Topics in Current Chemistry 365

Erin R. Johnson *Editor*

Density Functionals

Thermochemistry

Springer

365
Topics in Current Chemistry

Aims and Scope

The series Topics in Current Chemistry presents critical reviews of the present and future trends in modern chemical research. The scope of coverage includes all areas of chemical science including the interfaces with related disciplines such as biology, medicine and materials science.

The goal of each thematic volume is to give the non-specialist reader, whether at the university or in industry, a comprehensive overview of an area where new insights are emerging that are of interest to larger scientific audience.

Thus each review within the volume critically surveys one aspect of that topic and places it within the context of the volume as a whole. The most significant developments of the last 5 to 10 years should be presented. A description of the laboratory procedures involved is often useful to the reader. The coverage should not be exhaustive in data, but should rather be conceptual, concentrating on the methodological thinking that will allow the non-specialist reader to understand the information presented.

Discussion of possible future research directions in the area is welcome.

Review articles for the individual volumes are invited by the volume editors.

Readership: research chemists at universities or in industry, graduate students.

More information about this series at
http://www.springer.com/series/128

Erin R. Johnson
Editor

Density Functionals

Thermochemistry

With contributions by

A.D. Becke · S. Ehrlich · S. Grimme · A. Hansen ·
A. Heßelmann · B.G. Janesko · E.R. Johnson ·
M.A. Mosquera · A. Patra · J.P. Perdew · T. Risthaus ·
A. Savin · M. Steinmetz · A. Wasserman · B. Xiao

Springer

Editor
Erin R. Johnson
Department of Chemistry
Dalhousie University
Halifax, Nova Scotia
Canada

ISSN 0340-1022 ISSN 1436-5049 (electronic)
Topics in Current Chemistry
ISBN 978-3-319-19691-6 ISBN 978-3-319-19692-3 (eBook)
DOI 10.1007/978-3-319-19692-3

Library of Congress Control Number: 2015941274

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Preface

Density-functional theory (DFT) has its roots in the works of Hohenberg and Kohn [1], Kohn and Sham [2], and Slater [3]. Despite rumors of its untimely demise [4], DFT is presently enjoying ever greater application across a wide range of scientific disciplines, including chemistry, physics, biochemistry, and engineering. While density-functional approximations can be used for calculation of accurate structural and spectroscopic properties, they are most widely used for thermochemical predictions, such as heats of reaction, which are the subject of the present volume.

Becke's article applying gradient-corrected functionals to the "G1" set of atomization energies [5] was the key study which first attracted the interest of the computational-chemistry community to density-functional theory. The subsequent implementation and testing [6] in the Gaussian-92/DFT electronic-structure program marked the point where density functionals became a practical and attractive alternative to correlated-wavefunction theory for reaction thermochemistry. Because of their high accuracy and computational efficiency, DFT methods have become the standard approach for thermochemical predictions. Generalized gradient approximations (GGAs) and hybrid functionals are now implemented in a variety of commercial and open-source electronic-structure programs which can be readily used by chemical researchers.

In recent years, several other classes of functionals have been proposed. Notable developments include meta-GGAs, range-separated hybrid functionals, random-phase approximation functionals, and dispersion corrections, all of which have extended the range of systems to which one can reliably apply density-functional approximations. However, no functional is currently applicable to general thermochemical study of all types of chemical species and reactions. Particular challenges include systems with extensive electron delocalization or multi-reference character and thus the field is continuing to evolve.

This volume aims to reflect the current state of density-functional thermochemistry. It presents highlights of ongoing research at the frontiers of DFT and discusses the outstanding problems of delocalization error and strong correlation. The first few chapters give overviews of density-functional approximations as they

currently stand and push the boundaries of their applicability to increasingly complex systems, consisting of very large molecules (Chap. 1), surface chemistry (Chap. 2), and van der Waals dimers (Chap. 3). Next, we address the problems associated with judging the relative quality of functionals (Chap. 4). Finally, we consider research on non-conventional functionals, including recent successes of random-phase approximations for thermochemistry (Chap. 5), spin-density functionals and issues with derivative discontinuities (Chap. 6), and real-space strong-correlation models for exact-exchange-based DFT (Chap. 7), which may represent the future of the field.

In closing, I would like to thank Elizabeth Hawkins, Judith Hinterberg, Tanja Jaeger, and Arun Manoj Jayaraman at Springer for their help and for the opportunity to edit this volume. I would also like to thank the authors for their contributions, making this work possible, and the members of my research group for their support and camaraderie.

Merced, CA, USA Erin R. Johnson
February, 2015

References

1. Hohenberg P, Kohn W (1964) Inhomogeneous electron gas. Phys Rev 136:B864
2. Kohn W, Sham LJ (1965) Self-consistent equations including exchange and correlation effects. Phys Rev 140:A1133
3. Slater JC (1951) A simplification of the Hartree–Fock method. Phys Rev 81:385–390
4. Gill PMW (2001) Obituary: density functional theory (1927–1993). Aust J Chem 54:661–662
5. Becke AD (1992) Density-functional thermochemistry. I. The effect of the exchange only gradient correction. J Chem Phys 96:2155–2160
6. Johnson BG, Gill PMW, Pople JA (1992) Preliminary results on the performance of a family of density functional methods. J Chem Phys 97:7846–7848

Contents

Top Curr Chem (2015) 365: 1–24
DOI: 10.1007/128_2014_543

Published online: 24 April 2014

Accurate Thermochemistry for Large Molecules with Modern Density Functionals

Marc Steinmetz, Andreas Hansen, Stephan Ehrlich, Tobias Risthaus, and Stefan Grimme

Abstract The thermodynamic properties of molecules are of fundamental interest in chemistry and engineering. This chapter deals with developments made in the last few years in the search for accurate density functional theory-based quantum chemical electronic structure methods for this purpose. The typical target accuracy for reaction energies of larger systems in the condensed phase is realistically about 2 kcal/mol. This level is within reach of modern density functional approximations when combined with appropriate continuum solvation models and slightly modified thermostatistical corrections. Nine higher-level functionals of dispersion corrected hybrid, range-separated hybrid, and double-hybrid type were first tested on four common, mostly small molecule, thermochemical benchmark sets. These results are complemented by four large molecule reaction examples. In these systems with 70–200 atoms, long-range electron correlation is responsible for important parts of the interactions and dispersion-uncorrected functionals fail badly. When used together with properly polarized triple- or quadruple-zeta type AO basis sets, most of the investigated functionals provide accurate gas phase reaction energies close to the values estimated from experiment. The use of theoretical back-correction schemes for solvation and thermal effects, the impact of the self-interaction error for unsaturated systems, and the prospect of local coupled-cluster based reference energies as benchmarks are discussed.

M. Steinmetz, A. Hansen, S. Ehrlich, and S. Grimme (✉)
Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie der Universität Bonn, Beringstraße 4, 53115 Bonn, Germany
e-mail: marc.steinmetz@thch.uni-bonn.de; hansen@thch.uni-bonn.de; stephan.ehrlich@thch.uni-bonn.de; grimme@thch.uni-bonn.de

T. Risthaus
Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie der Universität Bonn, Beringstraße 4, 53115 Bonn, Germany

International NRW Graduate School of Chemistry, Wilhelm-Klemm-Str. 10, 48149 Münster, Germany
e-mail: tobias.risthaus@thch.uni-bonn.de

Keywords Thermochemistry benchmark, Density functional theory, Dispersion correction, Large molecules, Local coupled cluster

Contents

1 Introduction

In the past two decades Kohn–Sham Density Functional Theory (DFT) [1–5] has become a very important tool for understanding mechanistic problems in chemistry. At the heart of this topic is the proper energetic description of all chemical species involved in a reaction. This thermochemistry problem, which from a broader point of view may also contain transition states or non-equilibrium structures in addition to the normal minima, is the topic of this work. The evaluation of the performance of density functionals (DFs) by benchmarking for energetic properties is a crucial step prior to the investigation of a new system. The reason for this is the still somewhat empirical nature of current DF approximations and their non-systematic improvability.

Several molecular sets were developed over the last few years to test DFs for, e.g., atomization energies [6–8], non-covalent interactions (NCI) [9–11], or special reactions and kinetics [12–16]. Many of them were collected in the GMTKN30 [17] test set by our group to build a large benchmark set which includes the chemically most important properties of main-group chemistry. Less extensive benchmarks exist in the field of transition metal chemistry for which we mention a few examples [18–24].

The reference data used can be taken from experiment but nowadays it has become common practice to compute reference reaction energies at high Wave Function Theory (WFT) level (normally coupled-cluster) and compare these data for the same molecular geometry directly with DFT results. This procedure avoids the effects of temperature, conformations, solvents, and other uncertainties in the measurements, and is the preferred way in our group. However, it cannot always be applied to large molecules because the computationally demanding WFT calculations are intractable. Mixed approaches which combine experimental and theoretical data (back-correction schemes) represent a solution to the problem and will be discussed below.

Another important aspect in the context of DFT and larger chemical systems is the effect of the London dispersion interaction. As an electron correlation effect, dispersion has a fundamentally quantum chemical, complex many-particle origin but is chemically a local phenomenon. It often operates on a relatively long-range length scale where classical (atomic or other local) approximations perform well (and exchange is negligible) but also has a short-range component. The dispersion energy, and in particular its long-range (London) part, is not accurately described by common semi-local DFs [25–27] and dispersion corrections represent a very active field of research [28–30]. In this chapter we also want to highlight the importance of the dispersion energy in intramolecular cases and in thermochemistry generally. Even if dispersion alone is insufficient to form a stable chemical bond, it is clear that larger molecules are significantly more influenced by the intramolecular dispersion energy than smaller ones. Because dispersion as a special type of electron correlation effect is always attractive (energy lowering), this means larger molecules are thermodynamically stabilized by dispersion compared to small systems [31]. From this new concept it is concluded that large (preferably electron-rich and polarizable) functional groups can be used to stabilize thermodynamically (and not merely kinetically) weak bonds or reactive parts in a molecule. The chemical examples discussed at the end of this chapter illustrate this point.

As already mentioned, DFT has also become the "work-horse" of modern quantum chemistry because it represents a good compromise between computational effort and accuracy. However, the huge number of developed DFs to date shows that current approximations still suffer from several flaws and that the quest for finding a functional which comes close to the "true one" is still going on. In this context, we want to particularly focus on the fact that not every DF is equally well applicable to every problem [32]. This makes choosing the right functional for the right problem a tough task, even for experienced researchers in this field. Here, we want to shed light on the question whether very recent, newly or further developed functionals are accurate for thermochemistry and concomitantly robust, i.e., broadly applicable to various chemical problems. For evaluation we employ a combination of standard thermochemical benchmark sets and four "real" chemical reactions of molecules with about 70–200 atoms. Nine modern, higher-level functionals of dispersion corrected hybrid, range-separated hybrid, and double-hybrid type are tested.

2 Theory

2.1 Thermochemical Calculations in the Condensed Phase

A free reaction energy ΔG_r in solution, which is often measured experimentally under convenient equilibrium conditions for solvent X at temperature T, can be computed as

$$\Delta G_{\mathrm{r}} \approx \Delta E_{\mathrm{r}}(\mathrm{gas}) + \underbrace{\Delta G^{T}_{\mathrm{TRV}}(\mathrm{gas})}_{\substack{\mathrm{DFT,\,RRHO,}\\ \text{low-freq. mode}\\ \text{approx.}}} + \underbrace{\Delta\delta G^{T}_{\mathrm{solv}}(\mathrm{solvent}\,X)}_{\mathrm{COSMO\text{-}RS}} \tag{1}$$

where $\Delta E(\mathrm{gas})$ is the zero-point-vibrational exclusive reaction energy for the isolated molecules, $\Delta G^{T}_{\mathrm{TRV}}(\mathrm{gas})$ is the thermo-statistical correction from energy to free energy with translational, rotational, and vibrational contributions, and $\Delta\delta G^{T}_{\mathrm{solv}}$ is the solvation free energy contribution. For the latter term we employ the COSMO-RS continuum solvation model [33–35] throughout. It is based on single point calculations on the default BP86/def-TZVP [36–38] level of theory for optimized gas phase structures. Consistently, the rigid-rotor-harmonic-oscillator (RRHO) model is also based on gas phase structures. Note that the normally small effects stemming from changes of the structure and vibrational frequencies upon solvation are implicitly accounted for by the COSMO-RS parametrization. For charged or very polar species, where larger changes are expected, we usually compute the structure and frequencies at the DFT-D3/COSMO level. Low-lying vibrational modes ($<$100 cm^{-1}) are treated by a special rigid-rotor approximation in order to avoid numerical artifacts in the entropy calculations [39].

