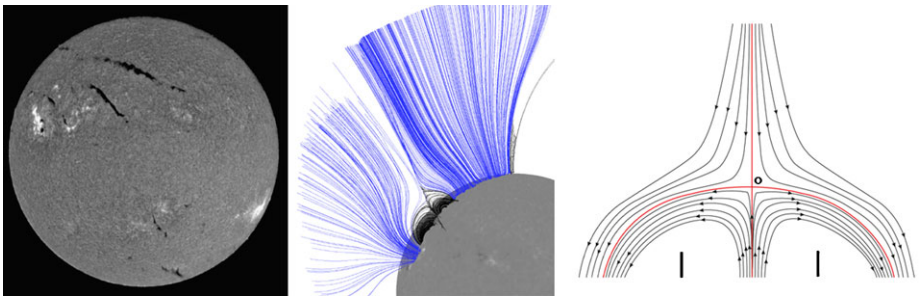


Figure 12 Left panel: BBSO Hα image of the long *twin filaments* in the northern hemisphere before they erupted on 1 August 2010. The eruptions were widely observed by SOHO, STEREO, SDO, and ground-based observatories. The separation in time between the eruptions of the two filaments was 11 h 30. Both twin filaments (filaments with the same chirality) were lying at the base of a pseudostreamer. Middle panel: PFSS extrapolation of the pseudostreamer magnetic field above two dextral filament channels beneath the depicted coronal arcades shown in black on 30 July 2010: rotated to limb view. Right panel: Magnetic configuration of the pseudostreamer and the position of the twin filaments (short vertical lines) relative to the null point, O, and magnetic separators (red lines).

2010. This event has been described in detail by Panasenco and Velli (2010), Schrijver and Title (2011), Török *et al.* (2011). The slight rolling motion of the erupting filaments toward the null point in the pseudostreamer configuration has been observed and reported in Török *et al.* (2011). Overall both CMEs, which originated under the pseudostreamer, propagated radially along the axis between the pseudostreamer branches.

4.2. 16 May 2007 South Polar Crown Filament Eruption

Figure 13 shows a PFSS extrapolation of the open field lines of the pseudostreamer and coronal loops overlying a filament under the southern lobe of the pseudostreamer on 15 May 2007. In the right column are anaglyph images from EUVI showing an erupting filament on 16 May 2007; the separation angle between STEREO-A and B is 8.14 degrees. To assist the reader, the upper edge of the filament is outlined by a red line and the lower edge by a light blue line. When one watches the 3D movies of this event created from STEREO/EUVI data, one can identify a rolling motion beginning to the east of the middle of the filament. A consequence of this rolling motion is that the bottom edge of the filament crosses gradually over the top edge. The roll appears to begin between the white Xs superposed on the blue line. It propagates in both directions away from the initial site (white Xs) and towards the apparent ends of the prominence as seen against the chromosphere. We also see that this erupting filament is erupting non-radially and rolling northward in the direction of the pseudostreamer null point (left panel in Figure 13).

5. Radial Eruptions Without the Roll Effect

The final examples are cases where there is no visible rolling of the filament, and both the filament and CME propagate radially outwards through the corona.

5.1. 3D Reconstruction of the Erupting Prominence on 28 February 2010

A polar crown prominence around the northern coronal hole (CH) erupted over the course of a day on 28 February 2010. The eruption was located on the far-side of the solar disk

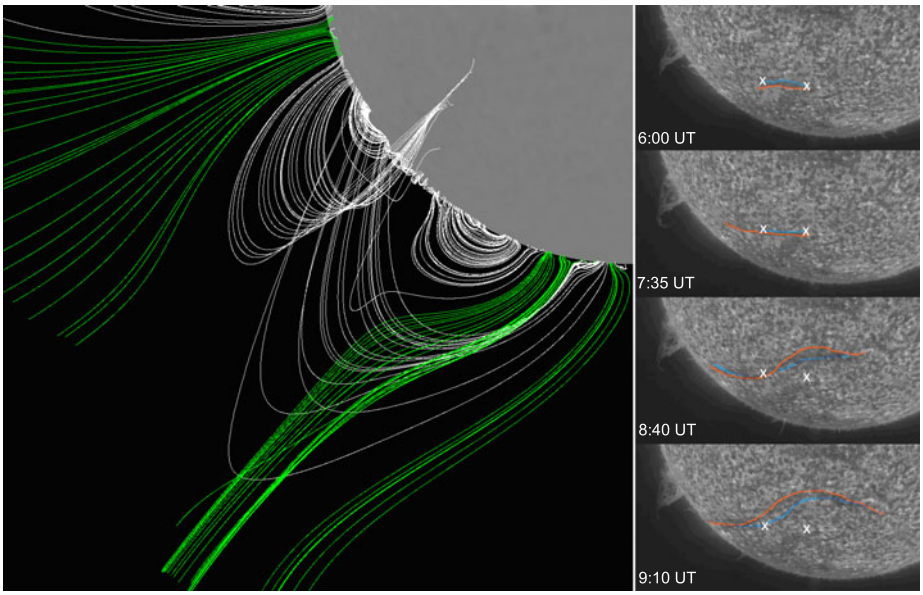


Figure 13 Left panel: PFSS extrapolation of the pseudostreamer magnetic field above the polar crown filament channel under the southern lobe of the pseudostreamer on 15 May 2007 06:06 UT rotated to limb view. Right panel: Images from STEREO/SECCHI/EUVI showing the rolling motion of the erupting filament on 16 May 2007. From this perspective the highest part of the pre-eruptive filament is shown by the red line and the lowest border by the blue line. The rolling motion is revealed as the initially highest part of the filament gradually becomes the lowest part and the approximate lowest part in the middle of the filament becomes the highest part. The roll appears to begin between the white Xs superposed on the blue line. It propagates in both directions away from the initial site (white Xs) and towards the apparent ends of the prominence as seen against the chromosphere.

but it was captured by the EUVI imagers onboard both STEREO satellites. The two EUV viewpoints revealed a very different behavior of the erupting prominence during its ascent, which is the reason we chose to analyze its 3D morphology in detail. In particular, EUVI-A shows a long piece of the polar crown prominence lifting off parallel to the surface. In contrast, the EUVI-B images show a very strongly kinked prominence (Figure 14).

It appears, therefore, that the apparent kinked morphology from one viewpoint could be just a projection effect. Since we have observations from two viewpoints, we can quantify this effect because we can localize the prominence in 3D space. Tie-pointing is the standard technique for this purpose and has been described before (Liewer *et al.*, 2009, 2011; Thompson, Kliem, and Török, 2012). The technique relies of the correct identification of the same structure in the two images. Filaments are ideally suited for this because of their highly detailed morphology. Because we are interested in the 3D shape of the large-scale structure of the prominence, we concentrate only on its 'backbone' in the EUVI 304 Å images. We use relatively few points to determine its orientation. We select six points for tie-pointing which are shown by the yellow crosses in Figure 14. To aid the reader, we connect some of those points with the dashed lines. We use the IDL routine *sunloop* which is part of the SECCHI software distribution in *Solarsoft*. The routine allows the user to select one feature in one viewpoint and then click on the same point (according to the use) on the other image. The program returns the location of the feature in 3D space defined by its longitude, latitude, and radial distance from the solar surface. Our results are shown in Figure 15 along with the

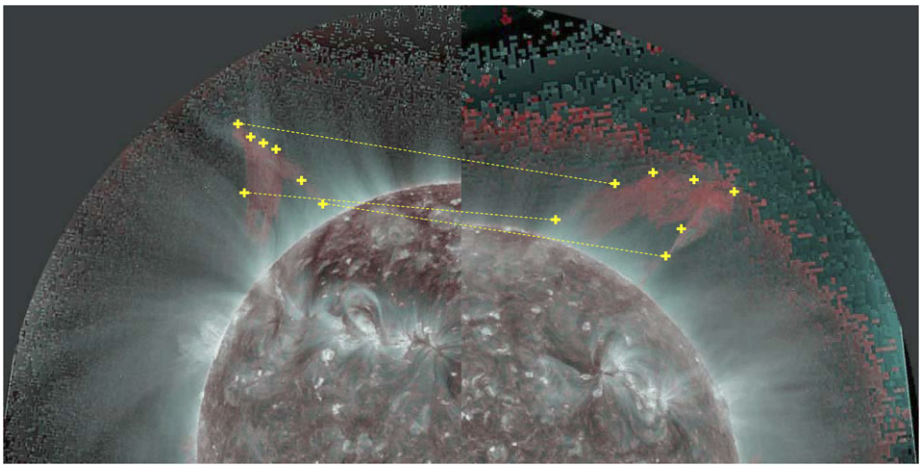


Figure 14 The 28 February 2010 eruption as seen by EUVI on STEREO-A (right) and -B (left). The images are composites of 304 Å (red) and 195 Å (silver) images. The time is 13:16 UT. The yellow stars mark the tie-pointing locations. The three yellow lines are visual aids to visualize the 3D configurations of the prominence.

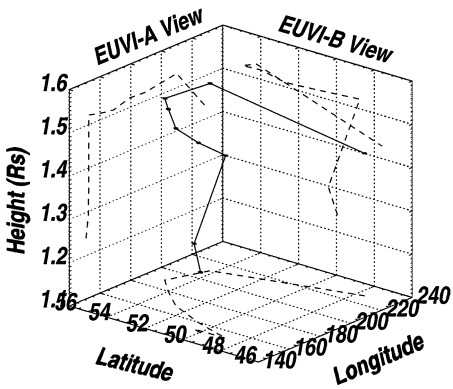


Figure 15 The 3D configuration of the prominence in the 28 February 2010 eruption using the points in Figure 14. The prominence has a small bend (about 6° wide) that creates the appearance of a kinked structure in the EUVI-A view. Note that the footpoints of the structure are not visible and hence have not been measured.

calculated projections of the prominence backbone for the EUVI-A and B views and on the surface.

The measurements verify that the appearance of a kink crossing is simply a projection effect. It is probably caused by the narrow shape of the prominence. The prominence is quite thin with a 70 deg longitudinal extent but only 8 deg of latitudinal extent. We cannot make a definitive measurement of the size of the prominence because its footpoints in the lower atmosphere are not visible from both EUVI telescopes. But we can easily trace the highest point along the backbone during the eruption. The resulting 3D height–time plots is shown in Figure 16 (left) where we see that the prominence is erupting quite radially; the latitude changes by only 4 deg while the point remains within a 10 deg longitudinal range. The height–time profile (Figure 16 (right)) is quite typical for such slow eruptions. We measure a very slow rising during the first seven hours followed by a rapid increase after 10 UT. We fit the height–time plot only for the period of the fast rise using a second degree polynomial. The resulting velocity profile is marked by the straight line of star symbols. We find that the

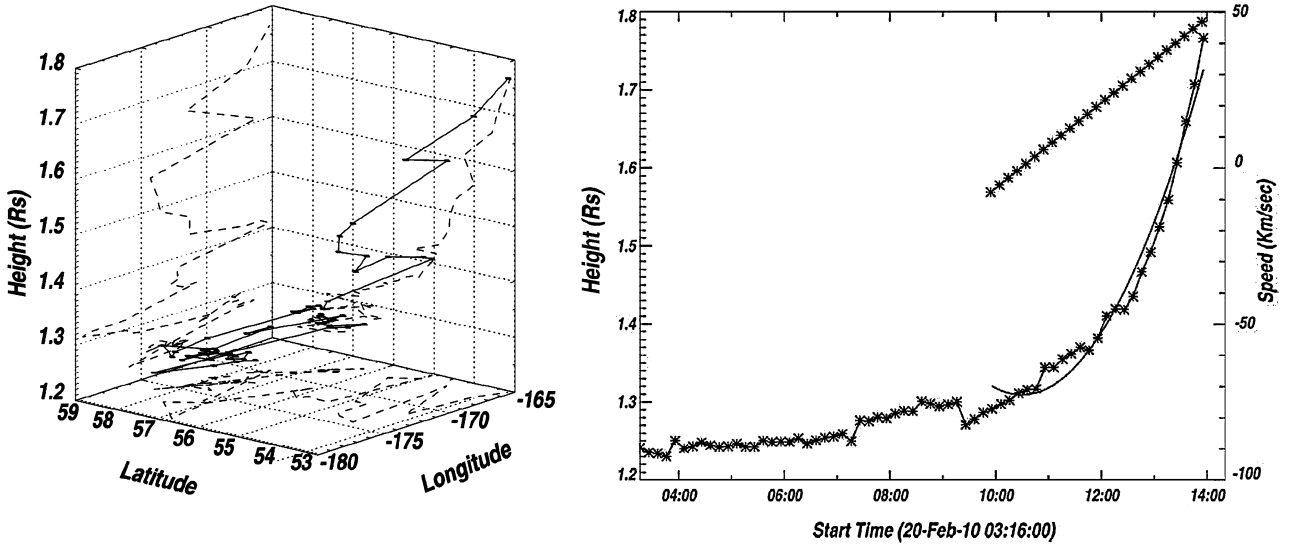


Figure 16 True radial height–time plots for the 28 February 2010 prominence eruption. The highest point of the structure was used. Left panel: 3D height–time plot of the prominence front. The eruption is deflected from radial by only 6° in latitude. The longitudinal variation is likely the result of errors in the identification of the same feature in both EUVI views. Right panel: Fit to the height–time plot. Only points after 10:00 UT were used. The velocity curve is shown by the straight line with stars and the values are shown on the right-hand side of the plot.

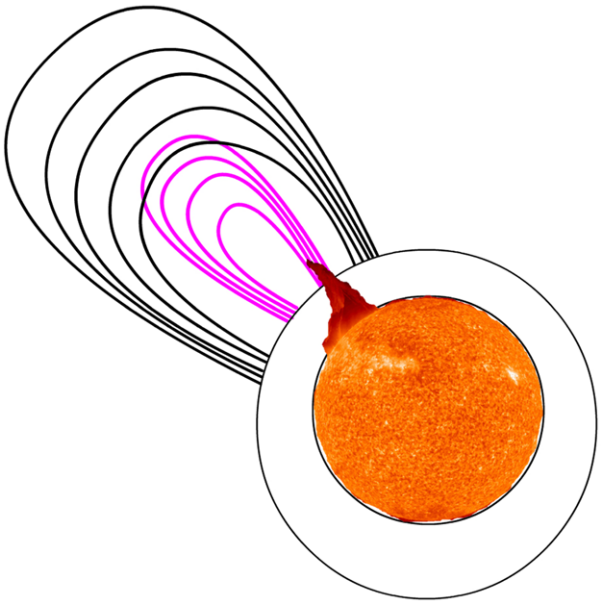


Figure 17 Radial eruption on 28 February 2010. Superposition of the outer boundaries of CME (black lines) and erupting prominence inside (magenta lines) as observed by STEREO-B/COR1 during different moments of the eruption and the solar disk with the prominence before eruption (STEREO-B/EUVI 304 Å).

prominence attains a speed of only 50 km s^{-1} at 14 UT when it leaves the EUVI field of view. The prominence and the corresponding CME propagate radially outwards through the corona (Figure 17).

In summary, the three-dimensional analysis of the 28 February 2010 prominence results in two important findings for this paper. First, we find that the kink morphology can be a result of projection effect which has important ramifications for past identification of kinked prominences. Second, we determine that the eruption is radial without any hint of equatorial deflection even though it is a slow prominent eruption (Figure 17). This is contrary to the prevailing consensus in the field, which holds that prominence eruptions, especially during solar minimum, tend to be deflected towards the magnetic equator. Although this is certainly true for some events, it should not be taken for granted for any event.

5.2. Radial Eruption on 6 July 2010

A polar crown prominence around the northern CH erupted on 6 July 2010. The eruption was captured by the EUVI imagers onboard both STEREO satellites and SDO. The erupting prominence appeared to rise very radially and showed very little bending of the prominence spine during this eruption (Figure 18). In Figure 19 we also examined the position of the filament before its eruption on 6 July (SDO/AIA) beneath the coronal fields generated by the PFSS model. The filament is centered symmetrically beneath the arcade of overlying coronal loops. Therefore, as the filament rises, the magnetic pressure of the magnetized filament cavity and overlying arcade should be approximately equal on the two sides of the filament. The absence of the lateral rolling motion of the filament spine is consistent with the deduced presence of the magnetic force balance between the fields on the two sides of the prominence.

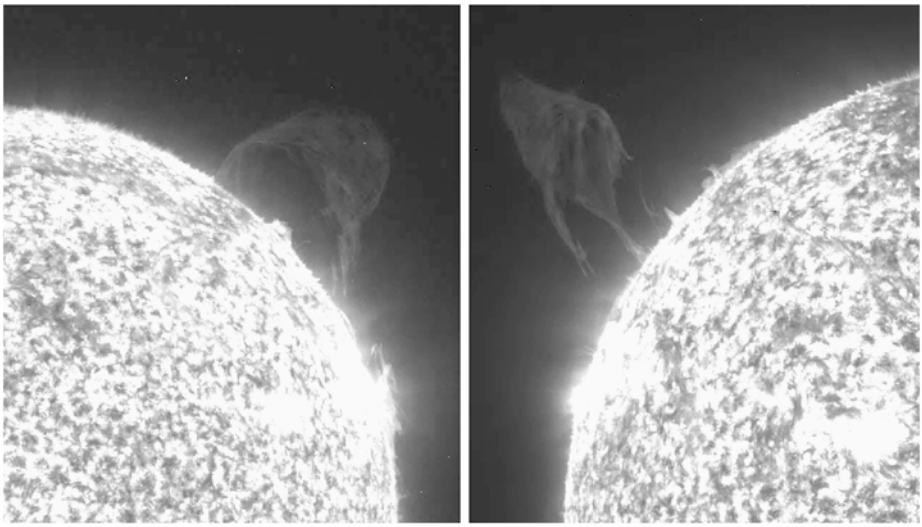


Figure 18 Radial eruption on 6 July 2010 seen from two perspectives. Left and right figures are STEREO-B and STEREO-A/EUVI 304 Å images captured during the eruption at 06:06:15 UT.

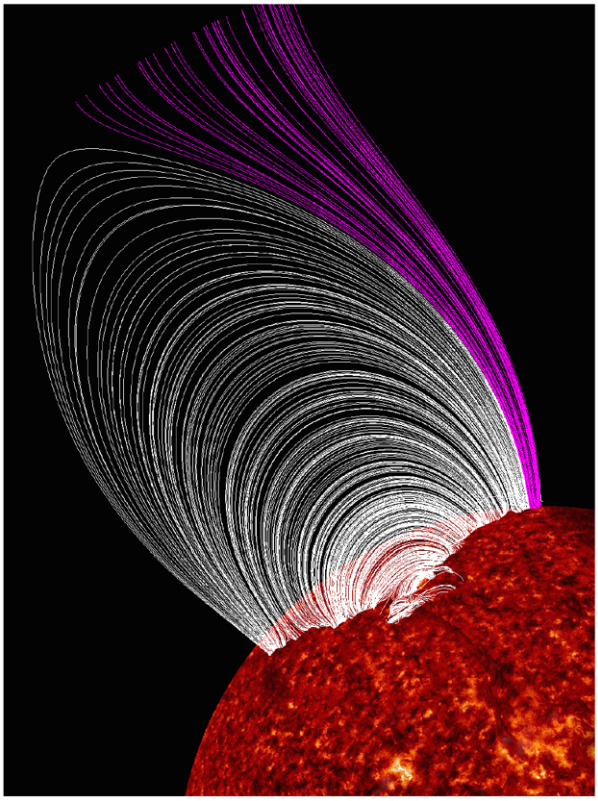


Figure 19 PFSS extrapolation of the open solar magnetic field and the coronal loops overlying the filament as observed by SDO 304 Å on 6 July 2010 at 01:14 UT. The filament is centered symmetrically beneath the arcade of overlying coronal loops.

6. Discussion

The path followed by the CME, and by the erupting filament immersed within it, depends on the macroscopic forces to which the erupting plasma is subject. Both the roll effect and the variable CME trajectories imply that the energy in the background corona is non-negligible compared to that liberated in the ejection. In the opposite case of negligible background energy, one would expect an essentially radial expulsion with a self-similar expanding spherical shock wave *à la* Sedov (1981). Describing the propagation of a part of a plasma cloud involves considerable difficulties, from the proper definition of the plasma cloud volume, to the separation of plasma cloud from the background field, to understanding the photospheric inertial effects of the parts of the plasma volume which remains connected to the Sun. Indeed, some or all of these problems are also intrinsic to numerical simulations. Consider then a volume of plasma as it is ejected and propagates outward into the solar corona. Parker (1957) focused on the possible shapes the plasma should take depending on the relative importance of kinetic and magnetic pressures, and then included the effects of the outward motion of the center of mass in a field with gradients. Assuming a general shape in the form of a deformable prolate spheroid, and using the virial theorem in its tensor form, he showed that if the internal pressure was dominated by the gas, rather than the magnetic field, the plasmoid would elongate into an infinitely thin shape aligned with the external magnetic field. Pneuman and Cargill (1985) showed that if the plasma β (ratio of kinetic to magnetic pressures) inside the body was very small, and it was initially in equilibrium with its surroundings, then it would move outward self-similarly, maintaining its shape. If we assume $\beta \ll 1$, the dominant force in determining the motion of the plasma center of mass will be the one arising from the magnetic field stresses, which, in the simplest case of a magnetically isolated plasma body, may be written as

$$\mathbf{F} = -\int_S \mathrm{d}\mathbf{S}\frac{B_\mathrm{e}^2}{8\pi}, \tag{1}$$

where S is the surface bounding the volume of the ejected material and B_e is the external magnetic field, as perturbed by the presence of the ejecting plasma itself. More generally the global force has a contribution coming from the non-diagonal part of the magnetic stress (magnetic tension) whose role is to eliminate the net force parallel to the magnetic field connecting the ejecta to the (arbitrarily defined at this point) outer magnetic field,

$$\mathbf{F}_i = -\int_S \mathrm{d}\mathbf{S}\frac{B_\mathrm{e}^2}{8\pi} + \int_S \frac{\mathbf{BB}\cdot\mathrm{d}\mathbf{S}}{4\pi}, \tag{2}$$

Parker (1957), Pneuman and Cargill (1985). For the cases of simple radial magnetic field background, these expressions lead to the so-called melon seed expulsion of diamagnetic plasmoids (see also Rappazzo *et al.*, 2005 for an application to blobs formed at the tips of helmet streamers).