Equation (1) can be rearranged to obtain approximate experimental reaction energies from condensed phase equilibrium measurements and the corresponding theoretical correction energy terms given above according to

$$\Delta E_{\mathrm{r}}(\mathrm{gas})^{\mathrm{exptl.}} \approx \Delta G_{\mathrm{r}}^{\mathrm{exptl.}} - \Delta G^{T}_{\mathrm{TRV}}(\mathrm{gas}) - \Delta\delta G^{T}_{\mathrm{solv}}(\mathrm{solvent}\,X). \tag{2}$$

This approach is used to obtain reference reaction energies for three examples discussed in Sect. 3.2 with which the various DF results are compared. If experimental values are not available, the quantum chemistry "gold standard," namely coupled-cluster with single and double excitations and perturbative triple excitations (CCSD(T)), is employed to obtain reference reaction energies directly. For large molecules it cannot be applied in its canonical form due a steep increase of the computational effort with the number of correlated electrons. However, CCSD(T) calculations on larger molecules become computationally feasible if local correlation approaches are applied. Among the growing number of local coupled cluster implementations, the recently published DLPNO-CCSD(T) method [40], which employs pair-natural orbitals (PNOs) and domain-based techniques, seems to be very promising. It shows near linear scaling of the computation time with the system size and, hence, molecular calculations with up to 200 atoms and reasonable valence triple-zeta AO basis sets are possible, although the computational demands are still significantly higher compared to DFT methods. It has been shown [40] and confirmed by us [41] that the errors due to the additional approximations in the DLPNO-CCSD(T) method are small ($<$1–2 kcal/mol) and well controllable. Using

a tight value of the electron pair cut-off ($T_{\mathrm{cutPairs}} = 10^{-5}\,E_{\mathrm{h}}$) to ensure that London dispersion interactions are captured properly, and estimating the remaining basis set incompleteness error by focal point analysis [42, 43], DLPNO-CCSD(T) offers a reliable means for obtaining accurate reference reaction energies if experimental values are missing.

The complete basis set (CBS) DLPNO-CCSD(T) results were estimated from the following standard additivity scheme for the electronic energy E in which a correction from a def2-TZVP (TZ) [44] calculation is added to the MP2/CBS result:

$$E(\mathrm{CCSD(T)/CBS}) \approx E(\mathrm{MP2/CBS}) + [E(\mathrm{CCSD(T)/TZ}) - E(\mathrm{MP2/TZ})]. \quad (3)$$

Here E(CCSD(T)) refers to a canonical CCSD(T) or for large molecules to a DLPNO-CCSD(T) energy.

For neutral systems and reaction energies in a "normal" range (10–40 kcal/mol), the errors of the back-correction scheme as well as residual errors in the DLPNO-CCSD(T)/CBS treatment (mostly due to basis set extrapolation errors and non-additivity effects) typically amount to 10% of ΔE, i.e., 2–3 kcal/mol. As will be shown below this is within the error range of most modern DFs. According to our experience, the often claimed "chemical accuracy" of about 1 kcal/mol for thermochemistry is unrealistic for large systems from the experimental as well as the theoretical point of view and the quoted 2–3 kcal/mol should be considered as a more realistic target.

2.2 The Functional "Zoo"

The "zoo" of density functionals [45] has grown tremendously over the years and it has become difficult even for an expert in the field to follow all developments continuously. About a decade ago it was basically sufficient to specify the choice of the Generalized Gradient Approximation (GGA) flavor in a DF with the amount of non-local Fock-exchange included (as specified by the mixing parameter a_x [46, 47]). Nowadays, the exchange-correlation energy E_{XC} is composed of more diverse components. We want to clarify this issue with the help of the general formula

$$E_{\mathrm{XC}} = E_X^{\mathrm{GGA}} + E_X^{\mathrm{NL}}(a_x, \mu) + E_{\mathrm{C}}^{\mathrm{GGA}} + E_{\mathrm{C}}^{\mathrm{NL}} \quad (4)$$

where $E_{\mathrm{XC}}^{\mathrm{GGA}}$ represents semi-local GGA exchange-correlation energy components, $E_X^{\mathrm{NL}}(a_x, \mu)$ is the non-local (NL) Fock-exchange energy determined by the global mixing parameter a_x and possibly the range-separation parameter μ, with the non-local correlation energy $E_{\mathrm{C}}^{\mathrm{NL}}$ describing mostly long-range London dispersion effects. These three parts constitute three independent "coordinate axes" which

define the functional space of modern DFT. Because there are many technically and theoretically different choices for the three parts, and since there is no universally accepted theory on how to combine them, the number of possible (and probably at least reasonably performing) DFs is huge. In the spirit of John Perdew's so-called "Jacob's Ladder" scheme [48, 49], however, one can make some general classifications (cf. Peverati and Truhlar [50]):

- Global hybrids like B3LYP [47, 51] or PBE0 [52, 53] employ standard GGA and Local Density Approximation (LDA) components and a fixed amount of Fock-exchange. The hybrids performing optimally for thermochemistry have small a_x values, typically in the range 0.1–0.3. They have been used extensively in chemistry and recently their theoretical foundation was discussed in some detail [54]. If properly corrected for London dispersion effects, e.g., by the atom-pairwise D3 [55, 56] or density-dependent VV10 [57] schemes (i.e., inclusion of $E_{\mathrm{C}}^{\mathrm{NL}}$, see below), good accuracy for thermochemical properties can be obtained [50, 58, 59].
- Range-separated hybrids (RSHs) split the two-electron operator based on the inter-electronic distance into short-range (SR) and long-range (LR) parts, which are then treated differently. Hirao and coworkers [60] realized this splitting as

$$r_{12}^{-1} = \underbrace{\mathrm{erfc}(\mu \cdot r_{12}) \cdot r_{12}^{-1}}_{\mathrm{SR}} + \underbrace{\mathrm{erf}(\mu \cdot r_{12}) \cdot r_{12}^{-1}}_{\mathrm{LR}}, \tag{5}$$

  where $\mathrm{erf}(x)$ is the error function, $\mathrm{erfc}(x) = 1 - \mathrm{erf}(x)$, and μ is an adjustable parameter controlling the switch between the two regimes SR and LR. Usually, the SR part is treated by a conventional GGA while the LR part is considered "exactly," i.e., Fock-exchange is taken. Again, such functionals can be dispersion-corrected for better performance like ωB97X-D or ωB97X-D3 [61, 62]. Note that simple RSHs derived from standard GGA components perform worse for thermochemistry than global hybrids [58] but improve reaction barriers.
- If $E_{\mathrm{C}}^{\mathrm{NL}}$ contains an orbital-dependent term which is computed by second-order perturbation theory, the functionals are called double-hybrids (DHs). The similar term "doubly-hybrid" for a linear combination of DFT and MP2 parts was first coined in Zhao et al. [63] (the term DH was first used in Neese et al. [64]). The first DF in this class was B2PLYP [65] (for earlier related mixtures of DF and MP2 components, see [63, 66, 67]). A general expression for the correlation energy E_{C} in modern DFs is given by

$$E_{\mathrm{C}} = (1 - a_{\mathrm{c}}) E_{\mathrm{C}}^{\mathrm{GGA}} + a_{\mathrm{c}} E_{\mathrm{C}}^{\mathrm{PT2}} + E_{\mathrm{C}}^{\mathrm{disp}}, \tag{6}$$

  where a_{c} is a local/non-local mixing parameter (in analogy to the Fock-exchange mixing parameter a_x), $E_{\mathrm{C}}^{\mathrm{PT2}}$ is the standard MP2 correlation energy expression

but evaluated with hybrid-GGA orbitals and eigenvalues, and E_C^{disp} is a further London dispersion energy correction. In the limit $a_c = 0$ a normal hybrid is obtained. For the PT2 part, different scale factors for same- and opposite-spin pair correlation energies in the spirit of the SCS-MP2 method [68, 69] can be used, increasing robustness in electronically complicated situations and possibly leading to computational savings. According to extensive benchmarks [58], DHDFs are the best performing methods in the functional "zoo." For a thorough review of DHDFs including extensive benchmarking, see Kozuch et al. [70].

- Dispersion effects are ubiquitous in matter and hence we briefly describe our recommended procedure for their computation. For reviews on dispersion corrections to DFT and other atom-pairwise approaches, see [28–30, 71–73]. In the VV10 scheme by Vydrov and van Voorhis [57] (application denoted by "-NL" or "-V"), which is based on earlier work by Langreth and Lundqvist [74], the non-local correlation (dispersion) energy takes the form of a double-space integral

$$E_C^{disp,\,VV10} = \frac{1}{2}\iint \rho(r)\phi\left(r, r'\right)\rho\left(r'\right)\mathrm{d}r\mathrm{d}r', \tag{7}$$

where ρ is the charge density and r and r' denote electron coordinates. The different flavors of such density-dependent corrections [57, 74, 75] only differ in the choice of the non-local correlation kernel $\phi(r, r')$. These kernels are physically based on local approximations to the (averaged) dipole polarizability at frequency ω (i.e., $\alpha(r, \omega)$). Knowing α at all (imaginary) frequencies leads automatically, via the famous Casimir–Polder relationship [76], to the long-range part of the dispersion energy. This clarifies the deep relation to atom-pairwise methods (e.g., the DFT-D3 [55] approach used here) which employ these coefficients as basic quantities and replace the charge density by atom-centered delta functions and the double-integral by a double-sum. The C_6 dispersion coefficient for induced dipole–dipole dispersion interacting fragments A and B is given by

$$C_6^{AB} = \frac{3}{\pi}\int_0^\infty \alpha(\mathrm{i}\omega)^A\alpha(\mathrm{i}\omega)^B\mathrm{d}\omega. \tag{8}$$

Higher-order dipole–quadrupole, quadrupole–quadrupole, ... coefficients (i.e., C_8, C_{10}, ...) can also be computed by similar formulas [77]. The C_6 coefficients (and derived C_8) in the D3 method were obtained from a modified form of this relation where the $\alpha(\mathrm{i}\omega)$ are computed non-empirically by time-dependent DFT and A and B are reference molecules from which atomic values are derived [55]. Because the reference system can also be a molecular cluster modeling a solid environment, special coefficients for atoms in the bulk can be derived [78] (for a discussion of these atom-in-molecules effects, see Johnson

[79]). The final form for the DFT-D3 two-body part of the dispersion energy employs the so-called Becke–Johnson (BJ) damping [56, 71] and truncates the expansion at C_8:

$$E_C^{\mathrm{disp,D3(BJ)}} = -\frac{1}{2}\sum_{A\neq B} s_6 \frac{C_6^{AB}}{R_{AB}^6 + f\left(R_{AB}^0\right)^6} + s_8 \frac{C_8^{AB}}{R_{AB}^8 + f\left(R_{AB}^0\right)^8}, \tag{9}$$

where $f(R_{AB}^0) = a_1 R_{AB}^0 + a_2$, and a_1, a_2, s_6 and s_8 are empirical parameters that have been determined by a fit to CCSD(T) interaction energies for typical NCI (i.e., the S66 set [10]). The VV10 functional also contains two empirical parameters (adjusting the long- and short-range behavior) of which only the latter is fitted for each DF [59]. The above form of the damping function (which is similar in VV10) ensures that, for small interatomic distances, the right constant limit of the dispersion energy is obtained [80]. The D3 or VV10 corrections can be added to all semi-local DFs that are dispersion devoid in the medium-range regime so that no significant double-counting effects occur. For DHDFs the correction is scaled to complement the contribution of the orbital-dependent PT2 part [17] or empirically fitted [70].

Although these DFs do not fully solve the "bouquet of DFT puzzles" [81] (see also [82]), the modern functionals considered in this work are expected to provide, for "non-exotic" electronic structure problems, on average a significantly higher accuracy than what was typically obtained a decade ago. This will be illustrated by selected sets from our GMTKN30 [17] database and a few "real-life" examples involving large but in a chemical sense prototypical systems. We consider three standard dispersion corrected hybrids with two flavors of dispersion correction (B3LYP-NL, PW6B95-D3, and PBE0-D3), two RSHs which have been constructed consistently by including dispersion in the empirical fittings ωB97X-D3 [62] and ωB97X-V [83], and three versions of double-hybrid DFs which are based on different construction principles (DSD-PBEP86-D3 [84], PWPB95-D3 [17], and PBE0-DH [85]). The highly parametrized M06-2X meta-hybrid DF [86] is included because of its widespread use and to investigate the question of how much accuracy is lost by including only the medium-range dispersion energy [87]. For comparison, dispersion uncorrected B3LYP results are also given. Table 1 provides an overview and some properties of the investigated functionals.

Rather important in practical applications is the amount of non-local Fock-exchange included as determined by the mixing parameter a_x (in RSHs the non-local exchange contribution normally reaches 100% for large inter-electronic distances). It determines the magnitude of the so-called self-interaction error (SIE) which leads to too low reaction barriers, too loosely bound electrons, and over-delocalized electronic structures [88–90]. In particular, these problems may arise in unsaturated, radical-containing structures. Most DHDFs employ larger a_x values and hence suffer less from SIE because the unwanted effects of the larger Fock-exchange contribution are compensated by the orbital-dependent correlation

Table 1 Overview of the investigated density functionals

DF	Type	Parameters	a_x	Dispersion	References
B3LYP-NL	GGA hybrid	4	0.2	Density	[47, 51, 57, 59]
PBE0-D3	GGA hybrid	3	0.25	Atom-pairs	[52, 53, 55, 56]
PW6B95-D3	Meta-GGA hybrid	9	0.28	Atom-pairs	[55, 56, 133]
M06-2X	Meta-GGA hybrid	33	0.54	Density[a]	[134]
ωB97X-D3	RSH	16	0.195728–1	Atom-pairs	[55, 62]
ωB97X-V	RSH	10	0.167–1	Density	[57, 83]
PBE0-DH	GGA DH[b]	None	0.5	Orbitals	[85]
PWPB95-D3	Meta-GGA DH[c]	8	0.5	Orbitals/ atom-pairs	[17, 55, 56]
DSD-PBEP86-D3	Meta-GGA DH[d]	6	0.7	Orbitals/ atom-pairs	[55, 56, 84]

[a]Long-range (asymptotic) part not included
[b] $a_c = a_x^3 = 0.125$
[c]Opposite-spin correlation energy only, $a_c = 0.269$
[d]Different scale factors for same- and opposite-spin correlation

energy. The influence of SIE on thermochemical properties and its structural dependence in large systems is much less clear than that of the dispersion energy.