These expressions are important because they show that regions in the corona separating strong and weak magnetic field regions, for example the boundaries of coronal holes, will act as walls repelling the motion of plasmoids from the strong field regions, whereas the neighborhoods of coronal null points (for example, between spine and fan for a pseudostreamer configuration), will be regions *towards* which an ejecting plasma will be accelerated, as they are regions of vanishingly small magnetic pressure. The dynamics of course is richer, and nonlinearly so, because as a consequence of such motions regions neighboring null points may develop strong current sheets, leading to attraction along a privileged axis (locally orthogonal to the tearing sheet) and repulsive in the plane of the sheet itself (a representative

example is given by the sympathetic eruptions simulated in Török *et al.*, 2011, where the breakout of one filament at first inhibits that of the other, but once sufficient flux is removed, the breakout of the second filament also follows). A detailed analysis of the acceleration of the filament and global CME requires separating the components of an eruption. Consider then the global CME as a large-scale plasmoid bubble, dominated by the cavity magnetic field, elongated along the filament axis. In this case one may apply the simplifications of a magnetically isolated system, compared to the rest of the coronal field, which in the case of a CME with small dimensions compared to the global coronal scales would lead to an expression for the force as something like

$$\mathbf{F} = -cV\nabla\frac{B^2}{8\pi} \tag{3}$$

with V the volume of the CME and c a constant $1.5 > c > 1$. This can explain departures from radial expansion due, for example, to the presence of a coronal hole, which occurs on large scales. For the filament itself, local deflections occurring in the neighborhood of the relatively low-lying null points in pseudostreamer or more general multipolar configurations, and complexities due to field lines not isolating the plasma entirely need to be taken into account. Differences in the acceleration of the filament (or filament spine) and the overall CME could therefore arise quite naturally in the initial phases of the eruption. More generally the expansion of the filament leads not only to macroscopic forces but also moments tending to rotate the prominence spine (responsible for the roll effect). Here the natural response of the filament should lead to deflection and alignment following the neighboring cavity field. As shown in the PFSS extrapolations, the asymmetries in the overlying arcade field correspond to both roll and deflection (Figure 20).

When multiple arcades are present below a pseudostreamer or a streamer configuration, the various polarity regions form a hierarchy of null points (and possible lines/fan structures). From the PFSS modeling we found that the lateral rolling motion of the filament is directed first towards the closest, *local* null point above the (asymmetric) coronal arcade before this arcade reconnects creating the CME flux rope. Once reconnection has occurred, the just-born CME flux rope then follows moving non-radially towards the higher, more distant *global* null points/lines separating larger-scale structures (the fan/spine null point of the large-scale pseudostreamer, or the tip of the helmet streamer for a globally bipolar configuration).

7. Conclusions

We have considered the global environment around CMEs, as can be approximated by the Potential Field Source Surface (PFSS) model. From our analyses, we have demonstrated the following relationships for eruptive solar events consisting of an erupting filament surrounded by a cavity and an enveloping CME:

- Both erupting filaments and their surrounding CMEs are non-radial only in the direction away from a nearby coronal hole and toward local and global null points. Due to the presence of the coronal hole, the global forces on the CME are asymmetric. The CME propagates non-radially in the direction of least resistance as we demonstrate by comparing low-latitude and high-latitude examples.
- The presence of lateral rolling motion in erupting filaments can account for their observed bending and twisting during eruption.

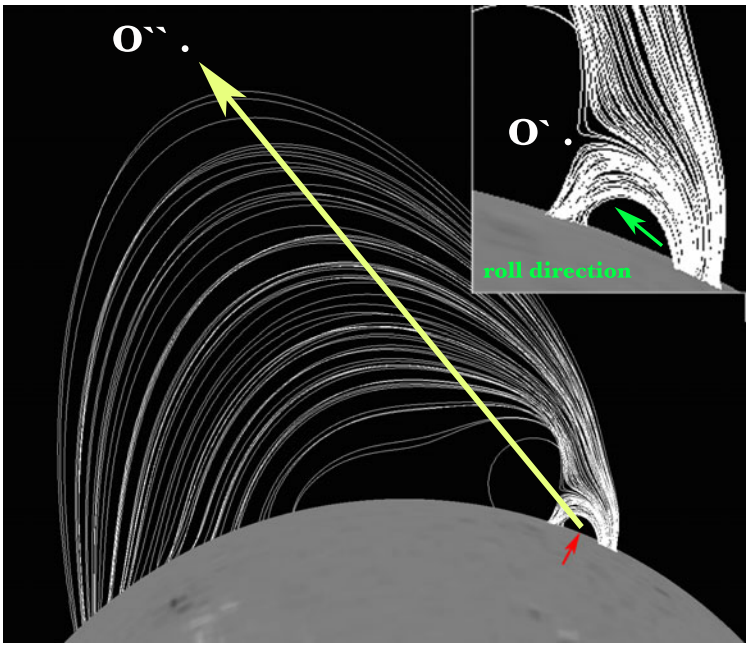


Figure 20 The PFSS extrapolation of the coronal loops overlying the filament channel indicated by the red arrow (30 April 2010; see Figures 10 – 11 for details). The green arrow indicates the direction of the rolling motion of the erupting filament, the yellow arrow indicates the direction of the CME deflection. O′ and O″ are local and global null points.

- The roll effect can depend, at least in part, on the asymmetries in the magnetic flux densities on the two sides of a polarity reversal boundary in a filament channel under a rising coronal loop system.
- The magnitude of the filament deflection from the radial propagation during its eruption is typically greater than the deflection of the corresponding CME. This difference can be explained by the different origins and character of filaments and CMEs:
  i) The spines of filaments exist as thin ribbons before their eruption and these ribbons roll, bend, and twist in characteristic ways during eruption.
  ii) Before their eruption, CME magnetic fields are tunnel-like systems of many fine coronal loops, a topology that allows an ascending loop system to reconnect and form a flux rope via magnetic reconnection below the associated and ascending filament ribbon. Such flux ropes therefore encapsulate the eruptive filaments, which still have some freedom to keep moving along the original direction of eruption, as defined prior to the flux rope formation.
- We confirm previous findings that the CME phase of rapid acceleration is initiated concurrently with the magnetic reconnection of a rising coronal loop system.
- In events with the roll effect and non-radial filament eruptions, but with radial CMEs, rolling motion and lateral deflection for the erupting filament is observed only during the early stages of the eruption.
- CME flux ropes commonly propagate radially for eruptive events associated with pseudostreamers.

- Non-radial and radial filament eruptions often show kinked shapes, which may, at least in part, be a consequence of projection effects from the rising ribbons of the filaments together with large-scale rolling motions.

Through modeling and comparison with observed events, we anticipate that major twists and non-radial motions in erupting prominences and CMEs will become predictable to the extent that their environments are well-defined and measurable.

Acknowledgements We are indebted to the SOHO, STEREO/SECCHI and SDO teams. O.P. and S.M. are supported in this research by the NASA grant NNX09AG27G. The work of M.V. was conducted at the Jet Propulsion Laboratory, California Institute of Technology under a contract from the National Aeronautics and Space Administration. A.V. is supported by NASA contract S-136361-Y to the Naval Research Laboratory. SOHO is a mission of international cooperation between ESA and NASA. The SECCHI data are produced by an international consortium of the NRL, LMSAL, and NASA GSFC (USA), RAL and Univ. Birmingham (UK), MPS (Germany), CSL (Belgium), IOTA and IAS (France). The AIA data used here are courtesy of SDO (NASA) and the AIA consortium.

References

Bemporad, A.: 2009, *Astrophys. J.* **701**, 298.
Brueckner, G.E., Howard, R.A., Koomen, M.J., Korendyke, C.M., Michels, D.J., Moses, J.D., *et al.*: 1995, *Solar Phys.* **162**, 357.
Cremades, H., Bothmer, V.: 2004, *Astron. Astrophys.* **422**, 307.
Cremades, H., Bothmer, V., Tripathi, D.: 2006, *Adv. Space Res.* **38**, 461.
Domingo, V., Fleck, B., Poland, A.I.: 1995, *Solar Phys.* **162**, 1.
Foukal, P.: 1971, *Solar Phys.* **19**, 59.
Filippov, B.P., Gopalswamy, N., Lozhechkin, A.V.: 2001, *Solar Phys.* **203**, 119.
Gaizauskas, V.: 1998, In: *New Perspective on Solar Prominences*, *IAU Colloquium 167* **CS-150**, 257.
Gaizauskas, V., Zirker, J.B., Sweetland, C., Covács, A.: 1997, *Astrophys. J.* **479**, 448.
Gopalswamy, N., Shimojo, M., Lu, W., Yashiro, S., Shibasaki, K., Howard, R.A.: 2003, *Astrophys. J.* **586**, 562.
Gopalswamy, N., Mäkelä, P., Xie, H., Akiyama, S., Yashiro, S.: 2009, *J. Geophys. Res.* **114**, A22.
Hood, A.W., Priest, E.R.: 1979, *Solar Phys.* **64**, 303.
Howard, R.A., Moses, J.D., Vourlidas, A., Newmark, J.S., Socker, D.G., Plunkett, S.P., *et al.*: 2008, *Space Sci. Rev.* **136**, 67.
Kaiser, M.L., Kucera, T.A., Davila, J.M., St. Cyr, O.C., Guhathakurta, M., Christian, E.: 2008, *Space Sci. Rev.* **136**, 5.
Kilpua, E.K.J., Pomoell, J., Vourlidas, A., Vainio, R., Luhmann, J., Li, Y., *et al.*: 2009, *Ann. Geophys.* **27**, 4491.
Lemen, J.R., Title, A.M., Akin, D.J., Boerner, P.F., Chou, C., Drake, J.F., *et al.*: 2012, *Solar Phys.* **275**, 17.
Liewer, P., Panasenco, O., Hall, J.R.: 2012, *Solar Phys.* doi:10.1007/s11207-012-0145-z.
Liewer, P.C., de Jong, E.M., Hall, J.R., Howard, R.A., Thompson, W.T., Culhane, J.L., Bone, L., van Driel-Gesztelyi, L.: 2009, *Solar Phys.* **256**, 57.
Liewer, P.C., Hall, J.R., Howard, R.A., de Jong, E.M., Thompson, W.T., Thernisien, A.: 2011, *J. Atmos. Solar-Terr. Phys.* **73**, 1173.
Martin, S.F.: 1990, In: Rudzjak, V., Tandberg-Hanssen, E. (eds.) *Proc. of the 117th Colloquium of IAU*, Springer, Berlin, 1.
Martin, S.F.: 1998, *Solar Phys.* **182**, 107.
Martin, S.F.: 2003, *Adv. Space Res.* **32**, 1883.
Martin, S.F., Lin, Y., Engvold, O.: 2008, *Solar Phys.* **250**, 31.
Martin, S.F., Panasenco, O., Engvold, O., Lin, Y.: 2008, *Ann. Geophys.* **26**, 3061.
Panasenco, O., Martin, S.: 2008, In: Howe, R., Komm, R.W., Balasubramaniam, K.S., Petrie, G.J.D. (eds.) *Subsurface and Atmospheric Influences on Solar Activity* **CS-383**, 243.
Panasenco, O., Velli, M.M.: 2010, AGU Fall Meeting Abstracts, A1663.
Panasenco, O., Martin, S., Joshi, A.D., Srivastava, N.: 2011, *J. Atmos. Solar-Terr. Phys.* **73**, 1129.
Parker, E.N.: 1957, *Astrophys. J. Suppl.* **3**, 51.
Pesnell, W.D., Thompson, B.J., Chamberlin, P.C.: 2012, *Solar Phys.* **275**, 3.
Pevtsov, A.A., Panasenco, O., Martin, S.F.: 2012, *Solar Phys.* **277**, 185.

Plunkett, S.P., Thompson, B.J., St. Cyr, O.C., Howard, R.A.: 2001, *J. Atmos. Solar-Terr. Phys.* **63**, 389.
Pneuman, G.W., Cargill, P.J.: 1985, *Astrophys. J.* **288**, 653.
Rappazzo, A.F., Velli, M., Einaudi, G., Dahlburg, R.B.: 2005, *Astrophys. J.* **633**, 474.
Schrijver, C.J., De Rosa, M.L.: 2003, *Solar Phys.* **212**, 165.
Schrijver, C.J., Title, A.M.: 2011, *J. Geophys. Res.* **116**, 4108.
Sedov, L.I.: 1981, *Similarity and Dimensionality Methods in Mechanics*, 9th edn., MIR, Moscow. English translation, 1985.
Sheeley, N.R., Warren, H.P. Jr.: 2012, *Astrophys. J.* **749**, 40.
Smith, S.F.: 1968, In: Kiepenheuer, K.O. (ed.) *Structure and Development of Solar Active Regions, Symposium no. 35*, D. Reidel, Dordrecht, 267.
Thompson, W.T., Kliem, B., Török, T.: 2012, *Solar Phys.* **276**, 241.
Török, T., Berger, M.A., Kliem, B.: 2010, *Astron. Astrophys.* **516**, A49.
Török, T., Panasenco, O., Titov, V.S., *et al.*: 2011, *Astrophys. J. Lett.* **739**, L63.
Velli, M., Hood, A.W., Einaudi, G.: 1990, *Astrophys. J.* **350**, 428.
Wang, Y.-M., Muglach, K.: 2007, *Astrophys. J.* **666**, 1284.
Wang, Y.-M., Muglach, K., Kliem, B.: 2009, *Astrophys. J.* **699**, 133.
Wang, Y.-M., Sheeley, N.R., Rich, N.B. Jr.: 2007, *Astrophys. J.* **658**, 1340.
Wood, P., Martens, P.: 2003, *Solar Phys.* **218**, 123.
Yeates, A.R., Mackay, D.H., van Ballegooijen, A.A.: 2007, *Solar Phys.* **245**, 87.
Zuccarello, F.P., Bemporad, A., Jacobs, C., Mierla, M., Poedts, S., Zuccarello, F.: 2012, *Astrophys. J.* **744**, 66.

DOI 10.1007/978-1-4899-8005-2_25
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0071-0

A Spatio-temporal Description of the Abrupt Changes in the Photospheric Magnetic and Lorentz-Force Vectors During the 15 February 2011 X2.2 Flare

G.J.D. Petrie

Received: 16 February 2012 / Accepted: 7 July 2012 / Published online: 2 August 2012

Abstract The active region NOAA 11158 produced the first X-class flare of Solar Cycle 24, an X2.2 flare at 01:44 UT on 15 February 2011. The *Helioseismic and Magnetic Imager* (HMI) instrument on the *Solar Dynamics Observatory* (SDO) satellite produces 12-minute, $0.5''$ pixel^{-1} vector magnetograms. Here we analyze a series of these data covering a 12-hour interval centered at the time of this flare. We describe the spatial distributions of the photospheric magnetic changes associated with the flare, including the abrupt changes in the field vector, vertical electric current and Lorentz-force vector acting on the solar interior. We also describe these parameters' temporal evolution. The abrupt magnetic changes were concentrated near the neutral line and in two neighboring sunspots. Near the neutral line, the field vectors became stronger and more horizontal during the flare and the shear increased. This was due to an increase in strength of the horizontal field components near the neutral line, most significant in the horizontal component parallel to the neutral line but the perpendicular component also increased in strength. The vertical component did not show a significant, permanent overall change at the neutral line. The increase in field strength at the neutral line was accompanied by a compensating decrease in field strength in the surrounding volume. In the two sunspots near the neutral line the integrated azimuthal field abruptly decreased during the flare but this change was permanent in only one of the spots. There was a large, abrupt, downward vertical Lorentz-force change acting on the solar interior during the flare, consistent with results of past analyses and recent theoretical work. The horizontal Lorentz force acted in opposite directions along each side of neutral line, with the two sunspots at each end subject to abrupt torsional forces relaxing their magnetic twist. These shearing forces were consistent with a contraction of field and decrease of shear near the neutral line, whereas the field itself became more sheared as a result of the field collapsing towards the neutral line from the surrounding volume. The Lorentz forces acting on the atmospheric volume above the photosphere were equal and opposite.

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

G.J.D. Petrie (✉)
National Solar Observatory, Tucson, AZ 85719, USA
e-mail: gpetrie@noao.edu

Keywords Magnetic fields · SDO/HMI · Solar flares · Solar photosphere

1. Introduction

Active region (AR) 11158 produced the first X-class flare of Solar Cycle 24, an X2.2 flare at 01:44 UT on 15 February 2011. The *Helioseismic and Magnetic Imager* (HMI) instrument (Schou *et al.*, 2012) on NASA's *Solar Dynamics Observatory* (SDO) satellite (Pesnell, Thompson, and Chamberlin, 2012) has been observing continuously since March 2010. HMI vector magnetograms covering several days including the time of this flare were released to the community in late 2011.[1] In this paper we analyze a series of these data covering a 12-hour interval centered at the time of the X2.2 flare.

SDO/HMI produces full-disk vector magnetograms with 0.5″ pixels every 12 minutes. The HMI instrument generates filtergrams in six polarization states at six wavelengths on the Fe I 617.3 nm spectral line. From these filtergrams, images for the Stokes parameters, I, Q, U, and V are derived. These are inverted for the magnetic vector components by the Very Fast Inversion of the Stokes Algorithm (VFISV) code (Borrero *et al.*, 2010). The 180° azimuthal field ambiguity is resolved using the "minimum energy" method (Metcalf, 1994; Leka *et al.*, 2009). The vector magnetograms specially released in late 2011 were derived by the HMI team from the HMI 720-second Stokes-parameter data series, running the VFISV inversion code with very strict convergence criteria to determine the optimal values, using more computing resources than the HMI pipeline is able to use routinely. The 180° ambiguity in the azimuthal field was computed with spherical geometry in a limited region with generous thresholds on field strength and very gradual annealing, also requiring significant additional computing resources. In this paper we use this data set to describe the abrupt and permanent field changes that occurred during the flare and characterize the associated Lorentz-force vector changes near the main neutral line of the region and within the neighboring sunspots. (Here a change is deemed "permanent" if its effects last at least several hours after the flare.)

After many decades of searching for evidence of flare-related field changes in the photosphere (*e.g.*, Severny, 1964; Zirin and Tanaka, 1981), abrupt, permanent photospheric field changes have been observationally linked to flares in the past two decades (Sudol and Harvey, 2005). Wang *et al.* (1992, 1994) found rapid and permanent field changes in flaring active regions, but a number of later studies produced inconclusive results; see the discussion in Wang (2006). Kosovichev and Zharkova (1999) reported a sudden decrease in magnetic energy near an X-class flare, during its impulsive phase. Then, Kosovichev and Zharkova (2001) reported on regions of permanent decrease of longitudinal magnetic flux in the vicinity of the magnetic neutral line near the 14 July 2001 "Bastille Day" flare and linked the change in flux to the release of magnetic energy. Wang and Liu (2010) studied 11 X-class flares for which vector magnetograms were available, and found in each case an increase of transverse field at the polarity-inversion line. Wang *et al.* (2012) and Sun *et al.* (2012) analyzed the HMI data for the 15 February 2011 X2.2 flare, the same data set studied in the present paper, and found similar behavior. These observations support the coronal implosion interpretation (Hudson, 2000; Hudson, Fisher, and Welsch, 2008; Fisher *et al.*, 2012) where, after a coronal magnetic eruption, the remaining coronal field contracts downward resulting in the field become more horizontal at the photospheric level. Fletcher and Hudson (2008)

[1] See http://jsoc.stanford.edu/jsocwiki/ReleaseNotes and http://jsoc.stanford.edu/jsocwiki/VectorPaper.

have given the only detailed explanation so far of how a coronal event could permanently alter a photospheric field.

Distinctive patterns of behavior have also been found in the behavior of sunspot magnetic fields during flares. Parts of the outer penumbral structures decay rapidly after many flares, while neighboring umbral cores and inner penumbral regions become darker (Wang *et al.*, 2004, 2005; Deng *et al.*, 2005, Liu *et al.*, 2005; Wang, Zhao, and Zhou, 2009). Transverse fields have been found by these authors to decrease in the regions of penumbral decay and to increase at the flare neutral lines. Li *et al.* (2009) found that during the 13 December 2006 X3.4 flare the mean inclination angle of the magnetic field increased in the part of the penumbra that decayed, whereas the inclination angle decreased in the part of the penumbra that was enhanced during the flare and near the magnetic neutral line. The magnetic twist of sunspot fields has been observed to decrease abruptly as a result of flares (Ravindra *et al.*, 2011; Inoue *et al.*, 2011). Gosain, Venkatakrishnan, and Tiwari (2009) found coherent lateral motion of the penumbral filaments near the neutral line using high resolution *Hinode* *G*-band images of the 13 December 2006 X3.4 flare, and speculated that these motions were due to impulsive horizontal Lorentz-force changes. Gosain and Venkatakrishnan (2010) investigated the evolution of the vector field during the 13 December 2006 X3.4 flare and found that, in the penumbra of the main sunspot, the observed field was more inclined than the equivalent potential field, and the difference between the observed and potential fields steadily increased before the flare, abruptly decreased during the flare and steadily increased again after the flare. AR 11158 featured much sunspot evolution during the time of the 15 February 2011 X2.2 flare (Liu *et al.*, 2011; Jiang *et al.*, 2012) but the associated magnetic field changes have not been studied in detail. We will do so in this paper, relating the sunspot field changes to those near the main neutral line and calculating the associated changes in the Lorentz-force vector.

Most of the studies described above have focused on field changes near neutral lines or in sunspots. Sudol and Harvey (2005) and Petrie and Sudol (2010) adopted a more general approach. Using one-minute the National Solar Observatory's *Global Oscillations Network Group* (GONG) longitudinal (line-of-sight) magnetograms, Sudol and Harvey (2005) characterized the spatial distribution, strength, and rate of change of permanent field changes associated with 15 X-class flares. By carefully co-registering the images they succeeded in tracing the field changes pixel-by-pixel and were able to show the spatial structure of the changes. They found that the majority of field changes occurred in regions where the field strength reached hundreds of Gauss which, given the $2.5''$ pixel^{-1} resolution of the data, suggests locations close to or within sunspots. Building on Sudol and Harvey's work, Petrie and Sudol (2010) analyzed one-minute GONG longitudinal magnetograms covering 77 flares of GOES class at least M5 and found some statistically significant correlations in the field changes. Exploring the relationship between increasing/decreasing longitudinal fields and azimuth and tilt angles at various positions on the disk, they noted that increasing/decreasing longitudinal fields do not correspond straightforwardly to decreasing/increasing changes in tilt angles. However, the overall distributions of longitudinal increases and decreases at different parts of the disk was found to be consistent with Hudson, Fisher, and Welsch's (2008) loop-collapse scenario.