2.3 Technical Details of Quantum Chemical Calculations

The most important issue to consider in practical calculations is the choice of the one-particle atomic orbital (AO) basis set which is used to expand the orbitals in the Kohn–Sham approach and to generate excitation spaces in PT2 or CCSD(T) treatments. While thermochemical results with hybrid DFs are typically already close to convergence with properly polarized triple-ζ type basis sets (e.g., def2-TZVP [44] or cc-pVTZ [91]), DHDFs require larger sets due to the presence of the perturbation term [58]. Therefore, we employ here in single-point energy computations the def2-QZVP basis deprived of *g*-functions on non-hydrogen atoms and *f*-functions on hydrogen and lithium atoms. In standard notation this basis reads for first- and second-row elements [7s4p3d2f]/[4s3p2d]. According to many tests performed over the years for various reactions (see, e.g., [24, 58, 92]), this basis set level provides reaction energies within 1–2 kcal/mol of the complete basis set (CBS) limit for the cases considered here. Because basis set effects are largest for energies but only moderate in structure optimizations (see, e.g., [93]), the def2-TZVP level is sufficient for the latter purpose and hence most structures are based on TPSS-D3/def2-TZVP optimizations. The structures in the subsets of the

GMTKN30 [17] database in Sect. 3.1 were taken without modification. For the palladium atoms (DIMPD reaction) an ECP (SD(28,MWB) [94]) (and respective ECP basis set) was used. In this case, the geometries used were obtained on the PBE-D3/def2-TZVP level. The MP2 energies used in the DLPNO-CCSD(T) calculations are based on def2-TZVP/def2-QZVP extrapolations according to the procedure of Halkier et al. [95, 96]. For the two-electron Coulomb integrals the RI-approximation [97–99] was used which speeds up the computations remarkably without any significant loss of accuracy when optimized auxiliary basis sets [100, 101] are used. The MP2 results are also based on RI-treatments [102] with the corresponding exchange-type auxiliary basis sets [103]. All orbital-dependent correlation energies were obtained within the frozen (chemical) core approximation. The numerical quadrature grid m5 was generally employed for the integration of the exchange-correlation contribution. For the M06-2X calculations with ORCA the larger grid 7 was used because results with this functional are known to be strongly grid-dependent [58, 104]. Moreover, it was recently reported [105] that the M06-2X functional produces artificially large basis set superposition errors (BSSE) even with very large AO basis sets indicating some numerical instability.

All electronic energy and frequency calculations were conducted with TURBOMOLE [106, 107] or ORCA [108, 109] codes which provide practically identical results for similar technical settings. The COSMO-RS corrections were obtained from the COSMOtherm [110] software package. For further numerical details and discussions of the back-correction scheme, see Grimme [39]. All calculations involved in the DLPNO-CCSD(T) treatment were conducted with the ORCA code.

3 Examples

3.1 Thermochemical Benchmark Sets

A very convenient and unbiased way to assess the "global" accuracy of DFs is using the so-called GMTKN30 [17] database developed in our group over several years [111]. This benchmark covers 30 subsets related to general main group thermochemistry, kinetics, and NCI. In total, it encompasses 1,218 single-point calculations and 841 data points (relative energies). It therefore turned out to be ideal for evaluation and development of DFT methods. Here we utilize only parts of the GMTKN30 database and concentrate on four prototypical benchmarks for "true" chemical reactions which are described below. Intermolecular NCI have been studied extensively in recent years [9, 10, 112, 113] and are not considered here. It is noted that all tested functionals in this work perform well for NCI as long as they are corrected for long-range London dispersion effects by, e.g., D3 [55, 56] or NL(VV10) [57, 59] methods. As usual, fixed molecular structures are used and all energies are vibrational zero-point energy exclusive, which can conveniently be compared to the result of a standard QC calculation.

The MB08-165 subset: the "mindless-benchmark" set (MB08-165) was introduced by Korth and Grimme. [16] It contains 165 randomly created so-called "artificial molecules" (AMs) with varying constituents. For these molecules, decomposition energies into their hydrides (for the main group elements 1–4) and homonuclear diatomics (main group elements 5–7) were calculated. For these reactions, estimated CCSD(T)/CBS reference values were computed. In contrast to other benchmark sets, MB08-165 is less biased towards certain chemical aspects, as it contains artificial systems only. Korth and Grimme assessed a variety of density functionals and could reproduce nicely the Jacob's Ladder scheme, with higher rung functionals yielding better results. We chose MB08-165 to be the first subset of our benchmark study as it can be regarded as one of the most important for general thermochemistry, in particular in difficult situations when the electronic structure is not fully clear. Although the stated objective of this investigation is to focus on non-exotic cases, we believe it is also important to test the limits of DFT. Compared to the other subsets, it contains a large number of reference values and rather large reaction energies (117 kcal/mol on average, with an energy range from −570.6 to 433.7 kcal/mol).

The G2RC subset: contains 25 reactions, whose reactants and products are part of the G2/97 set of heats of formation [6]. Based on vibrationally back-corrected experimental data from Curtiss et al. [6], reference energies were calculated. The G2RC set comprises 47 single point calculations and has an average absolute reaction energy of 50.6 kcal/mol, with an energy range from −1.0 to −212.7 kcal/mol. It contains relatively small molecules (benzene being the largest) but is non-trivial because chemically large structural changes (e.g., transformation of multiple to single bonds) occur. In contrast to MB08-165, however, it covers only the conventional chemical space.

The BH76 subset: a fusion of the HTBH38 [114] and NHTBH38 [12] databases by Truhlar and co-workers. HTBH38 contains forward and reverse barriers of 19 hydrogen atom transfer reactions. NHTBH38 comprises 38 barriers of 19 heavy atom transfer, nucleophilic substitution, unimolecular, and association reactions. Reference values are based on high-level W1 calculations and "best theoretical estimates" (see [12, 114] for more details). The combined BH76 test set involves 95 single point calculations and has an average barrier height of 18.5 kcal/mol, with an energy range from −15.5 to 106.2 kcal/mol. Because the reaction energy considered always involves a transition state in which chemical bonds are partially broken, the results are sensitive to the treatment of the SIE in approximate DF.

The ISOL24-6 subset: Huenerbein et al. recently published a new benchmark set containing 24 isomerization reactions (ISOL24 [15]) of large molecules covering a wide range of different compounds, like, e.g., a sugar, a steroid, an organic dye, hydrocarbons, and large molecules containing heteroatoms. As reference, estimated SCS-MP3/CBS(TQ) was used. In contrast to the popular ISO34 set [14], which is a part of GMTKN30, the large size of the molecules casts additional light on effects that are important in "real life" organic chemistry. These are in particular intramolecular London-dispersion effects. Charged systems are also considered. This set

Table 2 Statistical results for the thermochemical benchmark sets

	MB08-165		G2RC		BH76		ISOL24-6		All
	MAD	MaxD	MAD	MaxD	MAD	MaxD	MAD	MaxD	$\overline{\text{MAD}}$
B3LYP	8.2	32.0	2.6	6.9	4.7	11.0	2.5	4.1	4.5
B3LYP-NL	5.9	23.1	3.2	10.3	6.0	13.1	1.2	2.8	4.1
PBE0-D3	8.9	36.7	7.0	24.6	4.5	14.6	1.0	1.6	5.4
PW6B95-D3	4.8	16.9	3.4	10.2	3.4	9.2	1.5	3.4	3.3
M06-2X	4.8[a]	25.5[a]	2.6	6.9	1.3	3.7	1.3	2.1	2.5
ωB97X-D3	8.9	47.0	5.3	15.3	2.6	5.9	1.1	2.8	4.5
ωB97X-V	7.6	33.8	4.3	13.1	1.9	5.2	1.5	3.2	3.8
PBE0-DH	8.5	50.9	6.6	21.8	1.7	5.9	1.3	3.3	4.5
PWPB95-D3	2.6	13.0	2.3	7.6	1.7	5.2	1.0	2.4	1.9
DSD-PBEP86-D3	2.5	14.2	2.2	12.3	1.3	8.3	0.3	0.4	1.6

[a]Two cases were excluded because of SCF convergence problems

Mean absolute deviation (MAD), maximum absolute deviation (MaxD), and mean MAD ($\overline{\text{MAD}}$) are given in kcal/mol

has been reconsidered by Truhlar et al. [115] and for a selection of smaller systems (with 24–35 atoms) new, higher-level CCSD(T) reference values were computed. The energy range is from 4.7 to 33.5 kcal/mol and the average reaction energy is 13.6 kcal/mol.

The statistical data (mean and maximum deviations) for the four benchmarks sets are given in Table 2 along with an overall performance measure (mean of the four MAD values) for each DF. Because the sets are of varying complexity and the reaction energies are also of different magnitude, we mostly consider the relative performance of the DF and take the maximum deviation as a measure of robustness (which should be proportional to the number of expected outliers). We have also investigated a weighted MAD similar to that of Goerigk and Grimme [17], but have found virtually no difference in relative performances.

From the mean $\overline{\text{MAD}}$ one can identify three groups of DFs with increasing accuracy: (1) PBE0-D3, PBE0-DH, and ωB97X-D3 with values in the range 4.5–5.4 kcal/mol and a largest MaxD value >36 kcal/mol; (2) B3LYP-NL, ωB97X-V, and PW6B95-D3, with values in the range 3.3–4.1 kcal/mol and largest MaxD values in the range 17–34 kcal/mol; (3) M06-2X, PWPB95-D3, and DSD-PBEP86-D3 with MAD values of 1.6–2.5 kcal/mol and largest MaxD values of 13–25 kcal/mol. According to these four benchmarks, the PWPB95-D3 and DSD-PBEP86-D3 double-hybrids are clearly the best performers. They provide the smallest MAD for all sets, the lowest mean MAD, and always very low MaxD values. The other double-hybrid (PBE0-DH) performs less well, which is not unexpected because it misses a dispersion correction and the non-local perturbation contribution is much smaller than for the other two DHDFs. The only moderately good result for ωB97X-V is somewhat disappointing because this functional tries to account for all exchange-correlation effects, and is based on extensive parametrization and/or empirical searches for best performing functional components. The improved

accuracy of ωB97X-V compared to ωB97X-D3 at least shows that some progress can be made even on a relatively high accuracy level at a formal cost lower than for most DHDFs. The good performance of B3LYP with the density dependent dispersion correction is noteworthy and, as mentioned before [59], the combination of an over-repulsive hybrid GGA part with a (partly overbinding) medium-range correlation part (VV10) is beneficial due to systematic error compensation. B3LYP-NL yields mostly small MaxD values and only fails for the barrier heights (similar to PBE0-D3) due to the SIE introduced by the too small Fock-exchange component. Although in particular the LYP part in B3LYP is sometimes responsible for larger errors [17, 116], this functional in dispersion-corrected form (and this also holds for B3LYP-D3) can still be recommended for thermochemistry at least for comparison (error estimation) purposes. However, at a similar numerical complexity and empiricism level, PW6B95-D3 provides in general better results and hence represents our default hybrid DF. It is often close to the performance of M06-2X which unfortunately is numerically unstable as noted above. Nevertheless, it is clear that all "modern" functionals perform better than standard B3LYP which yields large errors, in particular for the MB08-165 and ISOL24-6 sets. The only moderately positive effect of the NL(VV10) dispersion correction to the B3LYP results is mostly rooted in the small molecules considered so far. Whether this picture also prevails in larger, more realistic situations will be discussed below.

3.2 Chemical Reaction Examples

In the following sections we discuss the results for four "real-life" chemical applications involving rather large systems: the dissociation energy of a substituted hexaphenylethane (HEXAPE), of a substituted, dimeric hydrochinone derivative (DHCH), and the ligand exchange and dissociation reaction of a dimeric palladium species (DIMPD). The last example considers activation of H_2 by a so-called frustrated Lewis pair (FLP) which is a very active field of chemical research [117]. The results for all tested functionals are given in Table 3; the reaction formulas and structures are shown graphically in each subsection.