The 15 February 2011 X2.2 flare has already been studied in several papers using a variety of observations and methods. As mentioned above, Wang *et al.* (2012) analyzed the HMI data for the 15 February 2011 X2.2 flare, the same data set studied in the present paper, and found an increase of transverse field and field inclination at the polarity-inversion line (see also Gosain, 2012; Sun *et al.*, 2012). Sun *et al.* (2012) calculated nonlinear force-free field models for the coronal field from the HMI vector measurements and argued that

the increase in magnetic shear observed at the photosphere is localized at low heights and the shear decreases above a certain height in the corona. Jiang *et al.* (2012) described the complex sunspot motions seen around the time of the flare in G-band images from the *Solar Optical Telescope* on the *Hinode* satellite and continuum intensity images from SDO/HMI, with particular emphasis on the clockwise motion of the positive sunspot neighboring the main neutral line. Beauregard, Verma, and Denker (2012) measured the horizontal proper motions with local correlation tracking using HMI continuum images and longitudinal magnetograms, and found shear flows along the main neutral line of a few 100 m s^{-1}. Schrijver *et al.* (2011) used multi-wavelength data from the *Atmospheric Imaging Assembly* (AIA) on SDO with high spatial and temporal resolution to analyze expanding loops from a flux-rope-like structure over the shearing polarity-inversion line between the central δ-spot groups of AR 11158 that eventually formed a coronal mass ejection moving into the inner heliosphere. Gosain (2012) used AIA observations to study the evolution of the coronal loops in the region and identified three distinct phases of the coronal loop dynamics during this event: a slow rise phase, a collapse phase and an oscillation phase. In this paper we will focus on characterizing in detail the vector-field changes that occurred during the flare by analyzing the three field components and the associated electric currents and Lorentz-force changes over a ten-hour interval centered at the start time of the flare.

The paper is organized as follows. In Section 2 we will present the vector fields observed by HMI before and after the main flare-related field changes took place. We will describe the differences between these vector fields in each spatial dimension and plot the vector field's evolution in time. We will describe in Section 3 the associated electric current changes that occurred during the flare. We will derive the accompanying Lorentz-force changes in Section 4 and discuss the likely causes of the changes. We will conclude in Section 5.

2. The Magnetic Field Vector Changes

The X2.2 flare began in NOAA AR 11158 on 15 February 2011 at 01:44 UT when the region was visible on the solar disk at about 20° south and 10° west of disk-center. The vector-field measurements were released by the HMI team in the form of 12-minute vector magnetogram images (B_r, B_θ, B_ϕ) in heliographic coordinates (r, θ, ϕ) on a 600×600 grid with pixel size 0.03°. The top panel of Figure 1 shows a spatial map of the vertical magnetic field component, B_r, before the flare (the image labeled 1:36 UT), with the corresponding horizontal field, (B_ϕ, $-B_\theta$), indicated by arrows. The magnetic field of the region had a complex structure with four major concentrations of intense field. The entire distribution was tilted with respect to the equator. The leading polarity concentration, which was positive, was the most equatorward and the lagging field concentration, which was negative, was the most poleward. So far this arrangement is in line with the Hale–Nicholson law. However, the most interesting part of the active region lay between the leading and lagging field concentrations. There was a bipolar structure composed of a positive leading field concentration and a negative lagging concentration at the same latitude, separated by a highly sheared, S-shaped neutral line that was tilted with respect to the equator at approximately the same angle as the region as a whole. The shear of the field at the neutral line is clearly visible in Figure 1, as are the clockwise and anti-clockwise circulations of the field in the negative and positive central field concentrations, respectively. It is in this central portion of the active region that most of the action occurred during the flare, as the bottom panel of Figure 1 shows. If we have observations of the photospheric vector field at two times, $t = 0$ before the field changes begin, and $t = \delta t$ after the main field changes have occurred, the magnetic vector changes due to the flare can be represented by the difference $\delta\mathbf{B} = \mathbf{B}(\delta t) - \mathbf{B}(0)$. The bottom panel of

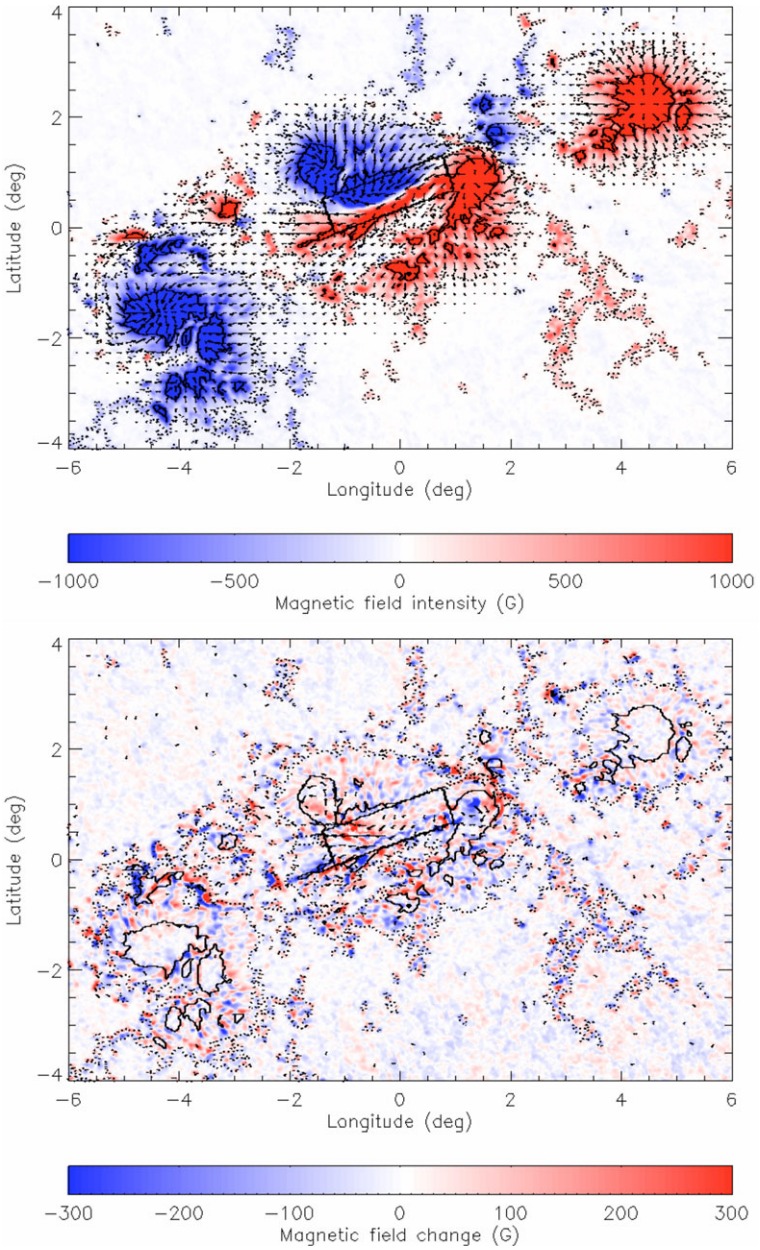


Figure 1 Shown are the vector magnetic field before the flare (top) and the vector-field changes (bottom). The vertical components, B_r and δB_r, are indicated by the color scales and the horizontal components by the arrows, with saturation values ± 1000 G in the top plot and ± 300 G in the bottom plot. Red/blue coloring represents positive/negative vertical field and field changes. The black rectangles mark the region near the neutral line that is used in subsequent analysis. The solid and dotted contours indicate strong ($|B_r| > 1000$ G) and quite strong ($|B_r| > 100$ G) fields, respectively.

Figure 1 shows a spatial map of the vertical magnetic field change, δB_r, with the horizontal field changes $(\delta B_\phi, -\delta B_\theta)$ indicated by arrows. This map was derived by subtracting the 1:36 UT from the 2:00 UT image for each of the three field components. The maps show that the vertical changes were mostly positive/negative on the negative/positive side of the neutral line, weakening the vertical field on both sides of the neutral line, while the horizontal changes point in approximately the same direction as the field itself near the neutral line, strengthening the horizontal field there. Meanwhile the vertical changes in the two central sunspots were mostly positive/negative in the negative/positive spot, weakening the vertical field in both spots. The horizontal changes were anti-clockwise in the negative central spot and clockwise in the positive central spot, weakening the azimuthal field component in each spot.

Figure 2 shows spatial maps of the changes in the field tilt $\tan^{-1}([B_\theta^2 + B_\phi^2]^{1/2}/B_r)$ and the total field strength $B = (B_r^2 + B_\theta^2 + B_\phi^2)^{1/2}$ during the flare. There was a clear increase in the tilt angle of the field near the neutral line during the flare, accompanied by a clear increase in total field strength near the neutral line, particularly at the east side of the rectangle in the plot. Also evident is a general decrease in total field strength in the surrounding field outside the rectangle, including much of the field in the two neighboring sunspots. This pattern of increasing tilt and field strength near the neutral line and decreasing tilt and field strength in the surrounding volume is consistent with field abruptly collapsing downwards towards the neutral line and field rushing in from the surrounding volume to fill the resulting void above the neutral line.

Figure 3 shows spatial maps of the changes in the unsigned vertical field, $|B_r|$, and in the horizontal field, $B_\mathrm{h} = (B_\theta^2 + B_\phi^2)^{1/2}$. There is some evidence of weakening of the vertical field during the flare but the distribution of changes is quite complex. The distribution of horizontal field changes is much more striking, with increasing horizontal fields clearly dominating the region near the neutral line surrounded by an area of decreasing horizontal fields. The increase in total field near the neutral line seen in Figure 2 was due to the large increase in horizontal field during the flare shown in the bottom panel of Figure 3. In fact this horizontal field increase mostly added field in the direction parallel to the neutral line as Figure 1 shows, and as is clearer in the corresponding temporal plots discussed below.

We next discuss the temporal profiles of the magnetic changes, shown in Figures 4 – 7. These and subsequent plots of temporal changes were derived by calculating area integrals of the field components over chosen photospheric areas in the 60 12-minute images, from 20:00 UT on 14 February to 7:48 UT on 15 February. For example, the top panel of Figure 4 shows the evolution of the integrated magnetic field strength

$$B^\mathrm{NL} = \int_{A_\mathrm{NL}} B \, \mathrm{d}A, \tag{1}$$

near the neutral line between 20:00 UT on 14 February and 7:48 UT on 15 February. The area A_NL corresponds to the rectangular region near the neutral line marked in Figure 1. The middle and bottom panels of Figure 4 show the equivalent integrals B_r^NL and B_h^NL of B_r and B_h. Near the neutral line, B^NL increased abruptly during the flare because of an increase in the horizontal field there. From the temporal plots it is clear that the vertical field near the neutral line did not change significantly overall during the flare, and the change that occurred did not have a permanent effect.

Figure 5 shows the evolution of the integrated horizontal magnetic field components parallel and perpendicular to the neutral line, $B_\parallel^\mathrm{NL}$ and $B_\perp^\mathrm{NL}$. These directions are defined by the black rectangle in Figure 1 and the integrals are evaluated over the area A_NL represented by this rectangle. The parallel direction is the direction of the long edges of the rectangle,

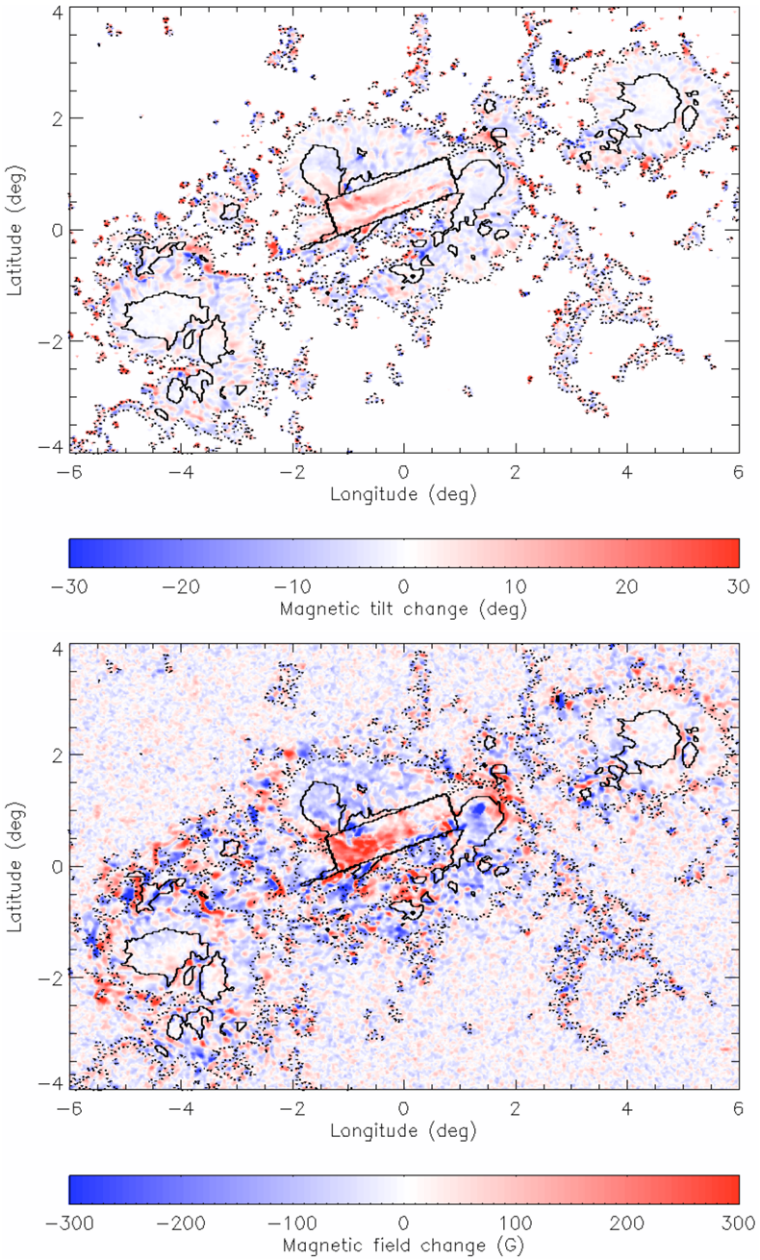


Figure 2 Plotted here are spatial maps of the changes in the field tilt (top) and the total field strength (bottom) during the flare. Red/blue coloring represents positive/negative change with saturation values $\pm 30°$ (top) and ± 300 G (bottom). The black rectangles mark the region near the neutral line that is used in subsequent analysis. The solid and dotted contours indicate strong ($|B_r| > 1000$ G) and quite strong ($|B_r| > 100$ G) fields, respectively.

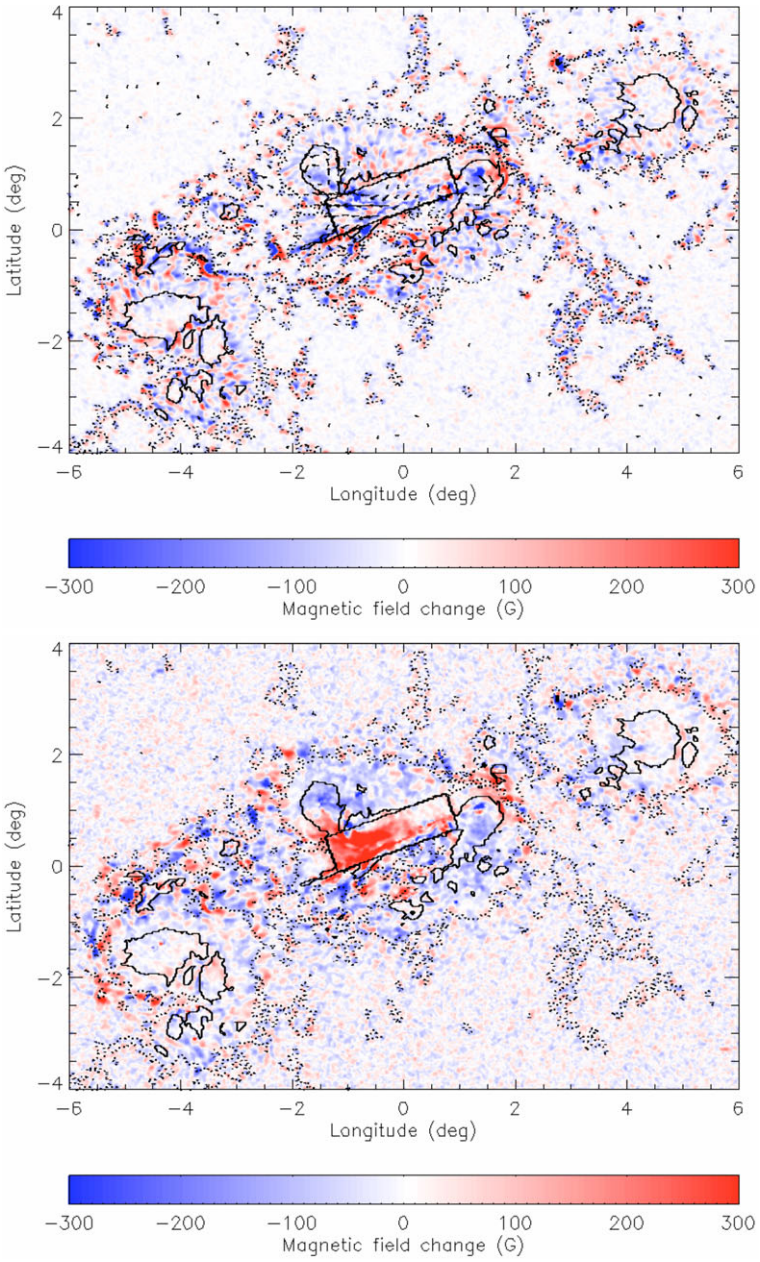


Figure 3 Shown here are spatial maps of the changes in the unsigned vertical field, $\delta|B_r|$ (top), and the total horizontal field changes, $\delta B_{\rm h}$ (bottom). Red/blue coloring represents positive/negative field change with saturation values ± 300 G. The arrows in the top plot are the same as those in the bottom plot of Figure 1. The black rectangles mark the region near the neutral line that is used in subsequent analysis. The solid and dotted contours indicate strong ($|B_r| > 1000$ G) and quite strong ($|B_r| > 100$ G) fields, respectively.

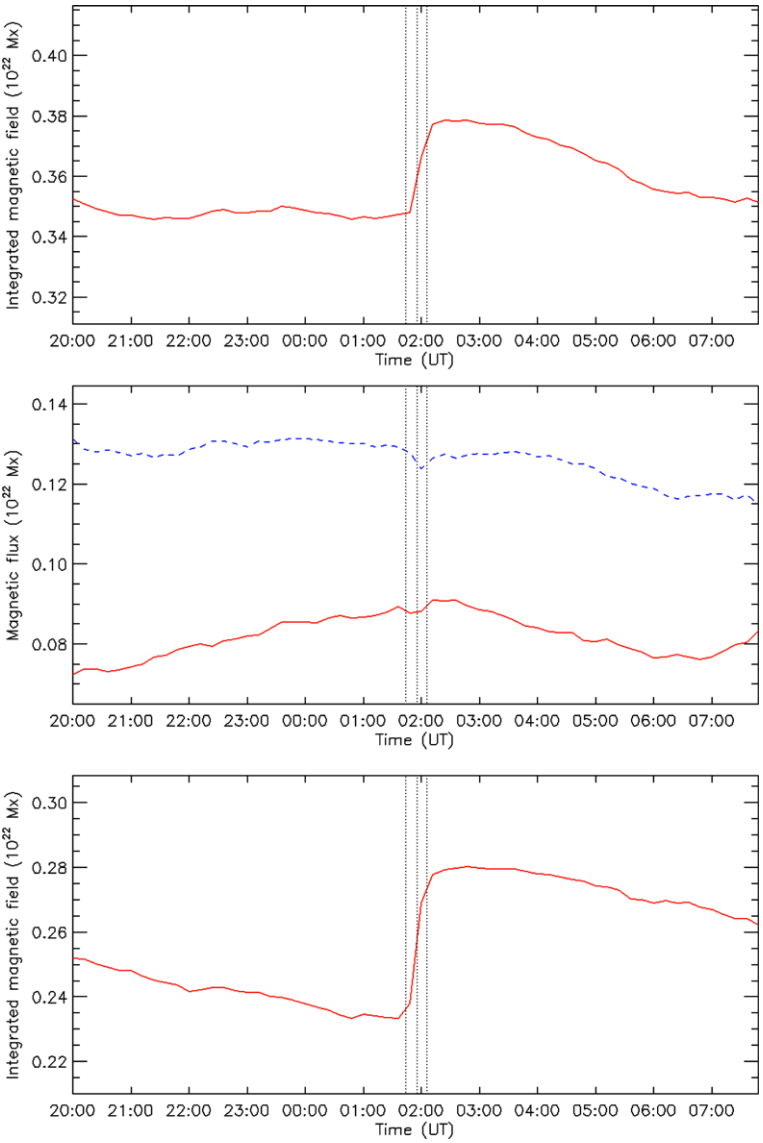


Figure 4 Shown here are the integrated magnetic field strength B^{NL} (top), the total vertical flux B_r^{NL}, (middle) and the integrated horizontal field $B_{\mathrm{h}}^{\mathrm{NL}}$, (bottom) near the neutral line are plotted as functions of time. The red/blue solid/dashed lines represent positive/negative field. The area of integration is indicated by the rectangle in Figure 1. The vertical lines represent the GOES flare start, peak, and end times.

pointing approximately west-north-west. The perpendicular direction is the direction of the short edges of the rectangle, pointing approximately north-north-east. The horizontal field increased during the flare near the neutral line, both parallel and perpendicular to the neutral line. The change in the horizontal component parallel to the neutral line was the most significant change. The pre- and post-flare evolution of the horizontal field was more steady in the parallel than in the perpendicular component. The total field near the neutral line de-

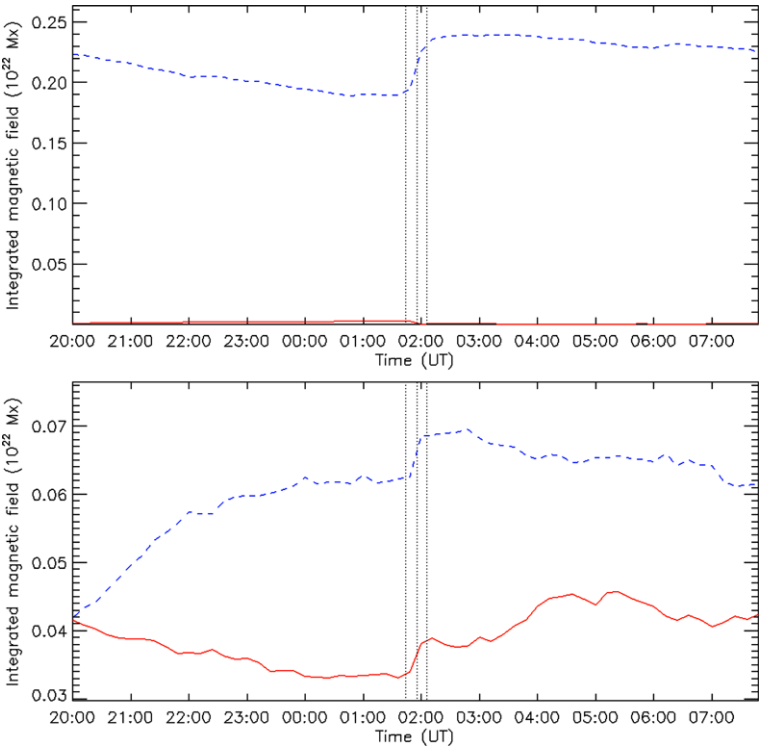


Figure 5 The integrals $B_{\parallel}^{\mathrm{NL}}$ (top) and $B_{\perp}^{\mathrm{NL}}$ (bottom) of the horizontal field components parallel and perpendicular to the neutral line are plotted as functions of time. The red/blue solid/dashed lines represent positive/negative field. The area of integration is indicated by the rectangle in Figure 1. The vertical lines represent the GOES flare start, peak, and end times.

creased steadily after the flare, reaching its pre-flare value about five hours after the flare. This decrease was due to changes in both horizontal and vertical components. The horizontal parallel component, however, remained significantly stronger five hours after the flare than its pre-flare value.

Flare-induced line profile changes can produce signatures that do not represent real changes in the magnetic field as discussed by Sudol and Harvey (2005) – see the bottom part of their Figure 1. Working with GONG 1-minute longitudinal field images, Sudol and Harvey (2005) and Petrie and Sudol (2010) fitted a $\tan^{-1}$ step-like function to the time profile of each pixel, applied selection criteria based on the quality of the function fits, and inspected the results for representative pixels by eye. For the HMI vector data this approach is not as helpful because the sensitivity of the HMI vector-field inversions is not as good as the sensitivity of the GONG data. Furthermore, the 12-minute time resolution does not resolve as many of the pixel field changes as the 1-minute GONG data do. Fits of $\tan^{-1}$ functions to HMI pixel time profiles are therefore not so useful in distinguising real field evolution from artifacts.

We have derived estimates of the field vector changes in two ways. We derived spatial maps by calculating pixel-by-pixel differences between before/after image pairs, the before image being composed of observations of the field before the published GOES start time of the X-ray flare and the after image deriving from observations of the field after the main

flare-related field changes have taken place. This calculation therefore excluded the time when most of the flare emission transients are expected to have occurred and so the derived difference maps are not expected to be significantly compromised by such artifacts. In the second calculation we integrated field components over many-pixel regions of interest such as the rectangle surrounding the main neutral line marked in Figure 1. We evaluated these integrals for all 60 magnetograms and formed time series of field integrals. This calculation does not avoid the time of the flare when emission might occur. We therefore need to be on the lookout for emission artifacts in the curves of the integrated quantities. Figures 4 and 5 show that in this data set there is no significant signature of an emission artifact – compare with the bottom part of Sudol and Harvey's (2005) Figure 1. In particular, the profiles of B^{NL}, $B_{\mathrm{h}}^{\mathrm{NL}}$ and $B_{\parallel}^{\mathrm{NL}}$ have clear, stepwise changes with no sign of an emission artifact. This shows that, while emission transients may have affected some of the pixels, the calculations of the integrated quantities plotted in Figures 4 and 5 were not significantly compromised by artifacts.

Besides occurring near neutral lines, abrupt field changes have been observed to occur in sunspots, as we discussed in Section 1. To accompany the above analysis of the changes in heliographic coordinates, separate analyses of the field changes were performed for the sunspots of the active region in local cylindrical coordinates (R, Θ, Z) with $R = 0$ located at the sunspot center in each frame. The sunspot locations were tracked from frame to frame by first approximating them as linear functions of time and then searching for the location where the radial field vanished and the integral of the surrounding radial field was maximized. Smooth functions of position resulted from these estimates.

Figures 6 and 7 show the magnetic field evolution near the negative and positive inner sunspots in the radial, azimuthal, and vertical directions as functions of time. Here the field components B_i are integrated over the negative and positive sunspot areas A^{NS} and A^{PS} to give the integrated field components B_i^{NS} and B_i^{PS}, where $i = R, \Theta, Z$. In the positive/negative sunspot all three integrated field components were positive/negative. A flux tube connecting the two spots would therefore be expected to have positive (right-handed) magnetic twist. Nonlinear force-free coronal field models extrapolated from these HMI photospheric vector magnetograms do indeed have positive relative magnetic helicity (Sun *et al.*, 2012). The positive sunspot had nearly twice as much integrated azimuthal field as the negative sunspot. The other components were closer to equal strength in the two spots, with the vertical component larger than the other components in each case. In the negative sunspot the dominant negative azimuthal component abruptly decreased during the flare and the positive azimuthal field continued a gradually increasing trend through the flare, so that the net integrated azimuthal field was small at the end of the flare. After the flare the dominant negative azimuthal component gradually returned to its original strength over the next several hours and the negative azimuthal field abruptly decreased, increasing the net integrated azimuthal field to almost its pre-flare value. The other two components did not show such a striking change. In the positive sunspot the dominant positive azimuthal component abruptly decreased during the flare and in this case the change was permanent. The negative integrated azimuthal field increased briefly during the flare but it was much smaller than the positive integrated azimuthal field throughout the time series. The other two components did not show such a striking change. In both spots the azimuthal field components changed more than the other components.

The magnetic twist of sunspot fields has been observed to decrease abruptly as a result of flares (Ravindra *et al.*, 2011; Inoue *et al.*, 2011). Modeling the emergence of a twisted flux tube from the interior through the photosphere into the corona, Longcope and Welsch (2000) predicted that the expansion of the tube into the corona would redistribute the twist, creating

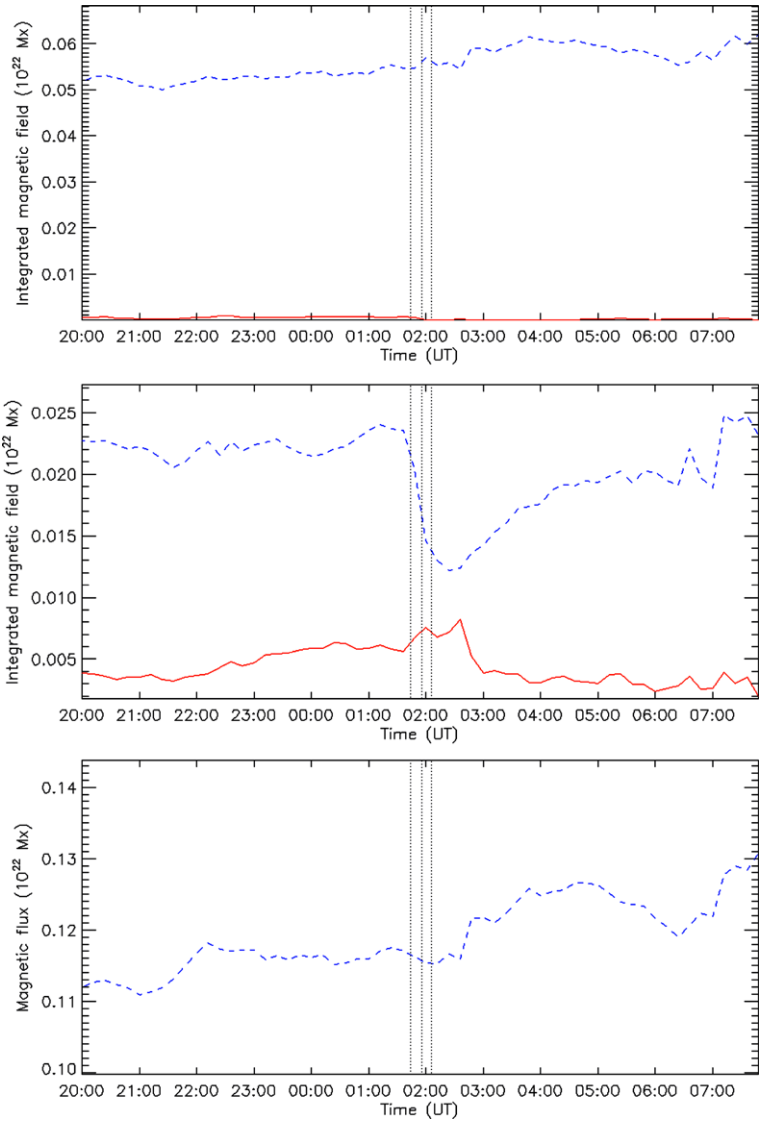


Figure 6 The radial integrated magnetic field B_R^{NS} (top), and the azimuthal, B_θ^{NS} (middle), and vertical, B_Z^{NS} (bottom) integrated field components in the inner negative sunspot are plotted as functions of time. The red/blue solid/dashed lines represent positive/negative field. The vertical lines represent the GOES flare start, peak, and end times.

an imbalance of torque at the photosphere-corona interface which would lead to a net rotation between the two photospheric footprints of the tube, reducing the coronal twist. The evolution of magnetic twist in emerging active regions observed by Pevtsov, Maleev, and Longcope (2003) was found to be in agreement with Longcope and Welsch's (2000) predictions. Magnetic helicity is not easily dissipated in the corona (Berger, 1984) and is believed to accumulate there until bodily removed by coronal mass ejections (Low, 2001). When twist

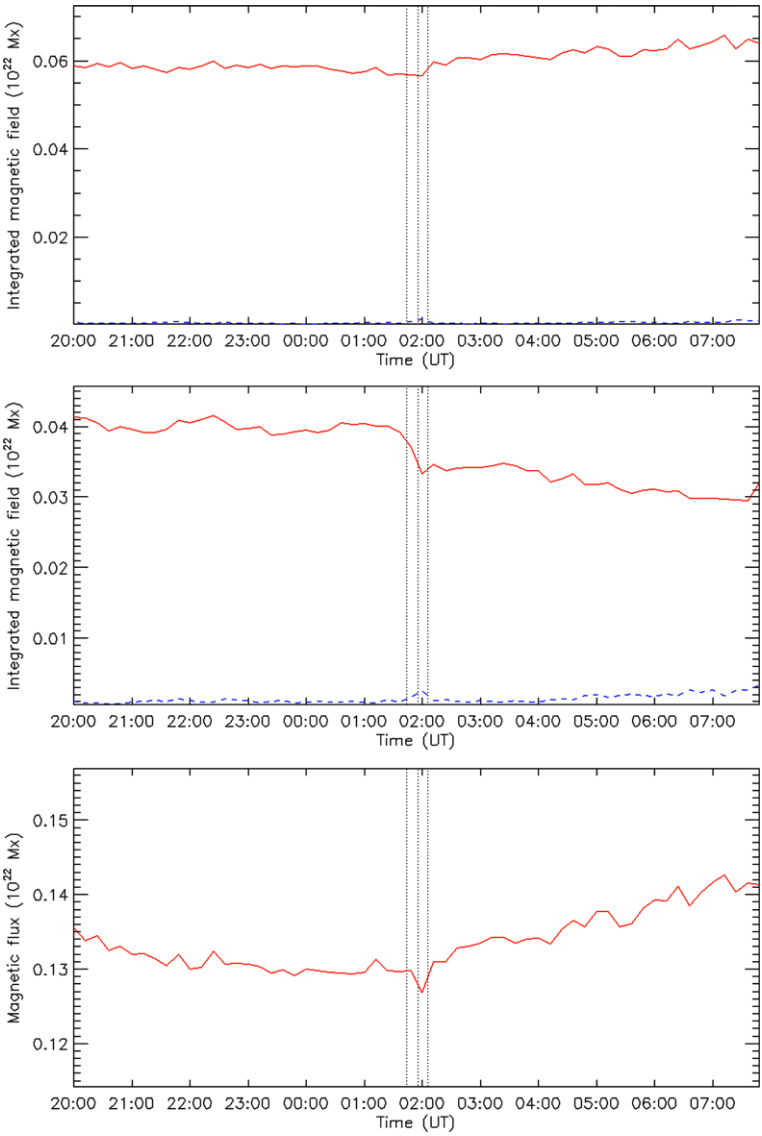


Figure 7 The radial integrated magnetic field B_R^{PS} (top), and the azimuthal, B_θ^{PS} (middle), and vertical, B_Z^{PS} (bottom) integrated field components in the inner positive sunspot are plotted as functions of time. The red/blue solid/dashed lines represent positive/negative field. The vertical lines represent the GOES flare start, peak, and end times.

is removed from coronal fields by coronal mass ejections, sub-photospheric fields could re-supply the twist until a new equilibrium is established, i.e., the rotation could be a reaction to the removal of twist from the active-region's magnetic field (Pevtsov, Maleev, and Longcope, 2003). In our data the azimuthal field decreased abruptly in both sunspots during the flare, in agreement with this theoretical picture. Subsequently, the azimuthal field in the negative spot increased steadily, returning to its pre-flare value in a few hours. The positive

spot had more azimuthal field than the negative spot throughout this series of observations, but showed no post-flare azimuthal field increase during the six hours of post-field-change observations analyzed here. It would be interesting to study the evolution of such fields over longer time intervals to see if the expected azimuthal field increases generally take place after flares.

We also computed spatial maps and temporal profiles of shear angles of the magnetic field. This is the angle between the observed horizontal photospheric field and the horizontal field of the unique potential field whose vertical component agrees with the observed vertical field distribution. This shear angle increased significantly as a result of the flare and the increases were mostly concentrated near the neutral line, following the pattern of the horizontal field changes. We do not show the plots here. The average shear angle near the neutral line as a function of time can be seen in Wang *et al.* (2012).

3. The Electric Current

Plotted in Figure 8 are spatial maps of the vertical current density J_r before and after the flare. Note the current reversal at the neutral line and the uniformity of sign on each side of the neutral line. One striking feature of Figure 8 is that in the two sunspots neighboring the neutral line the electric current is almost entirely of one polarity, positive/negative in the positive/negative sunspot. Analyzing 12 sunspots observed by *Hinode*, by the *Solar Optical Telescope/Spectro-polarimeter* instrument, Venkatakrishnan and Tiwari (2009) found that net electric currents were generally absent from their data set. Our results clearly do not fit into this pattern. Such results have been reported in the past. Examining vector magnetic field data from 21 active regions observed by the *Solar Magnetic Field Telescope* of the *Huairou Solar Observing Station* of Beijing Astronomical Observatory, Wheatland (2000) found that, while total active-region currents are well balanced, currents integrated over a given polarity of the magnetic field sum to quantities significantly different from zero, and so large-scale currents in active regions are typically unbalanced, implying that the magnetic field is not typically composed of isolated magnetic fibrils.

Comparing the two panels of Figure 8, the radial current changes do not show an obvious pattern. One small but striking change is that the two sunspots near the neutral line had small opposite-polarity concentrations of current near their centers after the flare-related field changes. These were not present before the flare.

Figure 9 shows the evolution of the integrated vertical current, J_r^{NL} near the neutral line. According to Figure 9 the negative vertical current near the neutral line decreased abruptly during the flare, although this change was not significantly greater than the background changes. The positive current continued its increasing trend through the start of the flare and began to decrease after the GOES peak time. The time of the flare marks a change from an increasing to an decreasing trend in electric current evolution near the neutral line. The total current of the region (not shown) did not change significantly at the time of the flare.

Figure 10 shows the electric currents in the positive and negative inner sunspots as functions of time. The electric currents of the two sunspots behaved quite differently. In the negative sunspot the dominant negative current abruptly decreased during and after the flare, but abruptly increased again to its pre-flare value less than an hour after the end of the flare. The dominant positive current in the positive sunspot showed no significant change during the flare but began a decreasing trend during the hours following the flare. In both sunspots the non-dominant polarity increased abruptly during the flare, because of the small opposite-polarity current concentrations appearing after the flare close to the centers of the sunspots in the lower panel of Figure 8. This feature survived long after the flare in the negative sunspot but quickly disappeared from the positive sunspot after the flare.

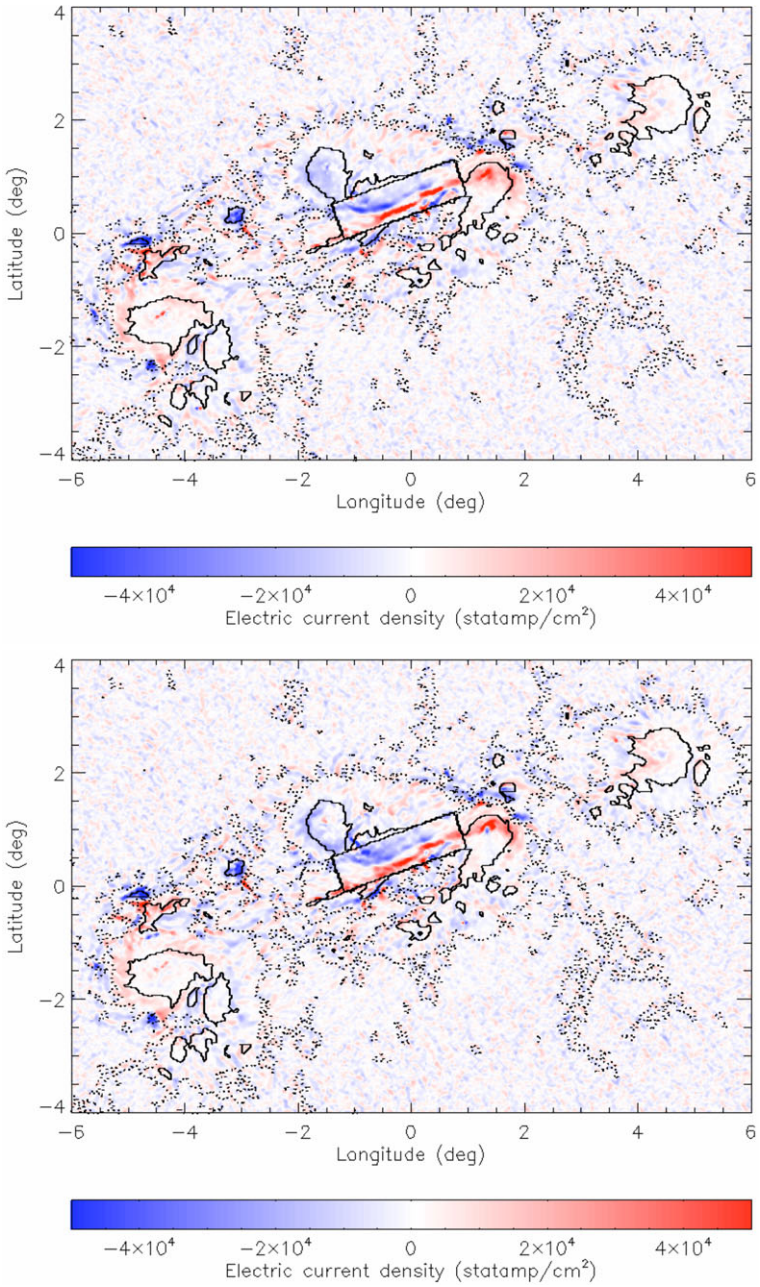


Figure 8 Shown are the vertical electric current density, J_r, before (top) and after (bottom) the main flare-related field changes. Red/blue coloring represents positive/negative vertical current with saturation values 5×10^4 statampère/cm^2. The black rectangles mark the region near the neutral line that is used in the analysis. The solid and dotted contours indicate strong ($|B_r| > 1000$ G) and quite strong ($|B_r| > 100$ G) fields, respectively.

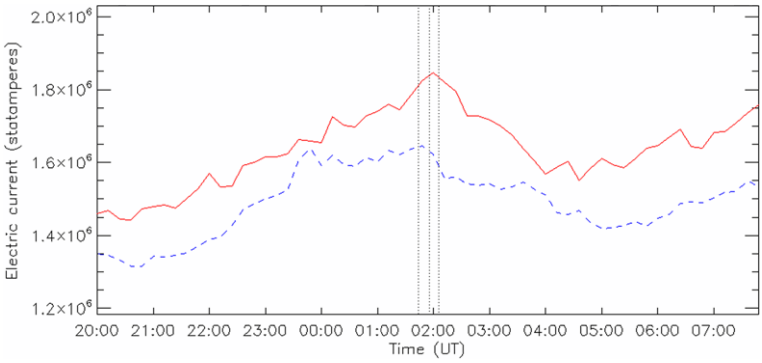


Figure 9 The vertical current density, J_r near the neutral line as a function of time is plotted here. The red/blue solid/dashed lines represent positive/negative current. The vertical lines represent the GOES flare start, peak, and end times. The area of integration for the neutral-line calculation is indicated by the black rectangle in Figure 1.

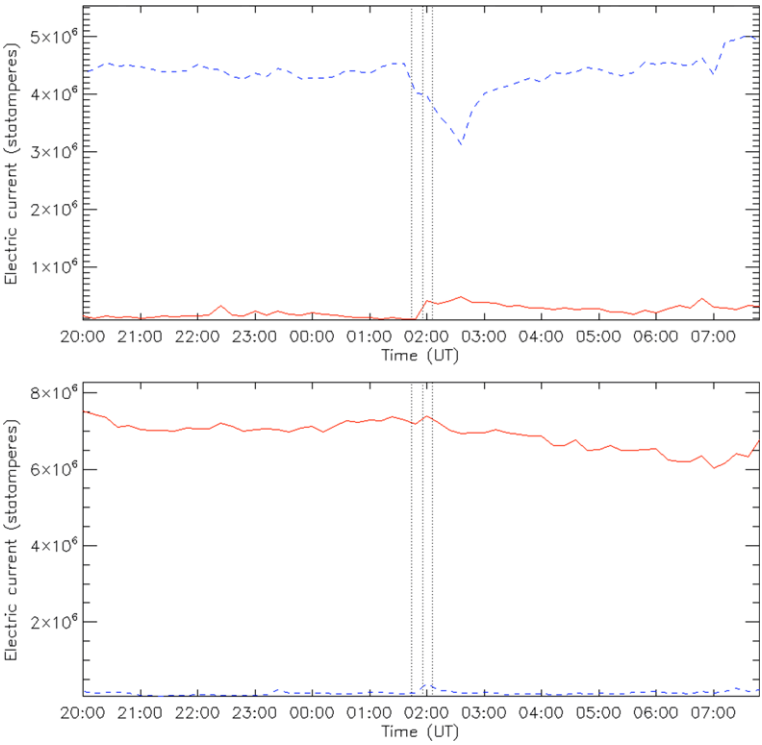


Figure 10 The vertical current density, J_z at the inner negative (top) and positive (bottom) sunspots as functions of time are plotted here. The red/blue solid/dashed lines represent positive/negative current. The vertical lines represent the GOES flare start, peak, and end times.