3.2.1 Hexaphenylethane

The predominant view in chemistry is that bulky groups in molecular structures are more repulsive rather than stabilizing. In particular the widespread misconception that a *tert*-butyl group only acts repulsively has recently been challenged by Grimme and Schreiner who re-investigated the textbook case of hexaphenylethane (Fig. 1) [118]. The stabilization of normal covalent bonds by dispersion in large systems is rather obvious from general considerations of the size and distance dependence of the dispersion energy [31]. Here we discuss a case in which the dispersion interaction between seemingly "innocent" ligands provides the main driving force for binding, meaning that without these forces the system would

Table 3 Results for the four chemical reaction energies ΔE (in kcal/mol)

	HEXAPE	DHCH	DIMPD	FLPH2
Reference	33 ± 4[a]	17 ± 2[b]	-32 ± 3[c]	-21.6[d]
B3LYP	−39.3	−14.6	−11.9	−18.6
B3LYP-NL	35.8	4.4	−34.8	−18.3
PBE0-D3	36.5	3.8	−41.4	−23.4
PW6B95-D3	30.5	1.4	−36.3	−19.3
M06-2X	28.2	8.3	−20.5	−19.1
ωB97X-D3	60.6	12.3	−42.1	−23.3
ωB97X-V	45.6	13.0	−36.6	−22.7
PBE0-DH	5.6	1.7	−28.8	−24.0
PWPB95-D3	41.2	12.6	−39.1	−19.3
DSD-PBEP86-D3	[e]	27.1	−43.7	−20.8

[a]Back-corrected ΔG value; see Grimme and Schreiner [118]. The experimental ΔG is not precisely known and a value of zero which is compatible with the observed radical concentration has been assumed

[b]Back-corrected experimental ΔG value; see Stephan and Erker [120]

[c]Back-corrected ΔG value from ITC measurements; see Hansen et al. [41]

[d]DLPNO-CCSD(T)/CBS value

[e]Not computed because of technical problems in the perturbation calculation

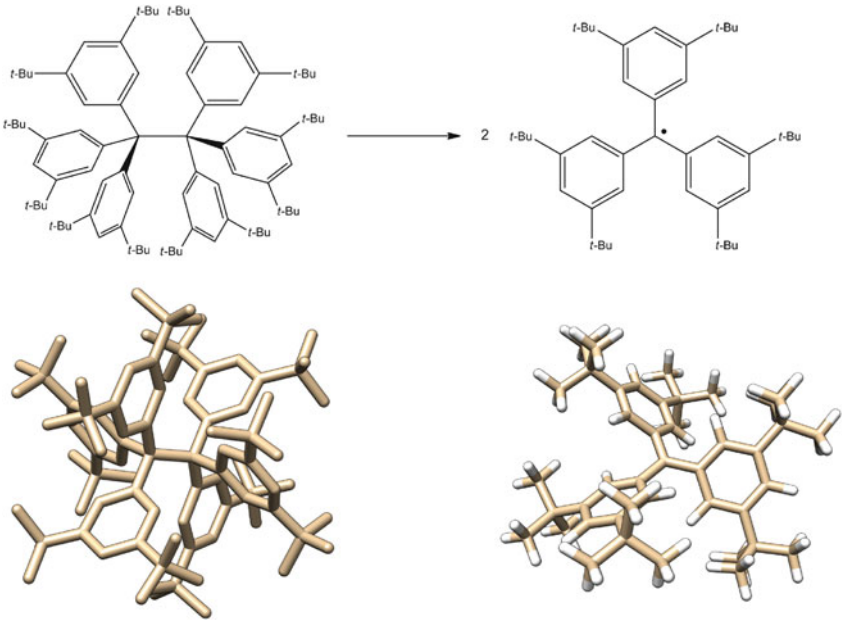


Fig. 1 Reaction formula and optimized structures for the all-meta substituted hexaphenylethane (HEXAPE). In the structure on the *lower left*, hydrogens have been omitted for clarity

spontaneously dissociate. Hexaphenylethane is known for its inability to form a stable $C(sp^3)$–$C(sp^3)$ single bond [118]. A delicate balance of covalent bonding, dispersion attraction, and Pauli repulsion forces between the phenyl rings and attached substituents can be expected when the central C–C bond is broken. However, why the parent molecule hexaphenylethane cannot be synthesized but the seemingly sterically more overcrowded all-*meta tert*-butyl substituted derivative can be isolated was an open question. It was shown that its thermodynamic stability and the instability of the parent molecule can be fully explained by dispersion-corrected DFT computations and that the *tert*-butyl groups stabilize the molecule compared to its radical fragments by as much as 40 kcal/mol. Because the system is very large (212 atoms), and involves open-shell species and various interaction types, we consider the computation of the dissociation energy ΔE as non-trivial. Unfortunately, the experimental ΔG is not precisely known. A value around zero with an error bar of ± 2 kcal/mol is compatible with the observations [118] from which, after back-correction, a gas phase ΔE value of 33 ± 4 kcal/mol for the formation of the two trityl-type radicals is deduced.

As can be seen from the results in Table 3, most functionals come close to this value, but we also note a rather large spread of the computed values (5.6–60.6 kcal/mol). Because of the missing dispersion effects in B3LYP, its error is huge (>70 kcal/mol) so that the molecule becomes unbound. In accordance with this observation, the lowest dissociation energies result from M06-2X and PBE0-DH. The former DF only includes the medium-range dispersion energy but misses the important long-range component and the DHF does not include enough dispersion by the perturbation part as noted above. The other examples discussed below support this finding and this underbinding tendency of M06-2X was also observed for non-equilibrium (stretched) van der Waals complexes [119]. Good performers with values within the error bar of the reference value are B3LYP-NL, PBE0-D3, and PW6B95-D3 with deviations of about 3 kcal/mol (10%). Considering the size and complexity of the system, this agreement between theory and experiment is very satisfactory.

3.2.2 A Zero Free Dissociation Energy Bond

It was recently shown experimentally [120] that the 2,6-di-*tert*-butyl-4-metoxyphenoxyl radical dimerizes in solution and in the solid state (Fig. 2). In the same study, a bond dissociation free Gibbs energy around zero (-0.2 ± 0.1 kcal/mol) was measured using optical and IR spectroscopies. The authors also provided results of DFT calculations in the gas phase in the supporting information. However, the calculations were done with a rather small basis set and only one DF was applied without any consideration of solvation effects. Thus, a more detailed theoretical investigation is justified and we take this example here to test our selection of modern DFs.

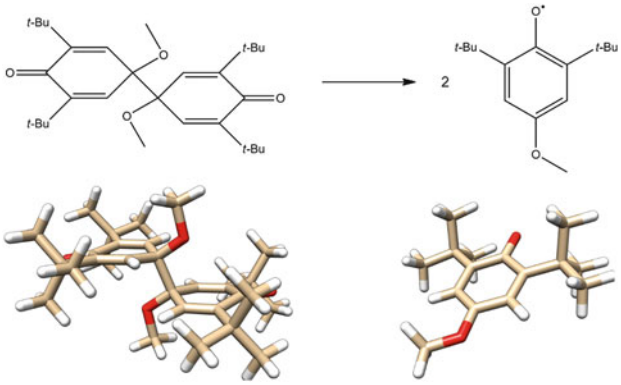


Fig. 2 Reaction formula and optimized structures for the substituted chinone dimer (DHCH)

Similar to the previous example, the dimer is stabilized by dispersion interactions between the *tert*-butyl substituents on the aromatic ring. While the stabilization here is surely smaller than in the case of the HEXAPE, the counteracting Pauli-repulsion is also smaller, resulting in a similar ΔG value in solution. As will be seen below, the DFT description of the monomers suffers more from spin-overdelocalization, which leads to an overstabilization of the monomers with respect to the dimer. This makes the system, though smaller than the previous one, very challenging.

The DFT results for the dissociation are given in Table 3. From the experimental free dissociation energy, a back-corrected value for ΔE of 17 ± 2 kcal/mol has been deduced. Thus, the central C–C bond is even weaker than in the hexaphenylethane example but, similarly, uncorrected B3LYP yields an unbound molecule due to missing dispersion effects. Good DF performers are in this case only the two range-separated DFs and one of the DHDFs with dispersion correction, namely PWPB95-D3. Even though these functionals counter the SIE to some degree, they still overestimate the stability of the radicals by 4–5 kcal/mol. It should be noted here that M06-2X performs better than most other functionals due to its large amount of Fock-exchange (54%), but cannot match the accuracy of the above-mentioned methods, which can in part be attributed to the missing long-range dispersion energy in this functional. The M06-2X value reported here deviates by about 3 kcal/mol from that given in the original publication (11.6 kcal/mol), which can be attributed to the basis set superposition error by the small basis used in the original publication leading to artificial overbinding. The strong overbinding of DSD-PBEP86-D3 by about 10 kcal/mol with respect to the reference value is probably related to the higher spin contamination of the monomer compared to the other functionals.

The system is a good example where consideration of both SIE and dispersion effects is necessary and that an accuracy of 2–3 kcal/mol or better poses a challenge for even the most sophisticated DFs.

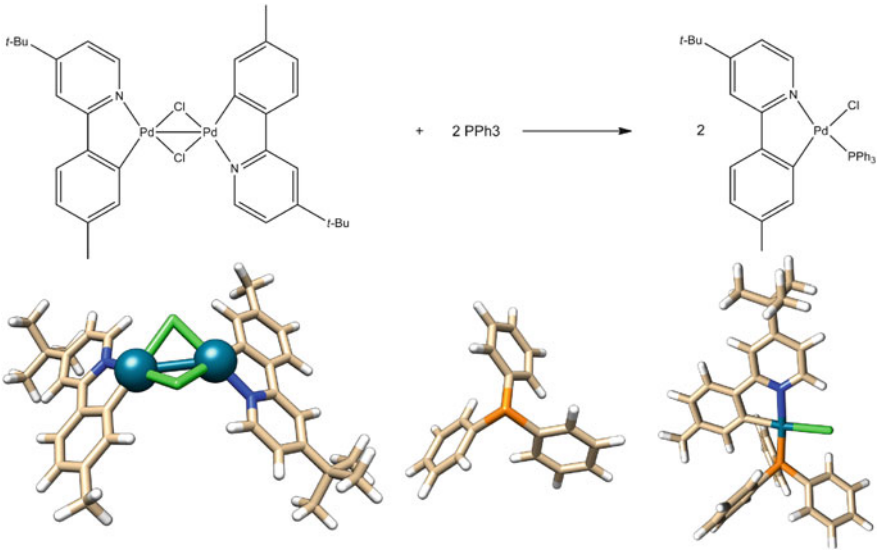


Fig. 3 Reaction formula and optimized structures for the palladium dimer ligand exchange reaction (DIMPD)

3.2.3 Transition Metal Complex

As a third example, a chemical reaction which involves a di-palladium complex was studied (see Fig. 3), for which the reaction enthalpy in solution was measured experimentally by Djukic et al. and back-corrected to pure electronic energies [41]. In this reaction, the Pd–Pd bond is quenched by a triphenylphosphane ligand, yielding the corresponding monopalladium complex. The best estimate experimental reference value for the reaction energy is $\Delta E = -32 \pm 3$ kcal/mol, whereat the largest source of error stems from the Gibbs free solvation energy correction (for a comprehensive discussion including different ligands and methods, see [41]). The relatively large uncertainty of ± 3 kcal/mol reflects the problems associated with obtaining reliable reference values for large systems, and a relative error of about $\pm 5\%$ is a realistic estimate. Hence, it should be emphasized once again that the term "chemical accuracy" has to be adjusted for the thermochemistry of large molecules which are the focus of this review. Several KS-DFT methods were tested concerning their ability to reproduce the experimental reference value of ΔE for the investigated reaction. For comparison, it should be noted that the plain HF method predicts the reaction energy qualitatively wrong (endothermic) which clearly shows that the driving force for this reaction is electron correlation. There is no strong steric cluttering around the Pd center, and thus it is easily accessible for the ligand and the base can easily approach the Lewis acid. It is also not the relatively large medium-range correlation but the significant long-range contribution to dispersion energy which renders this reaction particularly difficult. Not surprisingly, KS-DFT

methods without dispersion correction are unable to reproduce the experimental reference value, e.g., the hybrid functional B3LYP undershoots the reaction energy by more than 60% and also the M06-2X functional underbinds as noted before. The best performance of all tested methods is given by B3LYP with NL correction but the meta-hybrid PW6B95 with D3 correction and the ωB97X-V functional are also able to reproduce the reaction energy almost quantitatively. Moreover, the B3LYP-NL reaction energy is in reasonable agreement with the PW6B95-D3 result, thus indicating the D3 dispersion correction is physically sound. The two double hybrids PWPB95-D3 and DSD-PBEP86-D3 suffer from an overestimation of the reaction energy on the MP2 level and hence slightly overshoot the reaction energy. In contrast, the PBE0-DH functional, which does not include the D3 correction, underestimates the reaction energy due to the missing long-range dispersion interaction. It is encouraging to see that after describing the physics correctly, i.e., proper treatment of medium and long-range correlation effects, KS-DFT methods can reproduce the experimental reaction energy with sufficient accuracy, thus giving the right answer for the right reason. This finding is in agreement with a previous study of ligand substitution energies for transition metal complexes [23]. Thus, we are confident that the available modern KS-DFT methods are also well suited for studying the thermochemistry of larger transition metal complexes.

3.2.4 Dihydrogen Activation by an FLP

Activation of dihydrogen is typically a domain of transition metal chemistry and even Nature uses metal centered reactions to split the dihydrogen molecule in the hydrogenase enzymes. There is a recent development in the use of metal-free systems for H_2-activation. Stephan, Erker, and others have described so-called frustrated Lewis pairs (FLP), i.e., pairs of Lewis acids and bases, which do not quench each other due to the steric bulk of their substituents and heterolytically split the H_2-molecule [117, 121]. Phosphane/borane pairs such as the system considered here (see Fig. 4) react rapidly and effectively with H_2 to yield the corresponding phosphonium cation/hydridoborate anion pair [122]. These systems have been used as active metal-free hydrogenation catalysts [123] and an increasing number of related systems that appear in the literature can activate many other small molecules (e.g., alkenes, CO, CO_2, NO [124–127]). The quantum chemical description of FLP reactions has attracted a lot of interest recently [92, 128–132]. Here we study the reaction energy of the original $[B(C_6F_5)_3]/[P(t\mathrm{Bu})_3]$ system [122] with H_2. Because the thermodynamic properties have never been measured accurately (the reaction is practically irreversible in common solvents, i.e., $\Delta G \ll 0$) we take as reference the value of −21.6 kcal/mol from a DLPNO-CCSD(T)/CBS treatment which should be accurate to about 1–2 kcal/mol. Note that the reaction energy is calculated relative to the weakly

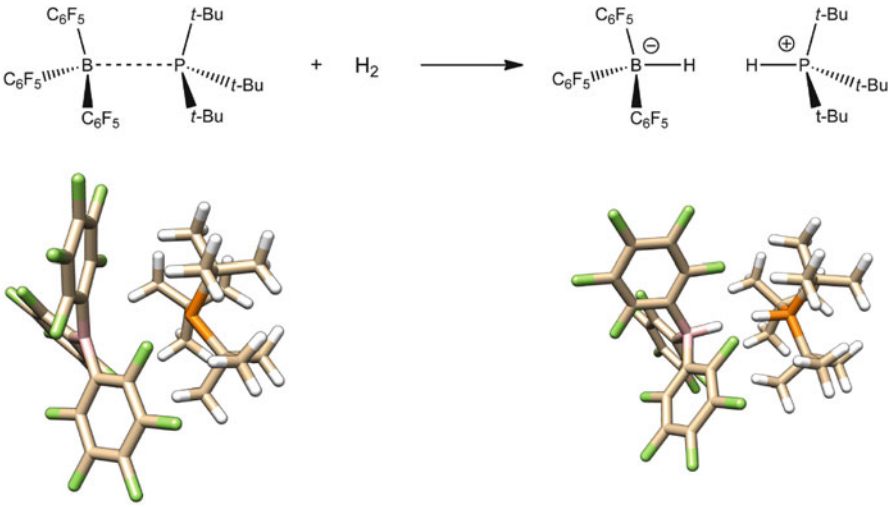


Fig. 4 Reaction formula and optimized structures for dihydrogen activation by a frustrated Lewis pair (FLPH2)

bound $[B(C_6F_5)_3]/[P(t\text{Bu})_3]$ donor-acceptor complex and separate H_2 and not with respect to all free reactands (which would lower the reaction energy by about 15 kcal/mol [129]).

As can be seen from Table 3, the DFT results are relatively close to each other and also to the reference value. The span of the values from −18.3 to −24 kcal/mol is much smaller than in the previous cases and the DFT results nicely bracket the DLPNO-CCSD(T)/CBS reference value. Although the structural changes are large, i.e., splitting of a strong single bond and formation of a zwitterionic structure with two new bond types, it seems that the FLP reaction is electronically rather easy to treat by DFT. Part of the reason for this somewhat surprising finding (which, however, has been noted before for a comparison of MP2, SCS-MP2, and B97-D methods [129]) is that the non-covalent interactions are similar in the FLP and the reaction product so that most errors in the theoretical treatment cancel. This view is supported by the relatively good results of uncorrected B3LYP. The best results with deviations of only about 1 kcal/mol are provided by ωB97X-V and DSD-PBEP86-D3. The worst performers are B3LYP-NL and M062-X with deviations of 3.3 and 2.5 kcal/mol, respectively.

4 Summary and Conclusion

The accurate description of the electronic energy part of the thermodynamic properties in large molecule reactions still represents some challenge to theory. However, the benchmarks presented and the four “real-life” chemical problems

show that significant progress has been achieved in DFT in recent years. The modern density functionals investigated from hybrid and double-hybrid rungs of "Jacob's Ladder" mostly perform very well when properly corrected for long-range London dispersion effects. The basic reason for this is that larger molecules are significantly more stabilized by the intramolecular dispersion energy than smaller ones, and in particular dissociation reactions definitely require dispersion corrections in DFT. This also explains why the long-range dispersion devoid M06-2X functional performs less well for all larger "real-life" examples. Note that these findings are not evident from still widely employed small molecule benchmark sets. Similar conclusions apply to the performance analysis of the dispersion-uncorrected PBE0-DH double-hybrid functional. In contrast, the "old-fashioned" B3LYP approach shows reasonably good performance in its tested NL(VV10)-corrected form. Somewhat better accuracy than with global hybrids – which include non-locality only for the exchange part – can be achieved with double-hybrids that include a perturbative orbital-dependent correlation energy. From the three tested variants, the PWPB95-D3 functional, which is also computationally efficient due to the use of opposite-spin orbital correlation only, shows the best overall performance. The two range-separated functionals tested from the ωB97X family provide good results in particular when the self-interaction error is relevant (e.g., for delocalized radicals or reaction barriers) but also some outliers are noted. This is tentatively attributed to the underlying Taylor-expansion of the B97-type GGA part. Nevertheless, we reiterate that the error estimates of the experimental measurements of substantial reaction energies in large molecular systems do not justify the supposition of a chemical accuracy of 1 kcal/mol. Rather, we suggest a relaxation to 2–3 kcal/mol and note that the modern functionals investigated are not too far away from this bound. Concerning just the electronic energy we conclude that DFT in combination with nowadays possible large-scale DLPNO-CCSD(T) calculations opens a bright future for theoretical thermochemistry.

However, comparisons to experimental data under typical conditions require inclusion of thermal and entropic effects in the gas phase together with corrections for solvation. Since solvation and entropic contributions almost never cancel and some reactions are entirely driven by solvation (e.g., those leading to zwitterions or ion pairs), their accurate account is mandatory. Comparison of the results from different continuum solvation models (not shown here) for large but not very polar systems indicates that an accuracy of 1–2 kcal/mol for the solvation free energy contribution to a reaction is not easy to achieve. Similar estimates are obtained for the error of the thermostatistical calculation of reaction entropies. Further work along these lines together with improved density functionals should allow routine calculations with 2 kcal/mol accuracy or better for even larger systems than treated in this review in the foreseeable future.

Acknowledgment Financial support from the DFG in various projects over the last few years is gratefully acknowledged.

References

1. Hohenberg P, Kohn W (1964) Phys Rev B 136:B864
2. Kohn W, Sham LJ (1965) Phys Rev 140:A1133
3. Parr RG, Yang W (1989) Density-functional theory of atoms and molecules. Oxford University Press, Oxford
4. Koch W, Holthausen MC (2001) A chemist's guide to density functional theory. Wiley-VCH, New York
5. Dreizler J, Gross EKU (1990) Density functional theory, an approach to the quantum many-body problem. Springer, Berlin
6. Curtiss LA, Raghavachari K, Redfern PC, Pople JA (1997) J Chem Phys 106:1063
7. Curtiss LA, Raghavachari K, Redfern PC, Rassolov V, Pople JA (1998) J Chem Phys 109:7764
8. Karton A, Tarnopolsky A, Lamère JF, Schatz GC, Martin JML (2008) J Phys Chem A 112:12868
9. Jurečka P, Šponer J, Černý J, Hobza P (2006) Phys Chem Chem Phys 8:1985
10. Řezáč J, Riley KE, Hobza P (2011) J Chem Theor Comput 7:2427
11. Řezáč J, Riley KE, Hobza P (2012) J Chem Theor Comput 8:4285
12. Zhao Y, González-García N, Truhlar DG (2005) J Phys Chem A 109:2012
13. Zhao Y, Ng HT, Peverati R, Truhlar DG (2012) J Chem Theor Comput 8:2824
14. Grimme S, Steinmetz M, Korth M (2007) J Org Chem 72:2118
15. Huenerbein R, Schirmer B, Moellmann J, Grimme S (2010) Phys Chem Chem Phys 12:6940
16. Korth M, Grimme S (2009) J Chem Theor Comput 5:993
17. Goerigk L, Grimme S (2011) J Chem Theor Comput 7:291
18. Averkiev BB, Zhao Y, Truhlar DG (2010) J Mol Cat A Chem 324:80
19. Zhao Y, Truhlar DG (2006) J Chem Phys 124:224105
20. Schultz NE, Zhao Y, Truhlar DG (2007) J Comput Chem 29:185
21. Jiang W, Laury ML, Powell M, Wilson AK (2012) J Chem Theor Comput 8:4102
22. Hughes TF, Harvey JN, Friesner RA (2012) Phys Chem Chem Phys 14:7724
23. Grimme S (2012) ChemPhysChem 13:1407
24. Steinmetz M, Grimme S (2013) ChemistryOpen 2:115
25. Pérez-Jordá JM, Becke AD (1995) Chem Phys Lett 233:134
26. Kristyán S, Pulay P (1994) Chem Phys Lett 229:175
27. Hobza P, Šponer J, Reschel T (1995) J Comput Chem 16:1315
28. Grimme S, Antony J, Schwabe T, Mück-Lichtenfeld C (2007) Org Biomol Chem 5:741
29. Klimeš J, Michaelides A (2012) J Chem Phys 137:120901
30. Grimme S (2011) Wiley Interdiscip Rev Comput Mol Sci 1:211
31. Grimme S, Huenerbein R, Ehrlich S (2011) ChemPhysChem 12:1258
32. Rappoport D, Crawford NRM, Furche F, Burke K (2009) In: Solomon EI, Scott RA, King RB (eds) Computational inorganic and bioinorganic chemistry. Wiley-VCH, New York, pp 159–172
33. Klamt A (1995) J Phys Chem 99:2224
34. Eckert F, Klamt A (2002) AIChE J 48:369
35. Klamt A (2011) Wiley Interdiscip Rev Comput Mol Sci 1:699
36. Becke AD (1988) Phys Rev A 38:3098
37. Perdew JP (1986) Phys Rev B 33:8822
38. Schäfer A, Huber C, Ahlrichs R (1994) J Chem Phys 100:5829
39. Grimme S (2012) Chem Eur J 18:9955
40. Riplinger C, Sandhoefer B, Hansen A, Neese F (2013) J Chem Phys 139:134101
41. Hansen A, Djukic JP, Bannwarth C, Grimme S, to be published
42. Császár AG, Allen WD, Schaefer HF III (1998) J Chem Phys 108:9751
43. Hobza P, Šponer J (2002) J Am Chem Soc 124:11802
44. Weigend F, Ahlrichs R (2005) Phys Chem Chem Phys 7:3297

45. Perdew JP (1999) In: Geerlings P, Proft FD, Langenaeker W (eds) The functional zoo. VUB University Press, Brussel, pp 87–109
46. Becke AD (1993) J Chem Phys 98:1372
47. Becke AD (1993) J Chem Phys 98:5648
48. Perdew JP, Schmidt K (2001) In: Doren VEV, Alsenoy CV, Geerlings P (eds) Density functional theory and its applications to materials, AIP conference proceedings, vol 577. American Institute of Physics, New York
49. Perdew JP, Ruzsinszky A, Tao J, Staroverov VN, Scuseria GE, Csonka GI (2005) J Chem Phys 123:062201
50. Peverati R, Truhlar DG (2014) Phil Trans R Soc A 372:20120476
51. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994) J Phys Chem 98:11623
52. Perdew JP, Burke K, Ernzerhof M (1996) Phys Rev Lett 77:3865, (1997) Erratum Phys Rev Lett **77**:1396
53. Adamo C, Barone V (1999) J Chem Phys 110:6158
54. Csonka GI, Perdew JP, Ruzsinszky A (2010) J Chem Theor Comput 6:3688
55. Grimme S, Antony J, Ehrlich S, Krieg H (2010) J Chem Phys 132:154104
56. Grimme S, Ehrlich S, Goerigk L (2011) J Comput Chem 32:1456
57. Vydrov OA, van Voorhis T (2010) J Chem Phys 133:244103
58. Goerigk L, Grimme S (2011) Phys Chem Chem Phys 13:6670
59. Hujo W, Grimme S (2011) J Chem Theor Comput 7:3866
60. Iikura H, Tsuneda T, Yanai T, Hirao K (2001) J Chem Phys 115:3540
61. Chai JD, Head-Gordon M (2008) Phys Chem Chem Phys 10:6615
62. Lin YS, Li GD, Mao SP, Chai JD (2013) J Chem Theor Comput 9:263
63. Zhao Y, Lynch BJ, Truhlar DG (2003) J Phys Chem A 108:4786
64. Neese F, Schwabe T, Grimme S (2007) J Chem Phys 126:124115
65. Grimme S (2006) J Chem Phys 124:034108
66. Zhao Y, Lynch BJ, Truhlar DG (2005) Phys Chem Chem Phys 7:43
67. Ángyán JG, Gerber IC, Savin A, Toulouse J (2005) Phys Rev A 72:012510
68. Grimme S (2003) J Chem Phys 118:9095
69. Grimme S, Goerigk L, Fink RF (2012) Wiley Interdiscip Rev Comput Mol Sci 2:868
70. Kozuch S, Martin JML (2013) J Comput Chem 34:2327
71. Becke AD, Johnson ER (2005) J Chem Phys 123:154101
72. Tkatchenko A, Scheffler M (2009) Phys Rev Lett 102:073005
73. Tkatchenko A, DiStasio RA Jr, Car R, Scheffler M (2012) Phys Rev Lett 108:236402
74. Dion M, Rydberg H, Schröder E, Langreth DC, Lundqvist BI (2004) Phys Rev Lett 92:246401
75. Lee K, Murray ED, Kong L, Lundqvist BI, Langreth DC (2010) Phys Rev B 82:081101
76. Casimir HBG, Polder D (1948) Phys Rev 73:360
77. Stone AJ (1997) The theory of intermolecular forces. Oxford University Press, Oxford
78. Ehrlich S, Moellmann J, Reckien W, Bredow T, Grimme S (2011) ChemPhysChem 12:3414
79. Johnson ER (2011) J Chem Phys 135:234109
80. Koide A (1976) J Phys B At Mol Phys 9:3173
81. Ruzsinszky A, Perdew JP (2011) Comput Theor Chem 963:2
82. Cohen AJ, Mori-Sánchez P, Yang W (2012) Chem Rev 112:289
83. Mardirossian N, Head-Gordon M (2013) Phys Chem Chem Phys. doi:10.1039/C3CP54374A. Accepted Manuscript
84. Kozuch S, Martin JML (2011) Phys Chem Chem Phys 13:20104
85. Brémond E, Adamo C (2011) J Chem Phys 135:024106
86. Zhao Y, Truhlar DG (2008) Acc Chem Res 41:157
87. Grimme S (2006) Angew Chem Int Ed 45:4460
88. Zhang Y, Yang W (1998) J Chem Phys 109:2604
89. Gritsenko O, Ensing B, Schipper PRT, Baerends EJ (2000) J Phys Chem A 104:8558
90. Mori-Sánchez P, Cohen AJ, Yang W (2006) J Chem Phys 125:201102