4. The Lorentz-Force Changes

We use the results of Fisher *et al.* (2012) to estimate the changes in the Lorentz-force vector acting on the volume below the photosphere as a result of the flare. The Lorentz force per unit volume $\mathbf{f}_L$ can be written as

$$\mathbf{f}_L = \nabla \cdot \mathbf{T} = \frac{\partial T_{ij}}{\partial x_j}, \tag{2}$$

where the Maxwell stress tensor $[T_{ij}]$ in local Cartesian heliographic coordinates is,

$$T = \frac{1}{8\pi}\begin{bmatrix} B_r^2 - B_\theta^2 - B_\phi^2 & 2B_r B_\theta & 2B_r B_\phi \\ 2B_r B_\theta & B_\theta^2 - B_r^2 - B_\phi^2 & 2B_\theta B_\phi \\ 2B_r B_\phi & 2B_\theta B_\phi & B_\phi^2 - B_r^2 - B_\theta^2 \end{bmatrix}. \tag{3}$$

Fisher *et al.* (2012) evaluated the total Lorentz force over an atmospheric volume surrounding an isolated flaring active region by integrating Equation (2) over this volume, whose lower boundary is identified with the photosphere, with upper boundary far above the photosphere, and side boundaries connecting these surfaces to form a closed volume V as shown in Figure 1 of Fisher *et al.* (2012). Evaluating the volume integral of Equation (2) using Gauss's divergence theorem then gives (Fisher *et al.*, 2012)

$$\mathbf{F}_L = \int_V \nabla \cdot \mathbf{T}\, \mathrm{d}V = \int_{A_{\rm tot}} \mathbf{T} \cdot \hat{\mathbf{n}}\, \mathrm{d}A, \tag{4}$$

where the area integral is evaluated over all surfaces of the volume, denoted by $A_{\rm tot}$, with unit normal vector $\hat{\mathbf{n}}$. As Fisher *et al.* (2012) argue, if the upper boundary of the volume is so far above the photosphere and the side boundaries are distant enough from the active region that the field integrals over these surfaces are negligible, then the surface integral of Equation (4) reduces to an integral over the photospheric lower boundary $A_{\rm ph}$ only. In this case, for the force acting on the volume below the photosphere, $\hat{\mathbf{n}} = \hat{\mathbf{r}}$ and $\mathbf{B} \cdot \hat{\mathbf{n}} = B_r$, and

$$F_r = \frac{1}{8\pi} \int_{A_{\rm ph}} \left(B_r^2 - B_{\rm h}^2\right) \mathrm{d}A, \tag{5}$$

and

$$\mathbf{F}_{\rm h} = \frac{1}{4\pi} \int_{A_{\rm ph}} (B_r \mathbf{B}_{\rm h})\, \mathrm{d}A. \tag{6}$$

Then, assuming that the photospheric vector field is observed over a photospheric area $A_{\rm ph}$ at two times, $t = 0$ before the field changes begin, and $t = \delta t$ after the main field changes have occurred, the corresponding changes in the Lorentz-force vector components between these two times are given by Equations (17) and (18) of Fisher *et al.* (2012):

$$\delta F_r = \frac{1}{8\pi} \int_{A_{\rm ph}} \left(\delta B_r^2 - \delta B_{\rm h}^2\right) \mathrm{d}A, \tag{7}$$

and

$$\delta \mathbf{F}_{\rm h} = \frac{1}{4\pi} \int_{A_{\rm ph}} \delta(B_r \mathbf{B}_{\rm h})\, \mathrm{d}A, \tag{8}$$

where at a fixed location in the photosphere

$$\delta B_{\rm h}^2 = B_{\rm h}^2(\delta t) - B_{\rm h}^2(0), \tag{9}$$

$$\delta B_r^2 = B_r^2(\delta t) - B_r^2(0), \tag{10}$$

$$\delta(B_r \mathbf{B}_{\rm h}) = B_r(\delta t)\mathbf{B}_{\rm h}(\delta t) - B_r(0)\mathbf{B}_{\rm h}(0). \tag{11}$$

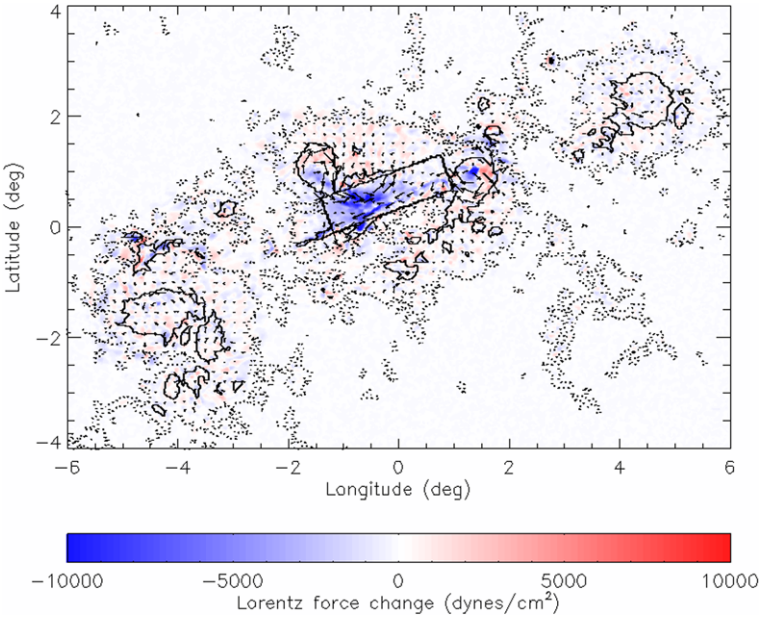


Figure 11 Shown are the Lorentz-force vector changes during the flare. The vertical component, δF_r is indicated by the color scale and the horizontal components by the arrows with saturation values 10^4 dynes/cm^2 for the color scale and 2.5×10^3 dynes/cm^2 for the arrows. Red/blue coloring represents positive/negative (upward/downward) Lorentz-force change. The black rectangle marks the region near the neutral line that is used in the analysis. The solid and dotted contours indicate strong ($|B_r| > 1000$ G) and quite strong ($|B_r| > 100$ G) fields, respectively.

The Lorentz force acting on the atmosphere above the photosphere is equal and opposite to the force acting on the volume at and below the photosphere (Fisher *et al.*, 2012).

Figure 11 shows a spatial map of the Lorentz-force changes F_r and $\mathbf{F}_\mathrm{h}$, derived by evaluating the integrals in Equations (7) and (8) pixel-by-pixel. The sums of the distributions shown in Figure 11 over the entire photospheric area gives the estimate for the total Lorentz-force vector described by Fisher *et al.* (2012). We argue that Figure 11 also gives a useful estimate of the spatial distributions of the Lorentz-force vector components across the region.

According to Equation (2) the divergence of the Maxwell stress tensor $\mathbf{T}$ gives the Lorentz-force density. Because the coronal plasma is likely to be an approximately force-free medium and the photospheric plasma is not, we expect $\mathbf{T}$ to be nearly divergence-free everywhere in the volume of integration V in Equation (4) except at the lower boundary, leading us to expect most contributions to the volume integral in Equation (4) to come from the lower boundary. We also expect the field components themselves to be stronger at the lower boundary than elsewhere in the volume because photospheric active region fields are significantly stronger than their coronal counterparts.

Evaluating the volume integral of Equation (4) over a single pixel, the volume of integration becomes a very tall, thin pillbox-shaped volume extending high into the atmosphere, whose photospheric footprint coincides with the single pixel. The contribution from the lower boundary is given by Equations (5) and (6). The contribution from the upper boundary is negligible, provided that this boundary is sufficiently high above the photosphere, but the contributions from the side boundaries need to be accounted for. For example, from

Equations (3) and (4), the contributions to F_θ from the walls with normal unit vectors $\pm\hat{\boldsymbol{\theta}}$ are

$$F_\theta^{(\pm\hat{\boldsymbol{\theta}})} = \pm\frac{1}{8\pi}\int_{A_{\pm\hat{\boldsymbol{\theta}}}} B_\theta^2 - B_r^2 - B_\phi^2 \,\mathrm{d}A. \tag{12}$$

The contributions to F_θ from the other boundaries can be readily be derived from Equations (3) and (4). These contributions involve coronal fields that are generally significantly weaker than their photospheric counterparts on the lower boundary. Furthermore, because the field distributions for major, well-resolved magnetic structures are likely to be similar on the two opposing side boundaries, situated a pixel-length apart, the contributions to Equation (4) are expected to approximately cancel each other and so their contribution can be neglected. For photospheric field structures composed of many pixels whose field patterns are well resolved, such as the region around the main neutral line and the major sunspots shown in Figure 1, the errors in the estimated Lorentz-force density changes shown in Figure 11 are expected to be small. Also the sum of these changes exactly matches the area integral described by Fisher *et al.* (2012).

In the example shown in Figure 11, the horizontal field changes $\delta\mathbf{B}_\mathrm{h}$ near the main neutral line are increases in horizontal field strength, $\delta B_\mathrm{h}^2 > 0$, and are significantly greater than the vertical field changes δB_r^2 (compare the middle and bottom panels of Figure 4). Equation (7) therefore leads us to expect the vertical Lorentz-force change to have been predominantly downward there. Figure 11 shows the spatial distribution of the change in the Lorentz-force components during the flare. As expected, near the main neutral line the Lorentz force clearly acted downwards into the photosphere. This behavior was anticipated to occur near neutral lines of flaring active regions by Hudson, Fisher, and Welsch (2008) and Fisher *et al.* (2012), and has been found in past estimates of Lorentz-force changes by Wang and Liu (2010) and Petrie and Sudol (2010). The two sunspots neighboring the neutral line appear to have undergone forces consistent with tilting motions towards the neutral line: their vertical force changes close to the neutral line were downward while those further from the neutral line were upward. Some evidence of corresponding field changes can be seen in Figure 1 where in these spots some vertical fields close to the neutral line became weaker and some far from the neutral line became stronger. However, the signature in Figure 11 is clearer.

The horizontal Lorentz-force changes also show clear patterns. Equation (8) implies that, wherever the vertical field does not change significantly compared to the horizontal changes and is positive/negative, the horizontal Lorentz-force changes $\delta\mathbf{F}_\mathrm{h}$ should be parallel/anti-parallel to the horizontal field changes $\delta\mathbf{B}_\mathrm{h}$. We already know from Figure 1 that on both sides of the neutral line $\delta\mathbf{B}_\mathrm{h}$ pointed eastward and approximately parallel to the neutral line. Figure 11 shows that the horizontal Lorentz-force change $\delta\mathbf{F}_\mathrm{h}$ acted in opposite directions along each side of the neutral line, with the changes on the southern positive side pointing eastward and those on the northern negative side westward as expected. Since the sheared field at the neutral line pointed eastward, these Lorentz-force changes are consistent with a reduction of the magnetic shear parallel to the neutral line. These horizontal Lorentz-force changes were directed against the shear flow pattern described by Beauregard, Verma, and Denker (2012). It seems that the steady shear flow pattern created, or at least strengthened, the magnetic shear whereas the horizontal Lorentz-force changes acted towards relaxing the shear. However, we know from Section 2 that the parallel field component increased significantly during the flare. The horizontal Lorentz-force pattern is a signature of the field contracting across the neutral line during the flare, tugging the photospheric fields on the two sides towards each other. The horizontal field that was added during the flare was associated

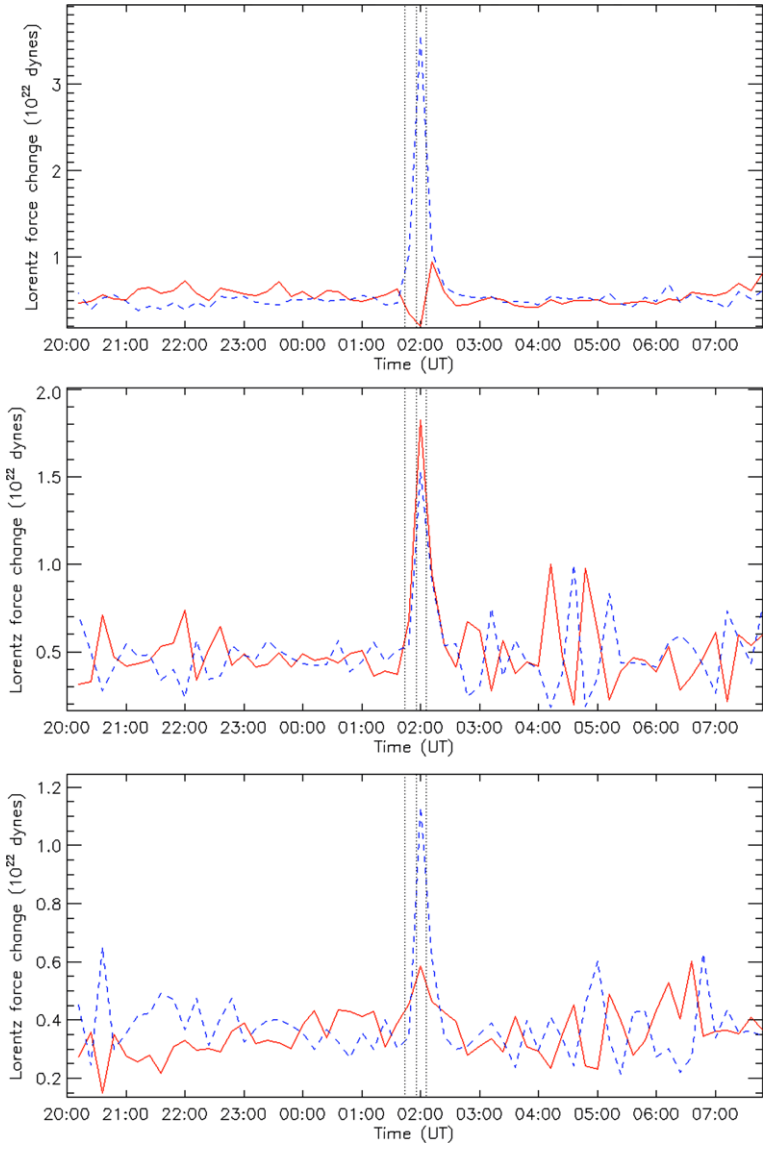


Figure 12 The Lorentz-force vector components in the vertical (δF_r^{NL}, top) and horizontal directions parallel ($\delta F_{\parallel}^{\mathrm{NL}}$, middle) and perpendicular ($\delta F_{\perp}^{\mathrm{NL}}$, bottom) to the neutral line are plotted as functions of time. The area of integration is indicated by the rectangle in Figure 1. The red/blue solid/dashed lines represent positive/negative force changes. The vertical lines represent the GOES flare start, peak, and end times.

with the large downward Lorentz-force change described above, and was caused by sheared field collapsing towards the photospheric neutral line from above. This suggests that the field near the neutral line underwent a contraction during the flare, both vertically towards the photosphere and horizontally along the neutral line. The vertical collapse had the dominant effect on the field changes near the neutral line.

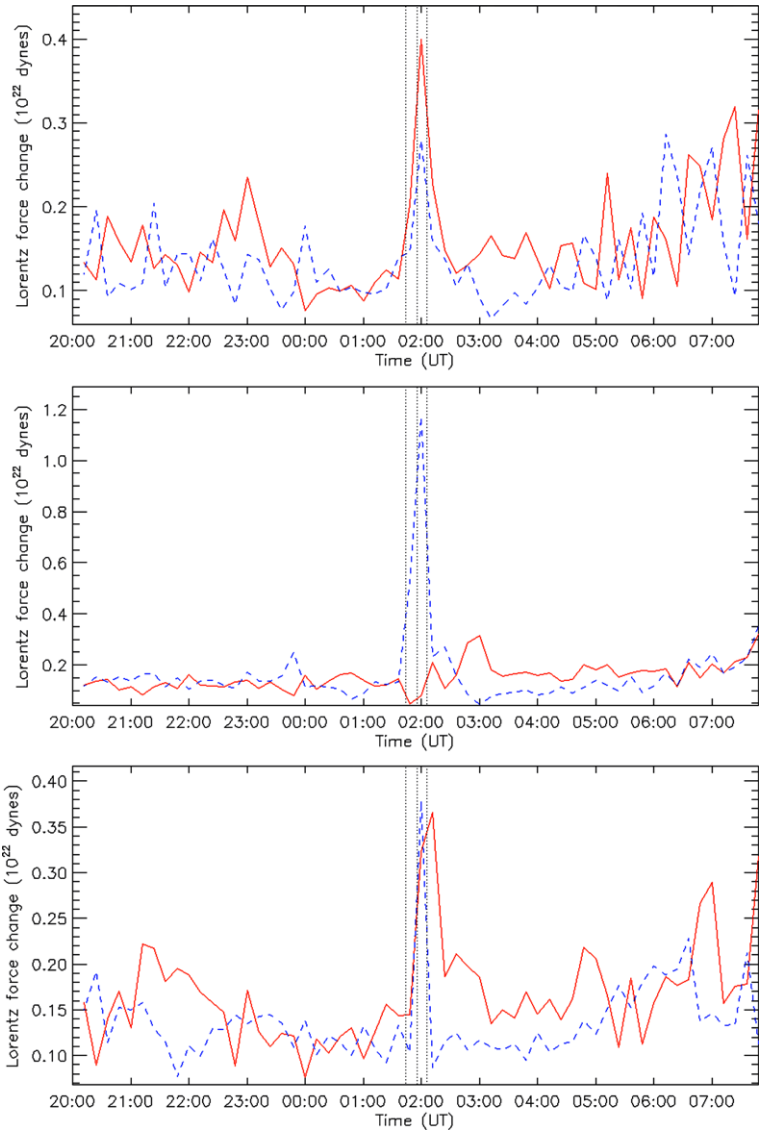


Figure 13 The Lorentz-force vector changes near the negative inner sunspot in the radial ($\delta F_R^{\rm NS}$, top), azimuthal ($\delta F_\Theta^{\rm NS}$, middle) and vertical ($\delta F_Z^{\rm NS}$, bottom) directions as functions of time. The red/blue solid/dashed lines represent positive/negative force changes. The vertical lines represent the GOES flare start, peak, and end times.

The horizontal force change patterns in the two sunspots near the neutral line are also striking. In both spots the horizontal changes had a dominant clockwise azimuthal component. These changes merged with the horizontal changes about the neutral line in acting westward on the northern side and eastward on the southern side of each spot. Because the vertical changes are smaller than the horizontal changes in Figures 6 and 7, from Equation (8) we expect the Lorentz-force change $\delta \mathbf{F}_{\rm h}$ to be parallel/anti-parallel to the horizontal

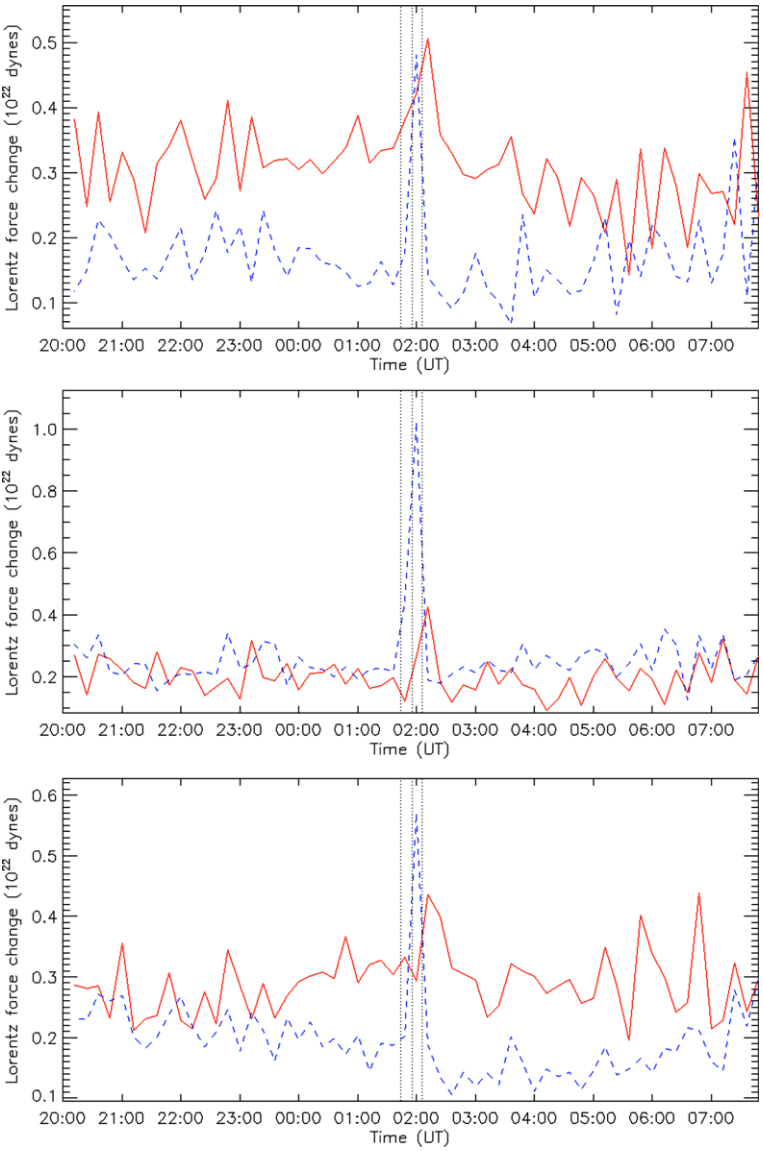


Figure 14 The Lorentz-force vector changes near the positive inner sunspot in the radial ($\delta F_R^{\rm PS}$, top), azimuthal ($\delta F_\Theta^{\rm PS}$, middle) and vertical ($\delta F_Z^{\rm pS}$, bottom) directions as functions of time. The red/blue solid/dashed lines represent positive/negative force changes. The vertical lines represent the GOES flare start, peak, and end times.

field changes $\delta\mathbf{B}_{\rm h}$ and this is indeed the case. As Figure 1 shows, the positive/negative spot showed clockwise/anti-clockwise horizontal field changes, and the horizontal Lorentz-force changes were clockwise in both spots as seen in Figure 11. Recall that the dominant azimuthal field changes were anti-clockwise in the negative spot and clockwise in the positive spot, reducing the azimuthal field in each spot. The clockwise Lorentz-force changes in both spots are signatures of the azimuthal field being abruptly removed from the spots from

above. Jiang *et al.* (2012) detected clear clockwise rotation in the sunspot proper motions and associated these rotations with the development of the positive-helicity spiral pattern of the positive spot's penumbral filaments and the shearing of the main neutral line. The horizontal Lorentz-force vector changes that we have derived from the vector-field changes are therefore directed in the same azimuthal direction as these proper motions but, being applied from above, they relaxed the field instead of twisting it.