91. Dunning TH Jr (1989) J Chem Phys 90:1007
92. Schirmer B, Grimme S (2013) Top Curr Chem 332:213
93. Grimme S, Steinmetz M (2013) Phys Chem Chem Phys 15:16031
94. Andrae D, Häußermann U, Dolg M, Stoll H, Preuss H (1990) Theor Chim Acta 77:123
95. Halkier A, Helgaker T, Jørgensen P, Klopper W, Koch H, Olsen J, Wilson AK (1998) Chem Phys Lett 286:243
96. Halkier A, Helgaker T, Jørgensen P, Klopper W, Olsen J (1999) Chem Phys Lett 302:437
97. Baerends EJ, Ellis DE, Ros P (1973) Chem Phys 2:41
98. Dunlap BI, Connolly JWD, Sabin JR (1979) J Chem Phys 71:3396
99. Eichkorn K, Treutler O, Öhm H, Häser M, Ahlrichs R (1995) Chem Phys Lett 240:283
100. Eichkorn K, Weigend F, Treutler O, Ahlrichs R (1997) Theor Chem Acc 97:119
101. Weigend F (2006) Phys Chem Chem Phys 8:1057
102. Weigend F, Häser M (1997) Theor Chem Acc 97:331
103. Weigend F, Köhn A, Hättig C (2002) J Chem Phys 116:3175
104. Hohenstein EG, Chill ST, Sherrill CD (2008) J Chem Theor Comput 4:1996
105. Mardirossian N, Head-Gordon M (2013) J Chem Theor Comput 9:4453
106. Ahlrichs R et al (2009) Universität Karlsruhe. See http://www.turbomole.com
107. Ahlrichs R, Bär M, Häser M, Horn H, Kölmel C (1989) Chem Phys Lett 162:165
108. Neese F (2013) ORCA—an ab initio, density functional and semiempirical program package, Ver. 3.0 (Rev 0). Max Planck Institute Energy Conversion Chemistry, Germany
109. Neese F (2012) Wiley Interdiscip Rev Comput Mol Sci 2:73
110. Eckert F, Klamt A (2010) COSMOtherm, Version C2.1, Release 01.11. COSMOlogic GmbH & Co. KG, Leverkusen, Germany
111. Goerigk L, Grimme S (2010) J Chem Theor Comput 6:107
112. Burns LA, Vázquez-Mayagoitia A, Sumpter BG, Sherrill CD (2011) J Chem Phys 134:084107
113. Risthaus T, Grimme S (2013) J Chem Theor Comput 9:1580
114. Zhao Y, Lynch BJ, Truhlar DG (2004) J Phys Chem A 108:2715
115. Luo S, Zhao Y, Truhlar DG (2011) Phys Chem Chem Phys 13:13683
116. Kruse H, Goerigk L, Grimme S (2012) J Org Chem 77:10824
117. Stephan DW, Erker G (2010) Angew Chem Int Ed 49:46
118. Grimme S, Schreiner PR (2011) Angew Chem Int Ed 50:12639
119. Goerigk L, Kruse H, Grimme S (2011) ChemPhysChem 12:3421
120. Wittman JM, Hayoun R, Kaminsky W, Coggins MK, Mayer JM (2013) J Am Chem Soc 135:12956
121. Stephan DW (2008) Org Biomol Chem 6:1535
122. Welch GC, San Juan RR, Masuda JD, Stephan DW (2006) Science 314:1124
123. Stephan DW (2009) Dalton Trans 3129–3136
124. Mömming CM, Frömel S, Kehr G, Fröhlich R, Grimme S, Erker G (2009) J Am Chem Soc 131:12280
125. Sajid M, Elmer LM, Rosorius C, Daniliuc CG, Grimme S, Kehr G, Erker G (2013) Angew Chem Int Ed 52:2243
126. Mömming CM, Otten E, Kehr G, Fröhlich R, Grimme S, Stephan D, Erker G (2009) Angew Chem Int Ed 48:6643
127. Cardenas AJP, Culotta BJ, Warren TH, Grimme S, Stute A, Fröhlich R, Kehr G, Erker G (2011) Angew Chem Int Ed 50:7567
128. Rokob TA, Hamza A, Pápai I (2009) J Am Chem Soc 131:10701
129. Grimme S, Kruse H, Goerigk L, Erker G (2010) Angew Chem Int Ed 49:1402
130. Rokob TA, Hamza A, Stirling A, Soós T, Pápai I (2008) Angew Chem Int Ed 47:2435
131. Rokob TA, Bakó I, Stirling A, Hamza A, Pápai I (2013) J Am Chem Soc 135:4425
132. Pu M, Privalov T (2013) J Chem Phys 138:154305
133. Zhao Y, Truhlar DG (2005) J Phys Chem A 109:5656
134. Zhao Y, Truhlar DG (2008) Theor Chem Acc 120:215

Top Curr Chem (2015) 365: 25–52
DOI: 10.1007/128_2014_555

Published online: 12 May 2014

Density Functional Theory Beyond the Generalized Gradient Approximation for Surface Chemistry

Benjamin G. Janesko

Abstract Density functional theory with generalized gradient approximations (GGAs) for the exchange-correlation density functional is widely used to model adsorption and reaction of molecules on surfaces. In other areas of computational chemistry, GGAs have largely been replaced by more accurate meta-GGA and hybrid approximations. Meta-GGAs and hybrids can ameliorate GGAs' systematic over-delocalization of electrons and systematic underestimate of reaction barriers. This chapter discusses extensions of meta-GGAs and screened hybrids to surface chemistry. It reviews evidence that GGAs' systematic underestimate of gas-phase reaction barriers carries over to reactions on surfaces, and that meta-GGAs and screened hybrids can improve results. It closes with recent applications and new work towards more accurate functionals for surfaces. These promising results motivate further exploration of meta-GGAs and screened hybrids for surface chemistry.

Keywords GGA • Meta-GGA • Reaction barrier • Screened hybrids • Surface chemistry

Contents

B.G. Janesko (✉)
Texas Christian University, 2800 S. University Dr., Fort Worth, TX 76129, USA
e-mail: b.janesko@tcu.edu

1 Introduction

1.1 Chemistry on Surfaces

Chemical reactions at surfaces and interfaces are central to many problems in chemistry. Topical examples include heterogeneous catalysis [1], surface-enhanced spectroscopies [2], battery technologies and hydrogen storage [3], and nanoscale devices [4]. Experimental studies of surface chemistry can be challenging. Measurements of molecules' adsorption, desorption, and reaction on single crystal surfaces under ultrahigh vacuum have yielded important insights [1] recognized by the 2007 Nobel Prize in Chemistry [5]. However, connections to practical surface chemistries often requires bridging the "pressure gap" between ultrahigh vacuum and industrially relevant high pressures, and the "materials gap" between single crystal surfaces and industrially relevant nanoparticles or catalysts. These gaps can play critical roles in surface chemistry. To illustrate, the Fischer–Tropsch process [6] for converting synthesis gas ($CO + H_2$) to long-chain hydrocarbons is catalyzed industrially by promoted and nanostructured cobalt [7] and iron [8] surfaces, but does not occur on single-crystalline surfaces under ultrahigh vacuum [9]. While there has been substantial progress in experiments bridging these gaps [10–12], experimental surface science remains challenging. Figure 3 of Maitlis [13] illustrates the contrast between well-defined experiments on homogeneous catalysts and the more "impressionistic" data available for heterogeneous catalysts.

1.2 Simulating Surface Chemistry

Simulations of molecules on surfaces have become essential tools for interpreting this impressionistic experimental data [14]. Such simulations typically require electronic structure calculations. Simulations of surface-enhanced spectroscopies combine classical electrodynamics models [15] with electronic structure calculations modeling how molecule–surface interactions shift a molecule's vibrational spectrum [16]. Simulations of nanoscale device chemistry generally center on modeling the electronic structure and properties of model devices [17]. Simulations of heterogeneous catalysis typically begin with electronic structure calculations on the geometries and adsorption energies of reaction intermediates on catalyst surfaces. The resulting potential energy surfaces can be applied in microkinetics models of reactions at realistic pressures, temperatures, and catalyst compositions [18–21].

Electronic structure calculations generally model a surface as either a periodic slab [22] or a finite cluster of surface atoms [23], potentially embedded in a simpler model background [24]. Cluster models are widely used for insulators, and are readily applicable to charged species. However, cluster models for metal surfaces often converge slowly with cluster size [25]. (Differences between calculations at different levels of theory can converge more rapidly [26].) Electronic structure approximations applicable to periodic slabs are desirable for treating metal, semiconductor, and insulator surfaces on an equal theoretical footing.

1.3 Choice of Electronic Structure Approximation

Electronic structure calculations on molecules, solids, and surfaces typically must approximate the many-body electron–electron interactions [27]. To be useful, an approximation should be both accurate enough to make experimentally meaningful predictions, and computationally inexpensive enough to treat the chosen system in a reasonable time using available computational resources. Several different electronic structure approximations have been applied to surfaces. One route uses first principles ab initio approximations for the many-electron wavefunction. Recent ab initio calculations on surfaces include coupled cluster [28–31] simulations of clusters of surface atoms, and quantum Monte Carlo (QMC) [26, 32–35] and random phase approximation (RPA) [36–42] simulations of periodic slabs. Unfortunately, ab initio methods typically have rather steep computational scaling [27], making them problematic for large and realistic model surfaces. A second route is to use tight-binding Hamiltonians. These can be applied to large systems, but require substantial empirical parameterization and can have limited transferability [43].

2 GGAs for Surface Chemistry

2.1 Density Functional Theory

The vast majority of electronic structure calculations on surfaces use Kohn–Sham density functional theory (DFT) [44, 45]. DFT can often provide acceptable accuracy at modest computational expense. Figure 1 illustrates DFT's dominance in modeling surface chemistry.

Kohn–Sham DFT models the ground state of an N-electron system as a reference system of N noninteracting Fermions corrected by a mean-field Hartree interaction,

$$E_{\mathrm{H}}[\rho] = \frac{1}{2}\int d^3\mathbf{r}\int d^3\mathbf{r}'\,\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}'-\mathbf{r}|}, \tag{1}$$

and an exchange-correlation (XC) density functional $E_{\mathrm{XC}}[\rho]$ containing all many-body effects. $\rho(\mathbf{r})$ is the probability density for finding an electron at $\mathbf{r}$. The Kohn–Sham reference system has the same $\rho(\mathbf{r})$ as the real system by construction. The exact $E_{\mathrm{XC}}[\rho]$ is a unique and variational functional of $\rho(\mathbf{r})$ [44]. The Hamiltonian of the non-interacting Fermions includes local and multiplicative Hartree and XC potentials, e.g., $v_{\mathrm{XC}}(\mathbf{r}) = \delta E_{\mathrm{XC}}[\rho]/\delta\rho(\mathbf{r})$. Practical calculations typically use separate ↑- and ↓-spin densities and Kohn–Sham orbitals. This work suppresses spin dependence for conciseness. All orbitals, densities, density matrices, and exchange energies are assumed to be spin polarized.

The accuracy and computational expense of a typical DFT calculation is determined by the one-electron basis set used to expand the reference system's wavefunction, by any approximate treatments of relativistic effects, reciprocal space, and core electrons, and by the choice of approximate XC functional. DFT's success is

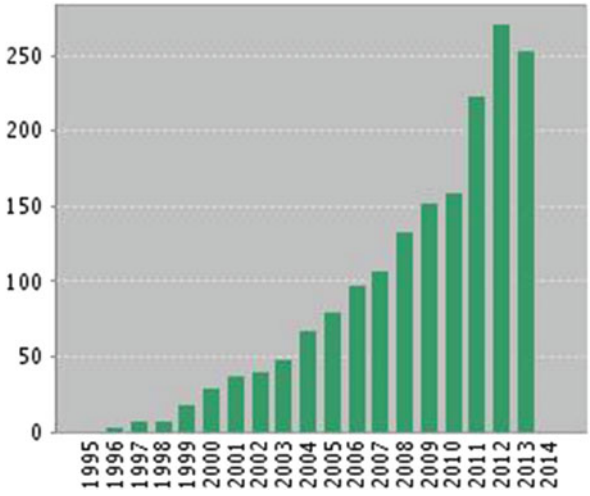


Fig. 1 Results of a Web of Science search for "DFT catalyst* surface*" illustrating published applications of DFT to heterogeneous catalysis [46]

largely attributable to the development of accurate and computationally tractable XC approximations [47]. Systematically convergent hierarchies of XC approximations (ab initio DFT) have been developed [48]. However, the vast majority of DFT calculations use an alternative "Jacob's Ladder" of approximations [49] based on the homogeneous electron gas (HEG, $\rho(\mathbf{r}) = \text{constant}$). The first "rung" of this ladder is the local spin-density approximation (LSDA) [50, 51] constructed to reproduce the numerically exact XC energy density [52, 53] of the HEG.

Generalized gradient approximations occupy the second rung of Jacob's Ladder. GGAs model the nonunique [54] XC energy density at point $\mathbf{r}$ as a function of $\rho(\mathbf{r})$ and its gradient:

$$E_{\mathrm{XC}}^{\mathrm{GGA}}[\rho] = \int d^3\mathbf{r}\, e_{\mathrm{XC}}^{\mathrm{GGA}}\left(\rho(\mathbf{r}), \left|\nabla\rho(\mathbf{r})\right|\right). \tag{2}$$

GGAs tend to improve upon the LSDA for total energies, atomization energies, and reaction barriers, and expand and soften bonds to (over)correct the LSDA's overbinding [55]. Their improved accuracy and computational simplicity makes them widely applied to periodic slab models for surfaces.

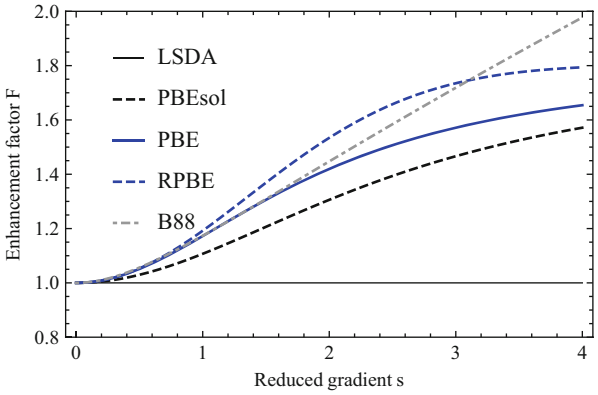
2.2 Design of GGAs

Unlike the LSDA, there is no single "best" choice of a GGA [55]. This flexibility has been exploited to construct GGAs which perform well for specific aspects of surface chemistry. In practice, a GGA's performance is often largely a function of its exchange enhancement factor F_{X}:

$$E_{\mathrm{X}}^{\mathrm{GGA}}[\rho] = \int d^3\mathbf{r}\, C_{\mathrm{X}}\rho^{4/3}(\mathbf{r})F_{\mathrm{X}}(s). \tag{3}$$

The "exchange" portion of $E_{\mathrm{XC}}[\rho]$ (4) is defined in terms of the expectation value of the electron–electron interaction operator, evaluated with the wavefunction of the noninteracting Kohn–Sham reference system [56]. Coefficient $C_{\mathrm{X}} = -\frac{3}{4}\left(\frac{6}{\pi}\right)^{1/3}$ is from the exact Kohn–Sham wavefunction of the HEG [57]. $s = |\nabla\rho(\mathbf{r})|/(2(6\pi^2)^{1/3}\rho^{4/3}(\mathbf{r}))$ is the unitless reduced density gradient.

Figure 2 illustrates representative GGAs' exchange enhancement factors and their performance for some properties relevant to surface chemistry. The LSDA is included for comparison. Similar to the LSDA, GGAs with small enhancement factors (PBEsol [62], AM05 [63], Wu-Cohen [64]) combine accurate lattice parameters with overestimated adsorption energies [65]. GGAs with larger enhancement factors (revPBE [66], RPBE [67], BLYP [68, 69]) improve adsorption energies [67] and molecular thermochemistry [70] at the expense of lattice parameters and geometries [71]. PBE [55] and PW91 [50, 72] provide intermediate performance



Functional	Lattice (pm) Ref. [58]	Bulk modulus (GPa)	CO/Pt(111) Eads (eV)	Barriers (kcal/mol)	GMTKN30 (kcal/mol)
LSDA	-4.59	35.74		17.5	12.0
PBEsol	-1.98	28.79	-2.1	13.2	7.4
PW91	1.95	7.46			
PBE	2.10	5.21	-1.6	10.5	5.2
RPBE,revPBE	4.19	-7.89	-1.3	8.1	4.5
BLYP			-1.0	9.6	5.3

Fig. 2 *Top*: Exchange enhancement factors of the LSDA and representative GGAs (F_X, Eq. (3)). *Bottom*: Mean errors in the shortest interatomic distances of 30 transition metals [58]; mean error in bulk moduli of 30 transition metals [58]; predicted chemisorption energies for CO on Pt(111) [60]; root-mean-square deviation in BH76 barrier heights and weighted mean absolute deviation in the GMTKN30 database of gas-phase properties (kcal/mol) [61]. Functionals are approximately sorted in order of increasing enhancement factor

[59]. None of these GGAs perform well for gas-phase reaction barrier heights, a fact central to what follows.

2.3 *Dispersion Corrected GGAs*

The locality of Eq. (2) prevents it from treating truly nonlocal correlation effects, including the asymptotic van der Waals interaction between distant closed-shell uncharged fragments [73]. Over the last decade, substantial resources have been invested into dispersion corrections for approximate XC functionals [74–76]. Dispersion corrections generally improve GGA simulations of molecule-surface adsorption [77], particularly for larger molecules such as coronene [78] and perylene derivatives [79, 80]. Dispersion corrections can be critical for some catalytic processes [37], some dissociation barriers [81], and some reactions of adsorbed molecules [82]. Dispersion corrections to adsorption energies are also important for apparent activation barriers on surfaces [83, 84]. However, dispersion

corrections generally do not correct GGAs' systematic underestimate of gas-phase reaction barriers. This can be seen, for example, in Tables S30–S31 of Goerigk and Grimme [85], where dispersion corrections slightly degrade PBE's performance for the BH76 benchmark set of gas-phase reaction barriers [85–87]. This assertion is supported by the widespread adoption of dispersion-corrected beyond-GGA functionals in computational chemistry [88–92]. Similar results are seen for reaction barriers of molecules on surfaces, evaluated with dispersion-corrected GGAs (Svelle et al. [93], Sect. 4). The studies reviewed below suggest that improving DFT for surface chemistry requires both dispersion corrections and beyond-GGA functional forms.

3 Successes of GGAs for Surface Chemistry

DFT calculations with generalized gradient XC functionals have been applied to an enormous array of problems in surface science. A comprehensive discussion of this literature would extend to several volumes, and is far beyond the scope of this work. This section presents a few representative successes connected to our work. Hafner [94], Greeley et al. [95], and Nørskov et al. [96] provide more extensive recent reviews.

GGAs are widely applied to determine how surface functionalization of nanostructures controls their properties. DFT calculations have mapped out the relationship between edge functionalization and electronic properties of graphene nanoribbons [4, 17, 97–99], have predicted how adsorbates open a bandgap in graphene electronics [100, 101], and have motivated studies of other quasi-1D semiconductors [102]. GGA calculations have also provided mechanistic insight into ammonia adsorption and reaction on Si(100) surfaces, a process relevant to chemical vapor deposition of silicon nitride for integrated circuits [103–107].

GGAs are also extensively applied to heterogeneous catalysis. Microkinetics models of ammonia synthesis over ruthenium nanoparticles, constructed from the RPBE and PW91 GGAs, reproduce experimental turnover frequencies to within an order of magnitude [108]. The calculated mechanisms also clarify the role of atomic steps for the rate-limiting N_2 dissociation [109]. Recent GGA calculations give new evidence for H_2-induced CO dissociation on Fischer–Tropsch catalysts [21, 110–114], a mechanism appearing to show significant structure sensitivity [115]. Studies of the C–C coupling step in the Fischer–Tropsch reaction point to the importance of surface carbide species [116]. Several recent simulations of the Fischer–Tropsch reaction considered the roles of adsorbed promoters which stabilize corrugated surfaces [117] and block graphitization [118], and adsorbed sulfur poisons which block CO dissociation [119]. These results have been expanded into complete microkinetics models of the Fischer–Tropsch reaction [21, 114, 120–123]. Other studies have considered methanol synthesis and the water-gas shift reaction [124] over some industrially relevant catalysts [18, 125, 126]. A recent combined computational and experimental study of methanation over nickel catalysts further

illustrates the value of GGA calculations for interpreting experiment [112]. GGA calculations have also provided atomic-scale explanations for the catalytic activity of gold nanoparticles, particularly bound to defective metal oxide supports [127–131]. Newer GGAs resolve the "CO/Pt(111) puzzle," in which standard GGAs' over-delocalization leads to qualitatively incorrect site dependence for CO adsorption on coinage metals [132]. Perhaps most importantly, GGA calculations have yielded general insights into periodic trends in adsorption [133, 134] and reactivity [135, 136] on catalyst surfaces.

A particularly interesting set of recent studies use GGA calculations to *design* new heterogeneous catalysts. Such studies often begin by identifying descriptors, such as adsorption or dissociation energies, which can be correlated with a catalyst's overall performance [137]. Typical calculations yield a "volcano plot" of catalytic activity vs descriptor, with optimal catalysts having intermediate values of the adsorption or dissociation energy [138]. GGA calculations of the descriptor on many model catalysts are then used to identify optimal candidates. GGA calculations on CO adsorption energies predicted that nickel-iron alloy catalysts could outperform more expensive pure Ni for CO methanation [139], a prediction subsequently verified by experiment [140]. GGA calculations were used to identify methylene chemisorption energies as a good descriptor for ethylene hydrogenation, and subsequently to identify novel nickel-zinc alloy catalysts [141]. GGA calculations on H_2 chemisorption energies were also used to identify near-surface alloy hydrogenation catalysts [142]. Nørskov et al. [143] reviews this active field.

4 Limitations of GGAs for Surface Chemistry

The successful applications reviewed above are all the more remarkable given the limitations of the GGA form. The large errors for gas-phase reaction barriers in Fig. 2 are not a special case, but are a general property of GGAs. GGAs systematically over-delocalize electrons [144–146] and overstabilize systems such as the stretched bonds of transition states. These errors are well known in the computational chemistry literature [86, 87, 147–154]. They have led to GGAs being almost entirely superseded in the computational chemistry community. (See, for example, the discussion of Fig. 1 in Burke [47].) The XC functionals which replaced them are discussed in Sect. 5.

The aforementioned difficulties of surface chemistry mean that there are relatively few experimental or computed benchmarks available to test GGAs' performance on surfaces. GGAs' limitations for surface chemistry are thus less well-characterized, and arguably less appreciated, than their limitations for molecules. However, available evidence strongly suggests that GGAs' errors for gas-phase barriers carry over to reactions of molecules on surfaces. The LSDA and the BP86 [50, 68] and BLYP GGAs underestimate QCISD(T) barriers for H_2 dissociation on

gas-phase silanes and a cluster model for Si(100) [155]. The PW91 GGA underestimates experimental [156] and quantum Monte Carlo [26] adsorption barriers for H_2 dissociative adsorption on Si(100). The PW91, PBE, and RPBE GGAs all underestimate the QMC barrier for H_2 dissociation on Mg(0001) [32]. PBE also underestimates QMC barriers for hydrogen abstraction by styrene radical on hydrogen-terminated Si(001) [33]. A series of studies by Bickelhaupt and coworkers show similar GGA errors for small organic molecules reacting with atomic Pd [157–159]. We found comparable errors in several GGAs' predicted dissociation barriers for H_2 dissociation on Au_3 and Ag_3 clusters [160]. The BP86 GGA severely underestimates completely renormalized coupled cluster reaction barriers for methanol oxidation on Au_8^- [31]. PW91 underestimates both coupled cluster reaction barriers for water splitting on an Fe atom, and RPA barriers for water splitting on Fe(100) [41]. PBE underestimates the barrier to O_2 sticking on a cluster model of Al(111) [161], and PBE and RPBE incorrectly predict barrierless O_2 dissociation on Al(111) slabs [162]. (While this is attributed to spin selection rules [162], recent hybrid DFT calculations [163] suggest that the GGA's limitations also play a role.) PBE and PW91 also incorrectly predict barrierless dissociation of H_2O_2 on cluster models of zirconium, titanium, and yttrium oxide surfaces [164]. Dispersion-corrected PBE calculations systematically underestimate the experimental reaction barrier to alkene methylation over a slab model zeolite catalyst [93]. We showed that the LSDA, the PBE, and revPBE GGAs underestimate the coupled cluster barrier for NH_3 dissociation on a cluster model of the reconstructed Si(100) surface, with differences between GGAs and beyond-GGA functionals persisting on larger clusters and periodic slabs [165]. GGAs also incorrectly predict the relative barriers to inter- vs intra-dimer NH_3 dissociation [104–106, 166, 167]. The LSDA and the PBE, PW91, and revPBE GGAs tend to underestimate diffusion Monte Carlo calculations for diffusion barriers (i.e., adsorption energies at different sites [168]) of adatoms on graphene [34, 169]. PBE also underestimates accurate reaction barriers for hydrogenation of graphene model compounds [30]. Other relevant GGA errors include overbinding of Cu on cluster models of the MgO(001) surface [170], and qualitatively incorrect spin distributions for defects on titania [171] and ceria [172] surfaces.