Figure 12 shows the Lorentz-force vector changes in the vertical and horizontal directions parallel and perpendicular to the neutral line as functions of time. These plots show the Lorentz-force change components in the vertical, horizontal parallel and horizontal perpendicular directions, $\delta F_R^{\rm NL}$, $\delta F_\parallel^{\rm NL}$ and $\delta F_\perp^{\rm NL}$, integrated over the area $A_{\rm NL}$ represented by the black rectangle shown in Figure 1. The parallel direction is the direction of the long edges of the rectangle, pointing approximately west-north-west. The perpendicular direction is the direction of the short edges of the rectangle, pointing approximately north-north-east. These plots show sharp signatures of the abrupt magnetic changes during the flare in all three directions, particularly the very large downward changes and the changes in both directions, of almost equal total size, parallel to the neutral line. The sizes of these force changes, about 3.5×10^{22} dynes downward near the neutral line and almost as much in the horizontal directions, is larger than those found in the previous estimates of flare-related Lorentz-force changes by Wang and Liu (2010) and Petrie and Sudol (2010). For the 26 July 2002 M8.7 flare Wang and Liu (2010) found a downward force change of 1.6×10^{22} dynes. Petrie and Sudol (2010) found a range of longitudinal force change estimates up to about 2×10^{22} dynes. Petrie and Sudol's estimates are likely to have been underestimates because they included only information on the longitudinal field component.

Figures 13 and 14 show the Lorentz-force vector changes $\delta F_i^{\rm NS}$ and $\delta F_i^{\rm PS}$, where $i = R, \Theta, Z$, integrated over the negative and positive inner sunspot areas $A_{\rm NS}$ and $A_{\rm PS}$ in the radial, azimuthal, and vertical directions as functions of time. In both cases the main effect of the flare is a large negative azimuthal (clockwise) force change at the time of the flare, consistent with the arrows in Figure 11. Sizable force changes in other components are also evident in the plots, but these are not as large and are of mixed sign.

5. Conclusion

We have analyzed in detail 12 hours of 12-minute SDO/HMI vector-field observations covering the first X-class flare in Cycle 24, the X2.2 flare at 01:44 UT on 15 February 2011. This data set has given us the first opportunity to resolve spatial and temporal changes of field direction and strength, and their associated Lorentz-force changes, in three spatial dimensions.

The main conclusions are:

i) Near the neutral line, the photospheric field vectors became stronger and more horizontal during the flare. This was due to an increase in strength of the horizontal field components near the neutral line. The increase in strength was most significant in the horizontal component parallel to the neutral line but the component perpendicular to the neutral line also increased in strength. The result was an increase in the shear of the field near the neutral line.

ii) Perhaps surprisingly, the vertical field component did not show a significant, permanent overall change at the neutral line to compensate for the strengthened horizontal field. Instead, the increase in field at the neutral line was accompanied by a compensating field decrease in the surrounding volume. The total photospheric field of the active region did not change significantly during the flare.

iii) The two sunspots near the main neutral line also showed significant field changes. In both cases the azimuthal field abruptly decreased during the flare but this change was permanent in only one of the spots.
iv) The vertical electric current density near the main neutral line steadily increased until the time of the flare, then steadily decreased for a few hours after the flare.
v) The vertical Lorentz force had a large, abrupt downward change during the flare. This is consistent with past observations and with recent theoretical work.
vi) The horizontal Lorentz force acted in opposite directions on each side of neutral line during the flare. The two sunspots at each end of the neutral line underwent abrupt torsional Lorentz-force changes that merged with the shearing pattern of the neutral-line force changes and were consistent with the relaxation of twist. The shearing forces were consistent with a contraction of the field and a decrease of shear near the neutral line, whereas the field itself became more sheared as a result of the field collapsing towards the neutral line from the surrounding volume.

Increased magnetic field tilts during flares at neutral lines have been detected many times in the past in vector measurements (Wang and Liu, 2010; Wang *et al.*, 2012) and also in the statistics of longitudinal measurements (Petrie and Sudol, 2010). The HMI vector data have enabled us to provide spatial maps of these changes, allowing us to show that the changes of field tilt are not the result of a simple rotation of the magnetic vector towards the neutral line but a transfer of magnetic field towards the photospheric magnetic neutral line from the surrounding volume. The associated Lorentz-force changes are also not consistent with a rotation of the magnetic vector towards the neutral line. The shearing pattern of the horizontal Lorentz forces and the related azimuthal forces in the neighboring sunspots would by themselves have reduced the shear of the field near the neutral line. However, they were accompanied by a strong downward force change associated with the field collapsing downward from the surrounding volume. It was this process that was decisive in increasing the shear of the field around the neutral line. Wang (2006) found from a study of high-cadence longitudinal magnetograms that some flares produced a decrease in magnetic shear along the main neutral line, while in other cases, the shear increased. If the two patterns of Lorentz-force change, horizontal shearing patterns that act to decrease the shear of the photospheric field and downward forces that increase the shear, generally occur during flares, then this might imply that the vertical forces dominate during some flares and the horizontal forces dominate during others. Whether the collapsing field is highly sheared or not seems also to be an important factor in determining the outcome. Nonlinear force-free field extrapolations have suggested that if the photospheric field shear increases during a flare, then the increase is localized at low heights and the shear decreases above a certain height (Jing *et al.*, 2008; Sun *et al.*, 2012; Liu *et al.*, 2012). This is consistent with a sheared structure collapsing towards the neutral line, leaving a void above that is filled by more relaxed field.

In the future this work will be extended by studying many more examples and deducing which general patterns tend to dominate during flares. We are not currently capable of predicting the behavior of active regions and the occurrence of flares and ejections. Only with a comprehensive and detailed study of the governing fields will this become possible. This work has suggested that horizontal and vertical Lorentz-force changes can have different effects on the shear of magnetic neutral-line fields, with the vertical changes dominating in the case of the X2.2 flare at 01:44 UT on 15 February 2011. In Petrie and Sudol's (2010) statistical study of GONG 1-minute longitudinal magnetograms covering 77 major flares the horizontal field changes (the longitudinal field changes observed near the limb) were larger than the vertical field changes (the longitudinal field changes observed near disk-center),

were usually negative on all parts of the disk investigated, and most of the derived Lorentz-force changes, particularly the largest ones, pointed downward. These results suggest that the horizontal field changes and associated vertical Lorentz-force changes tend to be, but are not always, the more important. By studying many more high-cadence vector-field measurements from HMI and the National Solar Observatory's *Synoptic Optical Long-term Investigations of the Sun* (SOLIS) telescope, we might be able to find out what the factor is that determines which force changes are dominant, and whether this factor is related to the flare productivity of an active region. For example, if the horizontal motions are more significant, will more twist be injected from below as predicted by Longcope and Welsch (2000), and will the region tend to flare again sooner? The relationship between sunspot twist and neutral-line shear is clearly non-trivial and will be investigated further. Sunspot twist, flux emergence and cancelation, and Hudson implosion all affect the distribution of Maxwell stresses in active-region magnetic fields in different ways, and a focused, observationally driven study of the interplay of these processes will reveal much of the basic dynamics of flares.

Acknowledgements I thank the referee for helpful comments that resulted in a much improved, clearer paper. SDO is a mission for NASA's Living With a Star program. I thank Sanjay Gosain and Alexei Pevtsov for discussions. This work was supported by NSF Award No. 106205 to the National Solar Observatory.

References

Beauregard, L., Verma, M., Denker, C.: 2012, *Astron. Nachr.* **999**, 1.
Berger, M.: 1984, *Geophys. Astrophys. Fluid Dyn.* **30**, 79.
Borrero, J.M., Tomczyk, S., Kubo, M., Socas-Navarro, H., Schou, J., Couvidat, S., Bogart, R.: 2010, *Solar Phys.* **251**, 267. doi:10.1007/s11207-010-9515-6
Deng, N., Liu, C., Yang, G., Wang, H., Denker, C.: 2005, *Astrophys. J.* **623**, 1195.
Fisher, G.H., Bercik, D.J., Welsch, B.T., Hudson, H.S.: 2012, *Solar Phys.* **277**, 59. doi:10.1007/s11207-011-9907-2.
Fletcher, L., Hudson, H.S.: 2008, *Astrophys. J.* **675**, 1645.
Gosain, S.: 2012, *Astrophys. J.* **749**, 85.
Gosain, S., Venkatakrishnan, P.: 2010, *Astrophys. J.* **720**, L137.
Gosain, S., Venkatakrishnan, P., Tiwari, S.K.: 2009, *Astrophys. J.* **706**, L240.
Hudson, H.S.: 2000, *Astrophys. J. Lett.* **531**, L75.
Hudson, H.S., Fisher, G.H., Welsch, B.T.: 2008, In: *Subsurface and Atmospheric Influences in Solar Activity*, *ASP Conf. Series* **383**, 221.
Inoue, S., Kusano, K., Magara, T., Shiota, D., Yamamoto, T.T.: 2011, *Astrophys. J.* **738**, 161.
Jiang, Y., Zheng, R., Yang, J., Hong, J., Yi, B., Yang, D.: 2012, *Astrophys. J.* **744**, 50.
Jing, J., Wiegelmann, T., Suematsu, Y., Kubo, M., Wang, H.: 2008, *Astrophys. J.* **676**, L81.
Kosovichev, A.G., Zharkova, V.V.: 1999, *Solar Phys.* **190**, 459. doi:10.1023/A:1005226802279.
Kosovichev, A.G., Zharkova, V.V.: 2001, *Astrophys. J. Lett.* **550**, L105.
Leka, K.D., Barnes, G., Crouch, A.D., Metcalf, T.R., Gary, G.A., Jing, J., Liu, Y.: 2009, *Solar Phys.* **260**, 83. doi:10.1007/s11207-009-9440-8.
Li, Y.X., Jing, J. Tan, C.G., Wang, H.: 2009, *Sci. China G.* **52**, 1702.
Liu, C., Deng, N., Liu, Y., Falconer, D., Goode, P.R., Denker, C., Wang, H.: 2005, *Astrophys. J.* **622**, 722.
Liu, C., Deng, N., Liu, R., Lee, J., Wiegelmann, T., Jing, J., Xu, Y., Wang, S., Wang, H.: 2012, *Astrophys. J. Lett.* **745**, 4.
Liu, Y., Hoeksema, J., Hayashi, K., Sun, X., Schuck, P., Muglach, K.: 2011, SPD meeting #42, #21.02, *Bull. Am. Astron. Soc.* **43**, Am. Astron. Soc.
Longcope, D.W., Welsch, B.T.: 2000, *Astrophys. J.* **545**, 1089.
Low, B.C.: 2001, *J. Geophys. Res.* **106**, 25141.
Metcalf, T.R.: 1994, *Solar Phys.* **155**, 235. doi:10.1007/BF00680593.
Pesnell, W.D., Thompson, B.T., Chamberlin, P.C.: 2012, *Solar Phys.* **275**, 3. doi:10.1007/s11207-011-9841-3.
Petrie, G.J.D., Sudol, J.J.: 2010, *Astrophys. J.* **724**, 1218.
Pevtsov, A.A., Maleev, V.M., Longcope, D.W.: 2003, *Astrophys. J.* **593**, 1217.

Ravindra, B., Venkatakrishnan, P., Tiwari, S., Bhattacharyya, R.: 2011, *Astrophys. J.* **740**, 19.
Schou, J., Scherrer, P.H., Bush, R.I., Wachter, R., Couvidat, S., Rabello-Soares, M.C., *et al.*: 2012, *Solar Phys.* **275**, 229. doi:10.1007/s11207-011-9842-2.
Schrijver, C.J., Aulanier, G., Title, A.M., Pariat, E., Delannée, C.: 2011, *Astrophys. J.* **738**, 167.
Severny, A.B.: 1964, *Annu. Rev. Astron. Astrophys.* **2**, 363.
Sudol, J.J., Harvey, J.W.: 2005, *Astrophys. J.* **635**, 647.
Sun, X., Hoeksema, T., Liu, Y., Wiegelmann, T., Hayashi, K., Chen, Q., Thalmann, J.: 2012, *Astrophys. J.* **748**, 77.
Venkatakrishnan, P., Tiwari, S.K.: 2009, *Astrophys. J. Lett.* **706**, L114.
Wang, H.: 2006, *Astrophys. J.* **649**, 490.
Wang, H., Ewell, M.W., Zirin, H., Ai, G.: 1994, *Astrophys. J.* **424**, 436.
Wang, H., Liu, C.: 2010, *Astrophys. J. Lett.* **716**, L195.
Wang, H., Liu, C., Deng, Y., Zhang, H.: 2005, *Astrophys. J.* **627**, 1031.
Wang, H., Qiu, J., Jing, J., Spirock, T.J., Yurchyshyn, V.: 2004, *Astrophys. J.* **605**, 931.
Wang, H., Varsik, J., Zirin, H., Canfield, R.C., Leka, K.D., Wang, J.: 1992, *Solar Phys.* **142**, 11. doi:10.1007/BF00156630.
Wang, J., Zhao, M., Zhou, G.: 2009, *Astrophys. J.* **690**, 862.
Wang, S., Liu, C., Liu, R., Deng, N., Liu, Y., Wang, H.: 2012, *Astrophys. J.* **745**, L17.
Wheatland, M.S.: 2000, *Astrophys. J.* **532**, 674.
Zirin, H., Tanaka, K.: 1981, *Astrophys. J.* **250**, 791.

DOI 10.1007/978-1-4899-8005-2_26
Reprinted from *Solar Physics* Journal, DOI 10.1007/s11207-012-0088-4

The Wave–Driver System of the Off-Disk Coronal Wave of 17 January 2010

M. Temmer · B. Vrsnak · A.M. Veronig

Received: 13 February 2012 / Accepted: 25 July 2012 / Published online: 31 August 2012

Abstract We study the 17 January 2010 flare–CME–wave event by using STEREO/SECCHI-EUVI and -COR1 data. The observational study is combined with an analytic model that simulates the evolution of the coronal wave phenomenon associated with the event. From EUV observations, the wave signature appears to be dome shaped having a component propagating on the solar surface ($\overline{v} \approx 280$ km s^{-1}) as well as one off-disk ($\overline{v} \approx 600$ km s^{-1}) away from the Sun. The off-disk dome of the wave consists of two enhancements in intensity, which conjointly develop and can be followed up to white-light coronagraph images. Applying an analytic model, we derive that these intensity variations belong to a wave–driver system with a weakly shocked wave, initially driven by expanding loops, which are indicative of the early evolution phase of the accompanying CME. We obtain the shock standoff distance between wave and driver from observations as well as from model results. The shock standoff distance close to the Sun ($< 0.3\ R_{\odot}$ above the solar surface) is found to rapidly increase with values of $\approx 0.03-0.09\ R_{\odot}$, which gives evidence of an initial lateral (over)expansion of the CME. The kinematical evolution of the on-disk wave could be modeled using input parameters that require a more impulsive driver (duration $t = 90$ s, acceleration $a = 1.7$ km s^{-2}) compared to the off-disk component (duration $t = 340$ s, acceleration $a = 1.5$ km s^{-2}).

Keywords Shock waves · Coronal mass ejection

Solar Dynamics and Magnetism from the Interior to the Atmosphere
Guest Editors: R. Komm, A. Kosovichev, D. Longcope, and N. Mansour

Electronic supplementary material The online version of this article (doi:10.1007/s11207-012-0088-4) contains supplementary material, which is available to authorized users.

M. Temmer (✉) · A.M. Veronig
Kanzelhöhe Observatory/IGAM, Institute of Physics, University of Graz, Universitätsplatz 5, 8010 Graz, Austria
e-mail: manuela.temmer@uni-graz.at

B. Vrsnak
Hvar Observatory, Faculty of Geodesy, University of Zagreb, Kačićeva 26, 10000 Zagreb, Croatia

1. Introduction

In large part, our knowledge of coronal mass ejections (CMEs) comes from coronagraph observations delivering white-light data. CMEs, as observed in white light, often exhibit a typical three-part structure, consisting of a bright rim encircling a dark cavity, and mostly followed by a bright core (Illing and Hundhausen, 1985). Therefore, by definition, a CME is a structured intensity enhancement observed in white light. The actual process that launches the ejection is probably connected to magnetic restructuring. This early evolution phase of a CME can often be observed in the extreme ultraviolet (EUV) as well as the soft X-ray regime in the form of expanding loop systems (*e.g.* Harrison and Lyons, 2000; Vršnak *et al.*, 2004).

CMEs, as they evolve and propagate away from the Sun, are able to drive magnetohydrodynamical (MHD) shocks in the corona that can be tracked by coronal type II radio bursts (Gopalswamy *et al.*, 1997; Magdalenić *et al.*, 2010). The formation of the shock itself is dependent on the time–speed profile of the CME as well as on the spatial distribution of the Alfvén speed in the solar corona, which in turn is related to the local magnetic field strength and density of the ambient plasma. To generate a shock, the CME must have a sufficiently high velocity with respect to the local Alfvén speed; such favorable conditions are assumed to be present in the middle corona over $\approx 2\ R_{\odot}$ (*e.g.* Gopalswamy *et al.*, 2001; Mann *et al.*, 2003). Recent studies showed that shocks driven by fast CMEs are observable in white-light data (*e.g.* Vourlidas *et al.*, 2003; Ontiveros and Vourlidas, 2009; Bemporad and Mancuso, 2010; Kim *et al.*, 2012) as well as EUV (*e.g.* Veronig *et al.*, 2010; Kozarev *et al.*, 2011; Ma *et al.*, 2011; Gopalswamy and Yashiro, 2011; Cheng *et al.*, 2012), and UV spectra (*e.g.* Raymond *et al.*, 2000; Bemporad and Mancuso, 2010).

The evolution of a three-dimensional (3D) dome connected to a surface shock wave is observed for the 17 January 2010 CME–flare event. It was studied in detail by Veronig *et al.* (2010), who showed that the surface as well as the off-limb structure are part of an evolving 3D wave dome formed by a weak shock. The surface wave propagated with a mean speed of $\approx 280\ \mathrm{km\,s^{-1}}$, whereas the upward moving part was of much higher speed of $\approx 650\ \mathrm{km\,s^{-1}}$ (Veronig *et al.*, 2010). The difference between the speed of the upward moving part of the wave and the on-disk signature was interpreted by Veronig *et al.* (2010) in the following manner: the upward moving part is driven all of the time by the outward moving CME, whereas the surface signature is only temporarily driven by the flanks of the expanding CME and then propagates freely.

A recent article by Grechnev *et al.* (2011), studying the same event, supports the result that the dome structure was actually a shock-driven plasma flow. Grechnev *et al.* (2011) simulated the evolution of the shock wave from which they concluded that most likely an abrupt eruption of a filament caused the weak shock. They compare this with a blast-wave scenario during which the wave is only briefly driven. Zhao *et al.* (2011) investigated, for the 17 January 2010 event, the relation between the surface wave speed, the CME speed, and the local fast-mode characteristic speed. They concluded that the observed CME front is in fact a wave phenomenon just like the EUV wave on the solar surface.

In this study, we focus on the kinematical evolution of the off-disk signature of the dome-shaped wave event and add new aspects not covered by previous studies. Using observations of the SECCHI instruments *Extreme Ultraviolet Imager* (EUVI) and COR1 on the *Solar Terrestrial Relations Observatory* (STEREO), we will show, by applying an analytical model with input parameters constrained by the observations, that the off-disk signature in fact consists of two components: a driver and a weakly shocked wave. The driver of the off-limb wave evolves from expanding loop structures and is interpreted as the CME; the observed

frontal part is interpreted as the shock wave ahead. In particular, we investigate the shock offset (standoff) distance for the wave–driver system.

2. Data

The EUVI instrument (Wuelser *et al.*, 2004) and the COR1 and COR2 coronagraphs are part of the *Sun Earth Connection Coronal and Heliospheric Investigation* (SECCHI: Howard *et al.*, 2008) instrument suite onboard the STEREO mission (Kaiser *et al.*, 2008), launched in October 2006. On 17 January 2010 STEREO-B/EUVI observed on the eastern limb a flare/CME event, associated with a dome-shaped structure that can be observed off-limb as well as on-disk. The upward moving dome is well observed in white-light coronagraph STEREO-B/COR1 and COR2 data. In the following study we use EUVI 171 Å and 195 Å filtergrams with a temporal cadence of 60 and 90 seconds, respectively, as well as COR1 white-light data with a cadence of five minutes. Using these instruments we can follow the event in EUVI out to 1.7 $R_{\odot}$ and in COR1 over the field of view (FOV) of 1.4 to 4 $R_{\odot}$. Accordingly, we focus on the low coronal signatures of the flare/CME event and its early evolution phase. We note that associated with the event is a high-frequency type II burst drifting from $\approx$310 MHz to $\approx$80 MHz during $\approx$03:51 – 03:58 UT. The source region of the event under study is active region AR 11041 located at S25E128, *i.e.* occulted as viewed from Earth.

In order to derive the kinematical evolution of the off-disk wave, we follow the dome-shaped structure. For this we use i) manual tracking of intensity enhancements and ii) perturbation profiles over the dome structure. The perturbation profiles are defined as intensity variations averaged over angular sectors along the propagation direction of the dome. Since the off-limb wave is not evolving radially from the solar surface (see also Grechnev *et al.*, 2011), the kinematical profiles are derived along the direction of motion which is -12° off the radial direction.

3. Analytical Model

Under the assumption that the coronal wave under study is a large-amplitude MHD wave, we show that the observed dome belongs to a wave–driver system. We simulate the observed kinematical profile of the dome structure by using the analytical model developed by Temmer *et al.* (2009), which we briefly describe in the following paragraphs.

In the model, the driver of the wave is a "synthetic" source surface which continuously emits MHD signals at time steps of Δt = ten seconds. The signals start to be emitted at t_0 and are iteratively followed at each time step $\Delta t = t_i - t_{i-1}$ during their evolution until the time t_i. We obtain the distance from the source region center $[d(t_i) = r(t_0) + x(t_i)]$ where $r(t_0)$ is the radius of the source surface at the time t_0 and $x(t_i)$ is the distance traveled by the signal from t_0 until t_i. The geometry of the driver is spherically symmetric and radially expanding with a radius $r(t)$ centered at height $h(t)$. The beginning of the shock formation in the model is determined when a later emitted signal overtakes the outermost one (for more details see Vršnak and Lulić, 2000).

For the driver of the off-disk wave we consider a source that may expand and move at the same time with a constant radius-to-height ratio $[r(t)/h(t)]$ acting as a combined bow shock–piston driver. What we call a bow shock is a shock wave that moves with the same speed as the driver; material can flow behind the driver. A piston-driven shock continuously compresses the wave ahead, leading to an increase in the shock standoff distance

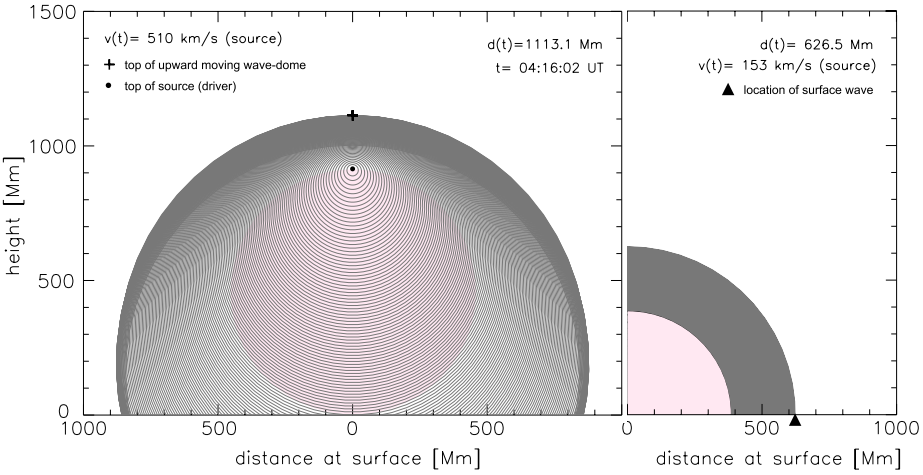


Figure 1 Snapshot of a model run for 04:16:02 UT. Gray circles indicate the propagation of the wave signals that the outermost part of the driver has sent every ten seconds starting from 03:48:32 UT. The driver velocity $[v]$ is given by the distance of the outermost part of the driver at time t with $d(t)$ (*cf.* Figure 5). The pink-shaded circle gives the size of the source (driver). Left: Applying the parameter set which simulates the upward moving dome of the wave, *i.e.*, the off-disk wave. Right: Applying the parameter set which simulates the on-disk wave.

and in the speed of the wave (for more details on the terminology see Vršnak, 2005). For completeness, we also simulate the on-disk surface wave for which we use a synthetic source expanding only in the lateral direction without upward motion; *i.e.*, plasma cannot flow behind the contact surface and the driver acts as a piston (Warmuth, 2007; Žic *et al.*, 2008). For the one-dimensional case, Vršnak and Lulić (2000) developed a simplified relation between the propagation speed of the surrounding plasma and the amplitude of the wave. From this it follows that the rest-frame speed of the wave signals $[w]$ is related to the flow velocity $[u]$ which is associated with the perturbation amplitude, and the local Alfvén velocity $[v_{A0}]$ as $w = v_{A0} + 3u/2$. Since we do not know the spatial distribution of v_{A0} in the corona we simply express the change of v_{A0} with distance using an exponential function (see Equation (2) in Temmer *et al.*, 2009). In this way, the exponential function regulates the decay of the wave signals. For more details on the model we refer the reader to Temmer *et al.* (2009).

On-disk coronal waves are assumed to be driven impulsively over a short time and then to propagate freely (*e.g.* Vršnak and Cliver, 2008); whereas, due to the upward movement of the CME, a separate mechanism acts on off-disk waves. Therefore, to simulate the on- and off-disk wave we will apply two different expansion mechanisms of the driver, which enables us to derive their physical characteristics separately.

Figure 1 gives a snapshot of a model run showing the wave signals (circles) that were emitted during the expansion of the source. The left panel of Figure 1 considers the simulation of the off-disk wave and shows the emitted signals for an upward moving (along the y-axis) and simultaneously expanding source. The snapshot presents the time step 04:16:02 UT at which the frontal part of the spherical source has a height of $h(t) = 914.8$ Mm and a radius of $\approx$450 Mm. This is comparable to observational results, as displayed in Figure 2. The right panel of Figure 1 shows the emitted signals for a source expanding in the lateral direction (along the x-axis), simulating the evolving flanks of the CME and the on-disk wave. From the model results we extract the kinematics of the solar

surface signal, *i.e.* the EUV wave, and the kinematics of the summit of the off-disk wave. We stress that the input parameters for the driver as well as the extracted kinematics of the simulated wave are constrained by observational results. Different model runs are performed until a best match is found between the model input/results and the observational results.

4. Results

Figure 2 presents composite images from EUVI 195 Å and COR1, which show the evolution of the surface wave as well as the dome of the wave moving outward from the Sun in an almost radial direction. The dome of the wave can be seamlessly connected to the surface wave, which supports the idea that it is part of the on-disk wave and not the frontal part of the erupting structure (see also Veronig *et al.*, 2010). In addition, a secondary intensity enhancement is observed behind the top part of the wave, which can be interpreted as the driver of the wave. The top part of the wave and the structure behind evolve and expand concurrently, since no black/white feature is visible in the running-difference image. This supports the idea that the top part of the wave and the secondary intensity enhancement belong together and form a wave–driver system. We note that the two components cannot be resolved beyond 2 $R_{\odot}$, which restricts our analysis to the early-evolution phase of the wave–driver system.

Figure 3 shows details of the active region as observed with EUVI-B 195 Å showing the evolution of different loop systems. In total we identify three loop systems (marked with red arrows in the top left panel) evolving in different directions from the active region: one to the northern direction, one radially away from the Sun, and one off-disk directed to the South. We note that the northern loop structure observed at 03:51:32 UT looks similar to a wave signature but may actually be a loop system pushed down by the lateral expansion of the central eruption (Patsourakos, Vourlidas, and Stenborg, 2010). At 03:52:47 UT a circularly shaped signature, presenting the coronal wave, appears (marked with a yellow line). In addition, some internal structures visible as intensity enhancements (black dashed line) are observed behind the wave front and ahead of the expanding loops, most probably resulting from compressed plasma. The on-disk wave evolves from the northern loop system, and the off-disk signature of the wave becomes visible at $\approx 0.28\ R_{\odot}$ above the solar surface ahead of the radially expanding loop structure. From the southern loop system less distinct coronal wave signatures evolve. We note that the surface coronal wave could be observed with highest intensity in the northward direction (Veronig *et al.*, 2010). During the early evolution of the off-disk wave, a clear spatial gap between the two intensity enhancements is observed which is increasing with time (see also Figure 2). We interpret the first intensity enhancement as the top part of the wave dome, the secondary intensity enhancement as the leading edge of the driver, *i.e.* the CME, and the spatial gap between as the shock standoff distance. This gives further evidence for a wave–driver system.

From base-ratio EUVI-B 171 Å images (Figure 4), we derive profiles of changes in the intensity relative to a pre-event image (03:40 UT). The intensity profiles are calculated by averaging the intensity variation of the image over an angle of 5° above the solar surface along their direction of motion. The left panel of Figure 4 shows a ratio image together with the region (yellow lines) over which the mean brightness is obtained. The derived averaged intensity is given as a function of distance above the solar surface in the right panel of Figure 4. This clearly shows a spatial gap in the brightness of the dome-shaped wave structure, which can be followed during three time steps. At 03:55 UT the wave and driver components first appear in the profile, having a similar relative intensity enhancement

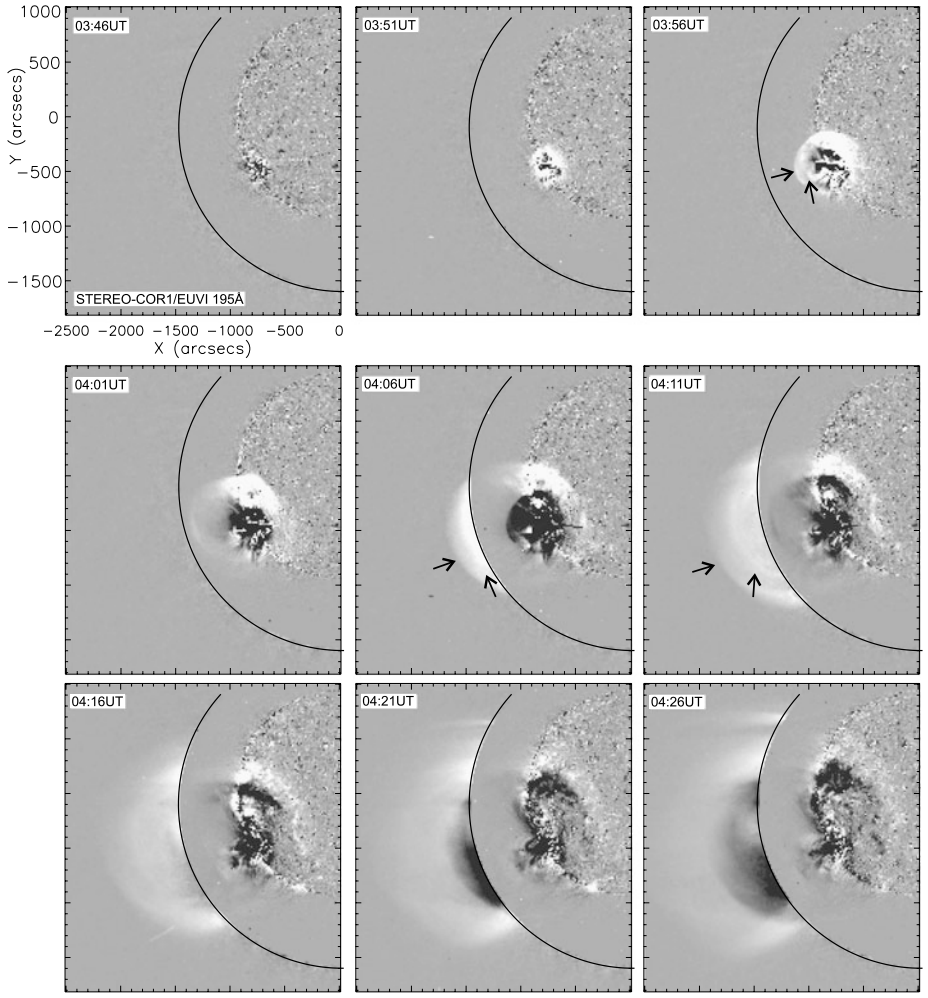


Figure 2 Composite EUVI 195 Å and COR1 observations from STEREO-B. The running-difference images show the evolution of the coronal wave as it propagates off-limb as well as a separate evolving structure below (black arrows). See Electronic Supplementary Material for the accompanying movie.

of ≈7 % above background level. The profiles of time steps 03:56 UT and 03:57 UT reveal that, relative to the driver component, the wave gains in intensity and the distance between them increases. This can be interpreted as compression of plasma ahead of the driver and steepening of the wave front, which propagates faster than the driver.

The ratio images become very noisy further out than $\approx 1.3\ R_{\odot}$ (*cf.* Figure 4). By using running-difference images we derive the distance–time profile of the dome wave over the entire FOV of EUVI-B 195 Å and 171 Å and beyond that of the COR1 data. We manually measure the top part of the dome as well as the secondary intensity structure along their propagation of motion. The derived kinematics of wave–driver of the off-disk wave is shown in Figure 5. The observational results of the solar surface wave are taken from the study by Veronig *et al.* (2010) and are calculated as the mean distance of the wavefronts from the derived wave center along great circles on the solar surface. Figure 5 also presents the

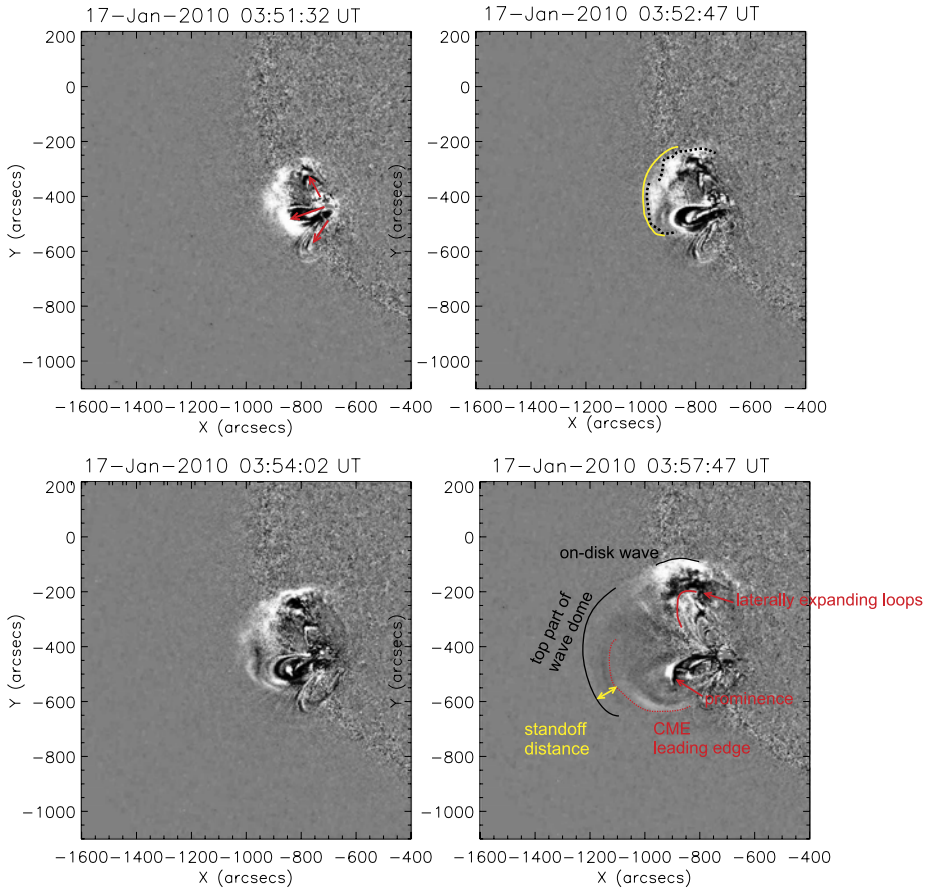


Figure 3 EUVI-B 195 Å running-difference images showing in detail the evolution of different loop systems and features (labeled) associated with the event.

simulated kinematics for the off-disk wave and its driver as well as for the on-disk wave, choosing model parameters constrained by observations. The first wave signals are emitted by the synthetic source at $t_0 = 03{:}48{:}32$ UT, which is a few minutes before the first wave signature could be identified (see Veronig *et al.*, 2010). We note that a type II radio burst appeared at 03:51 UT, indicating that the wave had to be launched well before (see also Grechnev *et al.*, 2011).

To simulate the kinematics of the upward moving off-disk wave together with its driver that best matches the observations we use a synthetic driver which accelerates over a time span of 340 seconds with $a = 1.5\ \mathrm{km\,s^{-2}}$, giving a final velocity of $v = 510\ \mathrm{km\,s^{-1}}$. The mean speed of the resulting wave measured $-12°$ off the radial direction is $\approx 600\ \mathrm{km\,s^{-1}}$ (*cf.* Grechnev *et al.*, 2011). In addition, the source size is set to be proportional to height at each time $[t]$ with $r(t)/h(t) = 0.1$. The surrounding Alfvén speed of the unperturbed plasma is chosen as $v_{A0} = 500\ \mathrm{km\,s^{-1}}$ (see Mann *et al.*, 2003). This type of source expansion acts as a combined bow shock/piston driver for the emitted signals. The decay of the signal follows an exponential function and is set, according to the best match between observational and model results, with a decay length of 220 Mm. The distance of the front of the wave minus

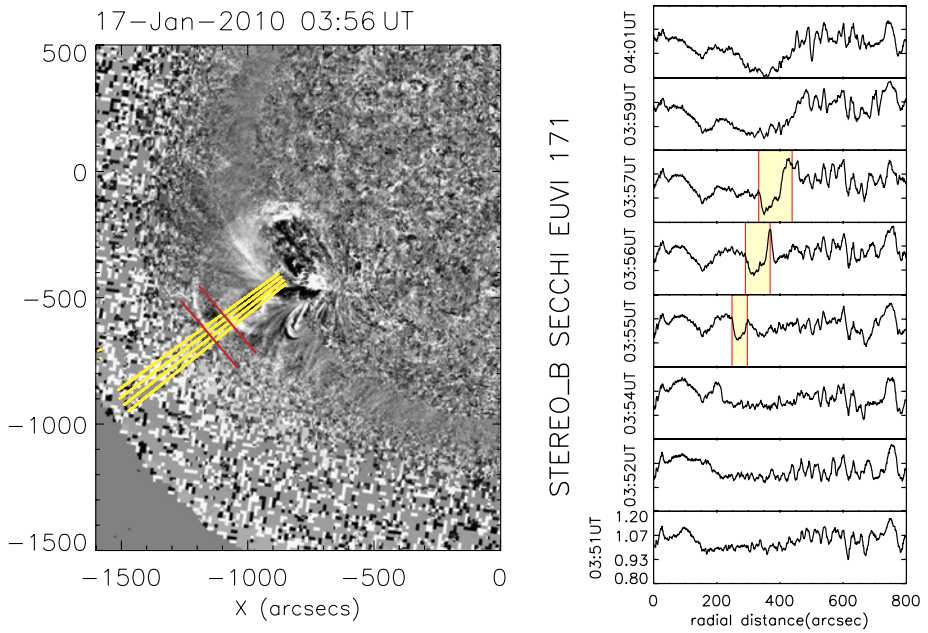


Figure 4 Left: EUVI-B 171 Å ratio image (03:56 UT – 03:40 UT) and intensity profiles of the off-limb structure measured along its propagation direction. Yellow lines mark a region of 5° over which the mean brightness is calculated; red lines mark the identified driver and wave components. Right: Profiles derived for different time steps showing changes of intensity relative to a pre-event image given as a function of distance above the solar surface. The standoff distance is marked with a yellow-shaded area; red lines mark the identified driver and wave components.

the top part of the source is defined as the standoff distance. The timing of shock formation is about 03:51:42 UT, which is close to the occurrence of the type II burst; after that the wave is freely propagating.

In order to mimic the surface signal (EUV wave) the source of the surface wave expands in a lateral direction and is fixed at the surface, which can be interpreted as a piston. The kinematics for this source is a synthetic profile accelerating over a time span of 90 seconds with an acceleration of $a = 1.7$ km s^{-2}, giving a final velocity of the driver of $v = 153$ km s^{-1}. The surrounding Alfvén speed of the unperturbed plasma is chosen as $v_{A0} = 200$ km s^{-1} (see Mann *et al.*, 2003). The decay length of the wave signals of 500 Mm is chosen to be consistent with observations given in Veronig *et al.* (2010), where the intensity profile strongly decreases in the range of 500 – 800 Mm. The timing of shock formation for the solar surface wave is about 03:50:02 UT, after which the wave is freely propagating.

Figure 6 shows the shock standoff distance between the driver and wave components for the dome-shaped structure. A good match is found between the standoff distance derived by manually tracking the wave–driver system and extracted from intensity profiles. The observed standoff distance shows a rather linear evolution up to 1 $R_\odot$. Beyond this distance, COR1-B observations (the last four data points) indicate a decreasing growth rate of the standoff distance, *i.e.* a certain "stagnation" of the growth.

The model results for the off-disk wave, presented by the black solid line in Figure 6, show a rather sharp increase of the standoff distance at heights below 0.3 $R_\odot$. Beyond this height, a nonlinear regime starts, characterized by a stagnation of the growth

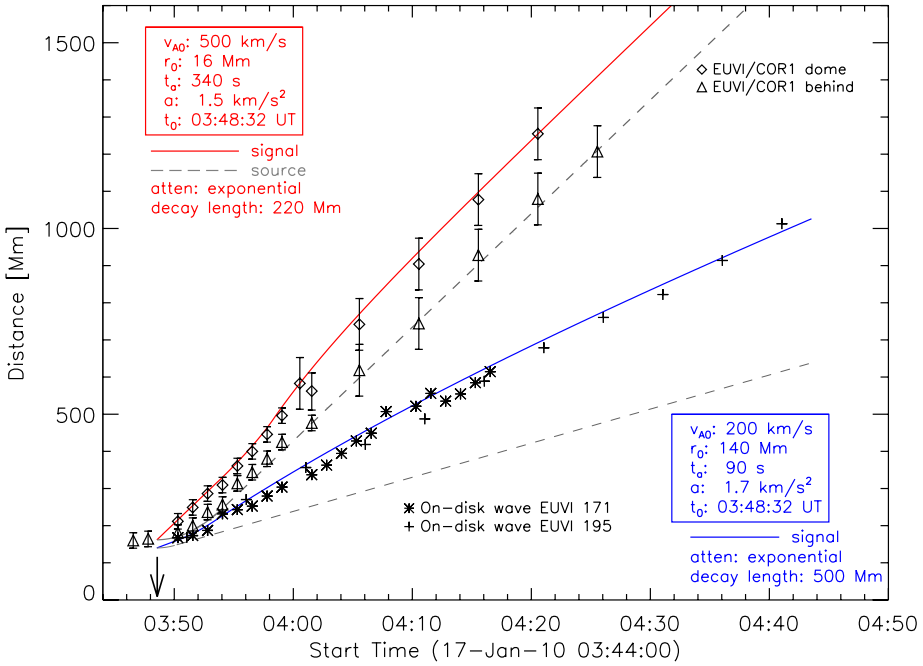


Figure 5 Distance–time plot for the on-disk wave, the top part of the wave dome, and the feature behind the wave. Overplotted is the outcome of model runs, which simulate the propagation of a wave, applying the parameters given in the legend. Red indicates the upward moving dome of the wave, dashed gray its driver, *i.e.* the feature behind the wave dome. In blue we present the EUV wave propagating at the solar surface and its driver as a dashed gray line.

rate, quite similar to that found from COR1-B measurements. This behavior can be attributed to the way that the source surface, from which the wave signals are emitted, behaves. The source moves upward and expands at the same time with $r(t)/h(t)$; hence the expansion is coupled to the kinematical characteristics of the synthetic source. The standoff distance therefore reflects the kinematical profile of the synthetic source. The stagnation beyond $\approx 1.2\ R_{\odot}$ results from the assumed decay of the wave, calculated by an exponential function with a decay length of 220 Mm. This can be interpreted as saying that the initial lateral (over)expansion of the CME (Patsourakos, Vourlidas, and Stenborg, 2010) acts as a piston close to the Sun, which results in an increase in the standoff distance since plasma material cannot flow behind the driver (Warmuth, 2007; Žic *et al.*, 2008). As the wave–driver system further evolves, it becomes more of a piston–bow shock type and the increase in distance between driver and wave is less strong. The standoff distance derived for the on-disk wave, presented as a gray line in Figure 6, shows a steep linear increase over distance. This reflects the (3D) piston mechanism of the laterally expanding source, which impulsively drives the wave, and as the strong expansion of the driver stops, the wave continuously separates from the driver.

Figure 7 shows an estimation of the shock standoff distance at 1 AU for which we simply extend, by using linear fits (assuming self-similar expansion of the CME), the results derived close to the Sun up to the Earth's location. For the model we use a linear fit to results obtained over the distance range $8-10\ R_{\odot}$. We apply a linear fit to all measured data points and, to consider the stagnation of the growth rate in standoff distance at larger distance, a linear fit

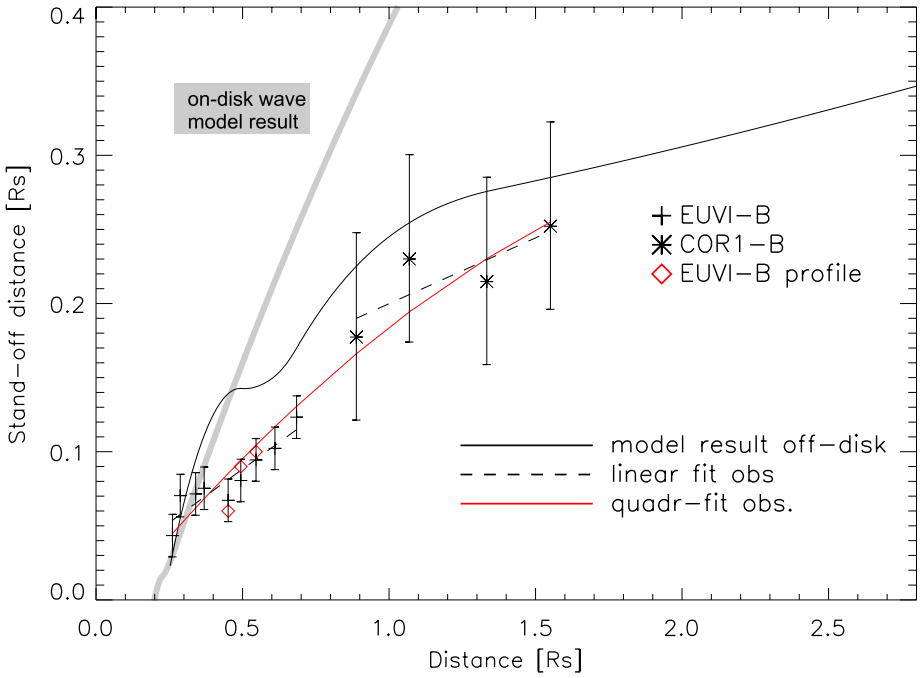


Figure 6 Standoff distance between the driver and off-disk wave *versus* propagation distance of the top part of the dome structure (starting from solar surface) derived from EUVI-B and COR1-B observations and the model (solid line). The dashed line gives a linear fit to the observational results separately performed on EUVI-B and COR1-B data, the red line a quadratic fit to all observational results. For completeness we show – as a gray line – the model results for the standoff distance of the on-disk wave–driver system *versus* distance of the wave front measured from the initiation location.

to COR1-B observations only (last four data points). From this we obtain a lower and an upper limit for the standoff distance at 1 AU lying in the range of $\approx 20-36\ R_{\odot}$. We note that the standoff distance derived from the model is very close to the linearly extrapolated standoff distance for COR1-B data points. Considering the wave speed of $v = 600\ \mathrm{km\,s^{-1}}$, this corresponds to a shock-CME time lag of the order of $6-11$ hours at 1 AU. This is consistent with reports of the thickness of the magnetosheath of ICMEs measured from *in-situ* data typically of the order of 0.1 AU (*e.g.* Russell and Mulligan, 2002). In a recent investigation, Maloney and Gallagher (2011) find a value of 20 $R_{\odot}$ for the shock standoff distance at 0.5 AU.

The standoff distance [Δ] between driver and wave is related to the speed and the size of the driver (*e.g.* Spreiter, Summers, and Alksne, 1966; Farris and Russell, 1994; Russell and Mulligan, 2002; Žic *et al.*, 2008). Physically, the Mach number [M] and the radius of curvature [R_{c}] of the nose of the driver control the standoff distance. Therefore, in Figure 8 our model results for the relative standoff distance [Δ/R_{c}] are presented as a function of M using $v_{\mathrm{A0}} = 500\ \mathrm{km\,s^{-1}}$. The model results are compared with one measurement of Δ/R_{c} determined at a time when the shock structure could be most clearly observed and the kinematical profile of the CME reached a constant speed of $\approx 600\ \mathrm{km\,s^{-1}}$. R_{c} is obtained by fitting a circle to the driver of the wave (see right panel of Figure 8). We derive for the relative standoff distance of the shock under study a value of $\approx 0.4 \pm 0.1 R_{\mathrm{c}}$. Considering a lower and an upper limit of the Alfvén speed ($300 < v_{\mathrm{A0}} < 500\ \mathrm{km\,s^{-1}}$) at the measured

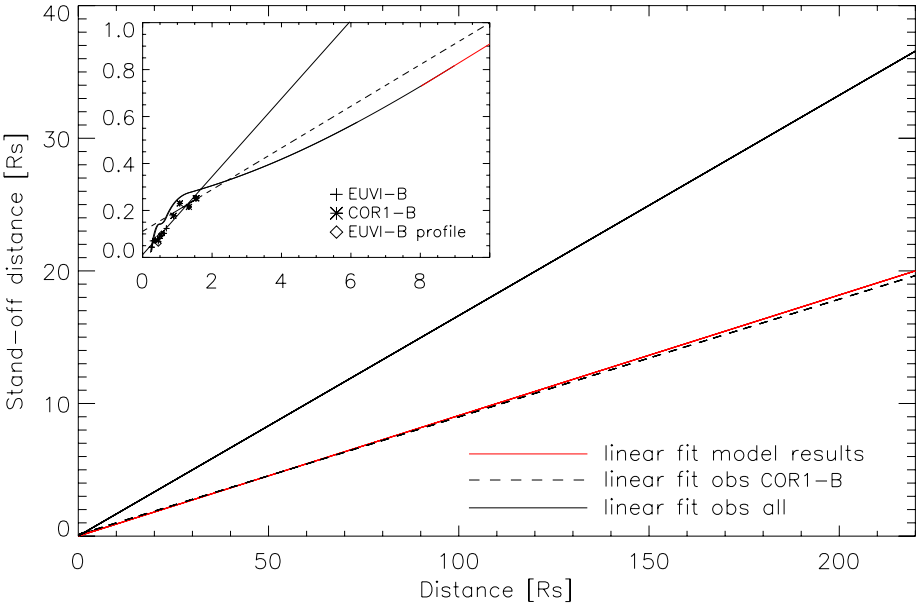


Figure 7 Standoff distance *versus* distance of top part of dome (starting from solar surface) over the Sun–Earth distance range as derived from linear fits to the model and observational results, respectively. The solid line indicates the linear extrapolation over all observed data points; the dashed line is a fit and its extrapolation to COR1-B data points only. The red line shows a linear fit to the model results over the distance range 8 – 10 $R_\odot$.

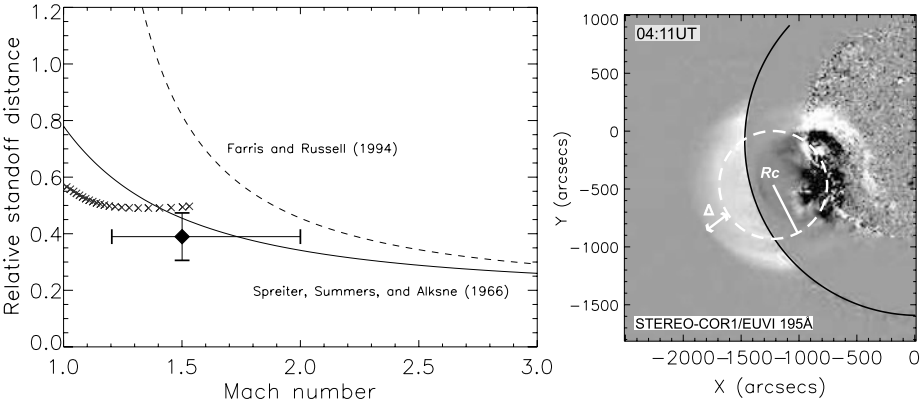


Figure 8 Standoff distance of a shock from a spherical obstacle with radius of curvature [R_c] (as shown in the right panel) expressed in terms of the radius of curvature for the models of Spreiter, Summers, and Alksne (1966) and of Farris and Russell (1994). See also Figure 6 in Russell and Mulligan (2002). The diamond gives the relative standoff distance [Δ/R_c] derived from observations together with error estimations for the uncertainty in the measurements as well as in the unknown value of v_{A0} and consequently M. The × symbols represent Δ/R_c derived from the model using $v_{A0} = 500\ \mathrm{km\,s^{-1}}$.

distance of $\approx 1.5\ R_\odot$ above the solar surface (Mann *et al.*, 2003), we obtain $1.2 < M < 2$. Hence, the vertical and horizontal error bars shown on the observational data point reflect the uncertainties in measurements as well as in the derivation of M, respectively.

For illustration only, in Figure 8 we also show the hydrodynamic model results of Spreiter, Summers, and Alksne (1966) and Farris and Russell (1994) as transformed by Russell and Mulligan (2002) in terms of R_c, which read

$$\Delta/R_c = 0.195 + 0.585M^{-2}$$

for the high Mach number approximation by Spreiter, Summers, and Alksne (1966) and

$$\Delta/R_c = 0.195 + 0.78\left(M^2 - 1\right)^{-1}$$

for the version by Farris and Russell (1994), which corrects the previous formula in the low Mach number regime.

Comparing the data shown in Figure 8, our model predicts much lower standoff distances at low Mach numbers than both models considered by Russell and Mulligan (2002). This is not surprising, since they considered entirely different physical situations. The models by Spreiter, Summers, and Alksne (1966) and Farris and Russell (1994) consider a stationary situation where the obstacle in a supersonic ambient flow has constant size and the ambient flow has steady speed. When applied to CMEs, these models describe a supersonic driver of a constant size and speed, *i.e.*, a standard bow-shock situation, and as $M \rightarrow 0$ the standoff distance increases $\Delta \rightarrow \infty$. In contrast, our model includes evolutionary aspect; *i.e.* it considers the acceleration stage of the source surface. Moreover, this stage, in addition to translatory motion, is characterized by expansion; *i.e.* it acts not only as a moving object, but also as a 3D piston. The latter effect is especially important, since it includes the effect of nonlinear evolution of the wave front, *i.e.* its steepening into a shock (*e.g.* Vršnak and Lulić, 2000). In this situation, the time and distance at which the shock forms and the magnetosheath thickness are determined by the acceleration–time profile of the driver, *i.e.* its kinematics (see, *e.g.*, Figure 4 in Vršnak and Lulić, 2000; Žic *et al.*, 2008).

Finally, note that most observational studies on standoff distances are related to fast CME events and Mach numbers > 1.5 (*e.g.* Maloney and Gallagher, 2011; Kim *et al.*, 2012; Gopalswamy and Yashiro, 2011). More observational studies are needed to give a more reliable conclusion about the evolution of the standoff distance for weak shocks.

5. Discussion and Conclusion

The 17 January 2010 event has been well observed, revealing a dome-shaped coronal wave structure. In this study we give evidence from observational and model results that the off-disk part of the wave actually consists of a driver and a wave component. The driver is interpreted as the CME, the frontal part as a weakly shocked wave. We derive that the shock standoff distance shows a linear evolution with a rather rapid increase below 0.3 $R_\odot$ above the solar surface. These results may be interpreted to mean that the initial lateral (over)expansion of the CME, which is short-lived ($\approx$ 70 seconds; see Patsourakos, Vourlidas, and Stenborg, 2010), acts as a piston driver to the shock, which leads to a rapid increase in the shock standoff distance. The piston nature of expanding CME flanks is also reflected in results from a recent study by Cheng *et al.* (2012). Using *Solar Dynamics Observatory* (SDO) observations, Cheng *et al.* (2012) analyzed the structural and kinematical evolution of a CME together with the separation process of a diffuse wave front from the CME flanks. The decoupling of the wave from the driver and the associated actual detection of the wave front occur when the CME expansion slows down.

Comparing our results to previous studies of off-disk waves moving in the radial direction away from the Sun, we find good agreement for several parameters. For the first observable we derive measurable standoff distance values of $0.03-0.06\ R_{\odot}$. Ma *et al.* (2011), who studied a low coronal shock wave using high spatial and temporal resolution data from SDO's *Atmospheric Imaging Assembly* (AIA), report a value of $\approx 0.03\ R_{\odot}$ for the thickness of the shocked layer. We find the first signatures of the shock at a distance of $\approx 0.28\ R_{\odot}$ above the solar surface, which is comparable to the result from Ma *et al.* (2011) of $0.23\ R_{\odot}$. From the model we also derive the timing of the shock signatures to be close to the observed type II radio burst. Cheng *et al.* (2012) refer to an almost simultaneous occurrence between a type II radio burst and the start of the separation process between wave and driver. For radio bursts, shock formation heights of $\approx 0.2\ R_{\odot}$ are derived by, *e.g.*, Magdalenić *et al.* (2010). We note that the height at which the shock forms is strongly dependent on the speed profile of the driver, *i.e.* the CME. Peak accelerations of CMEs occur at very small distances from their launch site: $< 0.5\ R_{\odot}$ above the solar surface (*e.g.* Temmer *et al.*, 2008, 2010).

Using an analytical model, the kinematical profile of both components, driver and wave, can be simulated by applying model parameters that are constrained by observations. In addition, we are able to simulate the on-disk wave using a pure piston-type expansion of the driving source, whereas the source of the off-disk wave behaves in the early evolution as a piston and then becomes more of a bow-shock type. We find that the on-disk wave requires a more impulsive driver ($t = 90$ seconds, $a = 1.7\ \mathrm{km\,s^{-2}}$) compared to the off-disk wave ($t = 340$ seconds, $a = 1.5\ \mathrm{km\,s^{-2}}$). These results lie between the findings of Grechnev *et al.* (2011), who obtain that the off-limb wave was probably excited impulsively from a filament eruption and then propagated freely, and those of Veronig *et al.* (2010), who conclude that the upward moving dome might have been driven all of the time.

The dome-shaped wave under study evolves from an eruption plus the deformation of different loop systems. In morphology, the dome-shaped wave is much more similar to a CME bubble than a separated loop system. In visible light, shock waves are reported as well-outlined and sharp boundaries (see Ontiveros and Vourlidas, 2009). For the evolution of a coronal surface wave, Temmer *et al.* (2010) reported that the wave was launched from two separate centers before it achieved a circular shape. We may speculate that the loop systems expand and are pushed aside due to the early evolution phase of the erupting structure. The magnetic loop structures form the "observable envelope" and are the first signatures of the evolving CME.

The current study shows that relatively slow drivers may cause weak shock waves low in the corona. These waves are visible in white light and may further propagate up to 1 AU. To investigate the evolution of shock standoff distances in interplanetary space, we require observations of the wave–driver system close to the Sun as well as their *in-situ* signatures.

Acknowledgements MT and AMV gratefully acknowledge the Austrian Science Fund (FWF): V195-N16 and P24092-N16. The research leading to these results has received funding from the European Commission's Seventh Framework Programme (FP7/2007-2013) under the grant agreements n° 263252 [COMESEP] and n° 284461 [eHEROES].

References

Bemporad, A., Mancuso, S.: 2010, *Astrophys. J.* **720**, 130. doi:10.1088/0004-637X/720/1/130.
Cheng, X., Zhang, J., Olmedo, O., Vourlidas, A., Ding, M.D., Liu, Y.: 2012, *Astrophys. J. Lett.* **745**, L5. doi:10.1088/2041-8205/745/1/L5.
Farris, M.H., Russell, C.T.: 1994, *J. Geophys. Res.* **99**, 17681. doi:10.1029/94JA01020.
Gopalswamy, N., Yashiro, S.: 2011, *Astrophys. J. Lett.* **736**, L17. doi:10.1088/2041-8205/736/1/L17.

Gopalswamy, N., Kundu, M.R., Manoharan, P.K., Raoult, A., Nitta, N., Zarka, P.: 1997, *Astrophys. J.* **486**, 1036. doi:10.1086/304556.
Gopalswamy, N., Lara, A., Yashiro, S., Kaiser, M.L., Howard, R.A.: 2001, *J. Geophys. Res.* **106**, 29207. doi:10.1029/2001JA000177.
Grechnev, V.V., Afanasyev, A.N., Uralov, A.M., Chertok, I.M., Eselevich, M.V., Eselevich, V.G., Rudenko, G.V., Kubo, Y.: 2011, *Solar Phys.* **273**, 461. doi:10.1007/s11207-011-9781-y.
Harrison, R.A., Lyons, M.: 2000, *Astron. Astrophys.* **358**, 1097.
Howard, R.A., Moses, J.D., Vourlidas, A., Newmark, J.S., Socker, D.G., Plunkett, S.P., *et al.*: 2008, *Space Sci. Rev.* **136**, 67. doi:10.1007/s11214-008-9341-4.
Illing, R.M.E., Hundhausen, A.J.: 1985, *J. Geophys. Res.* **90**, 275. doi:10.1029/JA090iA01p00275.
Kaiser, M.L., Kucera, T.A., Davila, J.M., St. Cyr, O.C., Guhathakurta, M., Christian, E.: 2008, *Space Sci. Rev.* **136**, 5. doi:10.1007/s11214-007-9277-0.
Kim, R.S., Gopalswamy, N., Moon, Y.J., Cho, K.S., Yashiro, S.: 2012, *Astrophys. J.* **746**, 118. doi:10.1088/0004-637X/746/2/118.
Kozarev, K.A., Korreck, K.E., Lobzin, V.V., Weber, M.A., Schwadron, N.A.: 2011, *Astrophys. J. Lett.* **733**, L25. doi:10.1088/2041-8205/733/2/L25.
Ma, S., Raymond, J.C., Golub, L., Lin, J., Chen, H., Grigis, P., Testa, P., Long, D.: 2011, *Astrophys. J.* **738**, 160. doi:10.1088/0004-637X/738/2/160.
Magdalenić, J., Marqué, C., Zhukov, A.N., Vršnak, B., Žic, T.: 2010, *Astrophys. J.* **718**, 266. doi:10.1088/0004-637X/718/1/266.
Maloney, S.A., Gallagher, P.T.: 2011, *Astrophys. J. Lett.* **736**, L5. doi:10.1088/2041-8205/736/1/L5.
Mann, G., Klassen, A., Aurass, H., Classen, H.T.: 2003, *Astron. Astrophys.* **400**, 329. doi:10.1051/0004-6361:20021593.
Ontiveros, V., Vourlidas, A.: 2009, *Astrophys. J.* **693**, 267. doi:10.1088/0004-637X/693/1/267.
Patsourakos, S., Vourlidas, A., Stenborg, G.: 2010, *Astrophys. J. Lett.* **724**, L188. doi:10.1088/2041-8205/724/2/L188.
Raymond, J.C., Thompson, B.J., St. Cyr, O.C., Gopalswamy, N., Kahler, S., Kaiser, M., Lara, A., Ciaravella, A., Romoli, M., O'Neal, R.: 2000, *Geophys. Res. Lett.* **27**, 1439. doi:10.1029/1999GL003669.
Russell, C.T., Mulligan, T.: 2002, *Planet. Space Sci.* **50**, 527. doi:10.1016/S0032-0633(02)00031-4.
Spreiter, J.R., Summers, A.L., Alksne, A.Y.: 1966, *Planet. Space Sci.* **14**, 223. doi:10.1016/0032-0633(66)90124-3.
Temmer, M., Veronig, A.M., Vršnak, B., Rybák, J., Gömöry, P., Stoiser, S., Maričić, D.: 2008, *Astrophys. J. Lett.* **673**, L95. doi:10.1086/527414.
Temmer, M., Vršnak, B., Žic, T., Veronig, A.M.: 2009, *Astrophys. J.* **702**, 1343. doi:10.1088/0004-637X/702/2/1343.
Temmer, M., Veronig, A.M., Kontar, E.P., Krucker, S., Vršnak, B.: 2010, *Astrophys. J.* **712**, 1410. doi:10.1088/0004-637X/712/2/1410.
Veronig, A.M., Muhr, N., Kienreich, I.W., Temmer, M., Vršnak, B.: 2010, *Astrophys. J. Lett.* **716**, L57. doi:10.1088/2041-8205/716/1/L57.
Vourlidas, A., Wu, S.T., Wang, A.H., Subramanian, P., Howard, R.A.: 2003, *Astrophys. J.* **598**, 1392. doi:10.1086/379098.
Vršnak, B.: 2005, *Eos Trans. AGU* **86**, 112. doi:10.1029/2005EO110004.
Vršnak, B., Lulić, S.: 2000, *Solar Phys.* **196**, 157. doi:10.1023/A:1005236804727.
Vršnak, B., Cliver, E.W.: 2008, *Solar Phys.* **253**, 215. doi:10.1007/s11207-008-9241-5.
Vršnak, B., Maričić, D., Stanger, A.L., Veronig, A.: 2004, *Solar Phys.* **225**, 355. doi:10.1007/s11207-004-4995-x.
Warmuth, A.: 2007, In: Klein, K.L., MacKinnon, A.L. (eds.) *Lecture Notes in Physics* **725**, Springer, Berlin, 107.
Wuelser, J., Lemen, J.R., Tarbell, T.D., Wolfson, C.J., Cannon, J.C., Carpenter, B.A., *et al.*: 2004, In: Fineschi, S., Gummin, M.A. (eds.) *SPIE* **CS-5171**, 111. doi:10.1117/12.506877.
Zhao, X.H., Wu, S.T., Wang, A.H., Vourlidas, A., Feng, X.S., Jiang, C.W.: 2011, *Astrophys. J.* **742**, 131. doi:10.1088/0004-637X/742/2/131.
Žic, T., Vršnak, B., Temmer, M., Jacobs, C.: 2008, *Solar Phys.* **253**, 237. doi:10.1007/s11207-008-9173-0.