A particularly instructive illustration of GGAs' strengths and limitations comes from the aforementioned careful and insightful study of ammonia synthesis over ruthenium [108]. As discussed above, microkinetics models constructed from the PW91 and RPBE GGAs both predicted overall turnover frequencies within an order of magnitude of experiment. However, PW91 predicted that the rate-limiting N_2 dissociation barrier was 0.6 eV lower than the RPBE barrier. This corresponds to an enormous ten orders of magnitude discrepancy in the room-temperature Arrhenius rate constant. At least one of the predicted mechanisms thus enjoyed substantial error cancellation between inaccurate adsorption energies and reaction barriers. The authors explicitly characterized this error cancellation, stating that PW91 calculations "increased the coverage ... and decreased the number of free sites for dissociation" [108]. While such error cancellations are acceptable for some

applications, improvements are clearly desirable. This result illustrates the handicaps faced by computational surface scientists, and motivates the development of new approximations.

Several groups have attempted to remedy these issues with new GGAs. Some explore parameterized GGA forms similar to those pioneered in Becke [173]. The BEEF-vdW dispersion-corrected GGA incorporates 31 empirical parameters fitted using Bayesian error estimation [174]. Calculations using this functional accurately treat a wide variety of properties, from small-molecule heats of formation and noncovalent interactions to lattice constants, bulk moduli, and chemisorption energies [174]. Applications to metal–carbon bond formation from hydrogenation of supported graphene [175] and chemisorption on zeolites [176] leverage its strengths for noncovalent interactions. A study of CO_2 hydrogenation even points to improvements over RPBE for some reaction barriers on surfaces [177]. However, BEEF-vdW still gives a 0.26 eV mean absolute error in representative gas-phase reaction barriers [150], comparable to RPBE (0.27 eV), and significantly larger than the B3LYP global hybrid (0.17 eV) [174]. This is especially noteworthy given that B3LYP is, for a hybrid functional, not particularly accurate for reaction barriers. (To illustrate, the HISS screened hybrid discussed in Sect. 5.3 gives mean absolute errors of 1.7 and 1.8 kcal mol^{-1} in the HTBH38/04 and NHTBH38/04 test sets of gas-phase reaction barriers [86, 87], significantly smaller than B3LYP errors 4.23 and 4.34 kcal mol^{-1} obtained with a different computational setup [87].) The SOGGA11 [178] and non-separable N12 [179] GGAs, which respectively incorporate 18 and 24 empirical parameters, also give fairly large errors in these test sets [180]. These errors are dramatically reduced by the empirical meta-GGAs discussed in Sect. 5.4 [180].

"Specific reaction parameter" GGAs interpolating between, for example, PW91 and RPBE have also been proposed [181]. The interpolations are not guaranteed to be transferable, and require fitting to known experimental values. Interpolations can also be problematic where the known experimental value is not bracketed by two different GGAs [182]. GGAs' systematic underestimate of reaction barriers appears to make this circumstance rather common. Indeed, specific reaction parameter functionals originally tuned a hybrid functional's GGA term and fraction of exact exchange [183], exploiting the effects discussed in Sect. 5.1. Unfortunately, it appears that this "Procrustean dilemma" (Perdew et al. [62]) is an inherent limitation of the GGA form. This fact motivates exploration of methods beyond the GGA.

5 Beyond the GGA

DFT calculations on small and medium-sized molecules almost exclusively use beyond-GGA approximations for the exchange-correlation functional [47]. Extension of these methods to surface chemistry offers a potential solution to the

dilemma presented in Fig. 2. This section focuses on screened hybrids and meta-GGAs, two beyond-GGA approximations that show particular promise for surface chemistry.

5.1 Hybrid XC Functionals

DFT's widespread adoption for computational chemistry [47] is arguably directly attributed to the ability of fourth-rung "hybrid" XC approximations to outperform Hartree–Fock theory and second-order many-body perturbation theory for chemical bond breaking. Hybrid functionals [56] incorporate a fraction of exact exchange:

$$E_{\mathrm{X}}^{\mathrm{ex}}[\rho] = -\frac{1}{2}\int d^3\mathbf{r}\int d^3\mathbf{r}'\,\frac{\left|\gamma\left(\mathbf{r},\mathbf{r}'\right)\right|^2}{\left|\mathbf{r}-\mathbf{r}'\right|}. \tag{4}$$

Here, $\gamma(\mathbf{r},\mathbf{r}')$ is the nonlocal one-particle density matrix of the noninteracting Kohn–Sham wavefunction, constructed from the occupied orbitals of the Kohn–Sham wavefunction $\{\phi_i(\mathbf{r})\}$ as $\gamma(\mathbf{r},\mathbf{r}') = \Sigma_i\phi_i(\mathbf{r})\phi_i{}^*(\mathbf{r}')$. These orbitals and density matrices are thus implicit density functionals. Equation (4) provides the exact XC functional for one-electron systems, where it exactly cancels the Hartree interaction.

Admixture of a fraction of Eq. (4) to a GGA is justified by an adiabatic connection between the real system and the noninteracting Kohn–Sham reference [56, 184]. Such admixture corrects GGAs' over-delocalization, improving the prediction of a variety of properties including reaction barrier heights [146]. GGA calculations with more localized Hartree–Fock orbitals [153, 185], self-interaction corrections [186], and explicit constraints on localization [187] all give related improvements.

Hybrid functionals' success can be rationalized in terms of GGAs' simulation of nondynamical correlation [188]. Briefly, the many-body correction to the Hartree interaction may be modeled as a "hole" $h_{\mathrm{XC}}[\rho](\mathbf{r},\mathbf{r}')$ in the electron density about an electron at point $\mathbf{r}$ [145]. The total Hartree + XC energy becomes

$$E_{\mathrm{HXC}}[\rho] = \frac{1}{2}\int d^3\mathbf{r}\rho(\mathbf{r})\int d^3\mathbf{r}'\,\frac{\rho\left(\mathbf{r}'\right) + h_{\mathrm{XC}}[\rho]\left(\mathbf{r},\mathbf{r}'\right)}{\left|\mathbf{r}-\mathbf{r}'\right|}. \tag{5}$$

The XC hole is delocalized in systems such as stretched H_2^+, where $h_{\mathrm{XC}}[\rho](\mathbf{r},\mathbf{r}') = -\rho(\mathbf{r}')$. Nondynamical correlation in stretched covalent bonds localizes the XC hole about $\mathbf{r}$, such that (for example) an electron on the left atom in stretched singlet symmetric H_2 pushes the other electron to the right atom. GGA exchange functionals use localized exchange holes by construction. Thus, *GGA "exchange" in practice models both exchange and nondynamical correlation* [189].

Unfortunately, this rather crude model tends to overestimate nondynamical correlation and overbind. It is especially problematic in stretched H_2^+ and other

odd-electron bonds [190]. (Notions of self-interaction error [144, 191, 192], delocalization error [193], exact constraints on the XC hole [194], and symmetry breaking [195] provide additional insights into these effects.) Exact exchange admixture tunes this model, providing a surprisingly effective treatment of chemical bonding.

It is worth noting another useful aspect of hybrid functionals. Generalized Kohn–Sham calculations with a nonlocal XC potential $v_{\mathrm{XC}}(\mathbf{r},\mathbf{r}') = \delta E_{\mathrm{XC}}/\delta\gamma(\mathbf{r},\mathbf{r}')$ approximate the exact functional's derivative discontinuity, allowing occupied-virtual orbital energy gaps to better approximate fundamental band gaps better [196–198]. Kohn–Sham calculations with hybrid functionals require the local and multiplicative effective potential $\delta E_{\mathrm{X}}^{\mathrm{ex}}[\rho]/\delta\rho(\mathbf{r})$, typically constructed with optimized effective potential methods [199–201] or variants thereof [202]. Cohen et al. [197] provides a particularly useful illustration of the differences between Kohn–Sham and generalized Kohn–Sham calculations.

5.2 *Hybrid Functionals' Limitations*

Hybrid functionals have three limitations which are particularly relevant to surface chemistry. First, the optimal fraction of exact exchange is not known a priori, but depends on the role of nondynamical correlation in the system or property of interest. In practice, simple global hybrids typically require ~20% exact exchange for accurate thermochemistry [203, 204], ~40–50% exact exchange for kinetics [205], 100% exact exchange far from finite systems where Eq. (4) is the exact XC functional [206–208], and relatively small exact exchange admixtures for many organometallic properties [209, 210]. Figure 3 illustrates how this is problematic for reactions on metal surfaces. Other implications for surface chemistry are discussed in Hafner [211]. (The "+U" method, where the Kohn–Sham reference system includes a Hubbard repulsion on special sites [212], has similar issues centered on choosing the magnitude of U [213].) There has been substantial interest in overcoming this limitation through system-dependent [214–219] or position-dependent [220–222] exact exchange admixture, or through more sophisticated mean-field models of the XC hole [223, 224]. However, existing position-dependent "local hybrids" are not unambiguously more accurate than global hybrids [222, 225], and system-dependent exact exchange admixture can introduce size consistency issues [217].

Other limitations for surface chemistry arise from the long-range piece of exact exchange (large $|\mathbf{r}-\mathbf{r}'|$ in Eq. (4)) Evaluating this long-range contribution is computationally expensive in metallic systems where the Kohn–Sham $\gamma(\mathbf{r}, \mathbf{r}')$ delocalizes over a large range of $|\mathbf{r}-\mathbf{r}'|$ [226]. (More sophisticated treatments of this term have been proposed [227, 228].) Additionally, long-range exact exchange is exactly cancelled by higher order electron correlation effects in the HEG, and

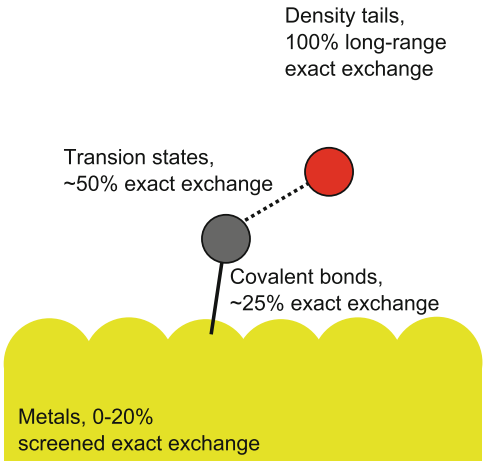


Fig. 3 Schematic of the optimal admixture of exact exchange in various regions of a reaction on a metal surface

approximately cancelled in metals [229, 230]. There are thus relatively few global hybrid DFT calculations on periodic metal slabs [161, 163, 231–233].

5.3 Screened Hybrids

Screened hybrid functionals [234, 235] rigorously [236] cut off the problematic long range of Eq. (4), sacrificing [237] an exact treatment of density tails [206–208] to enable facile treatments of periodic systems. The HSE06 screened hybrid [235, 238, 239] includes 25% of the error-function-screened exact exchange:

$$E_{\mathrm{X}}^{\mathrm{SR-ex}}[\rho] = -\frac{1}{2}\int d^3\mathbf{r}\int d^3\mathbf{r}'\,\frac{\left|\gamma\left(\mathbf{r},\mathbf{r}'\right)\right|^2}{\left|\mathbf{r}-\mathbf{r}'\right|}\mathrm{erfc}\left(\omega\left|\mathbf{r}-\mathbf{r}'\right|\right). \tag{6}$$

This is combined with 75% screened PBE exchange and 100% long-range PBE exchange and PBE correlation. Screened PBE exchange is constructed from explicit models [240, 241] of the angle- and system-averaged PBE exchange hole of Eq. (5):

$$E_{\mathrm{X}}^{\mathrm{SR-PBE}}[\rho] = \int d^3\mathbf{r}\, e_{\mathrm{X}}^{\mathrm{SR-PBE}}(\rho(\mathbf{r}),\nabla\rho(\mathbf{r})), \tag{7}$$

$$e_{\mathrm{X}}^{\mathrm{SR-PBE}}(\rho(\mathbf{r}),\nabla\rho(\mathbf{r})) = -\frac{1}{2}\rho(\mathbf{r})\int d^3\mathbf{r}'\,\frac{h_{\mathrm{X}}^{\mathrm{PBE}}\left(\rho(\mathbf{r}),\nabla\rho(\mathbf{r}),\left|\mathbf{r}-\mathbf{r}'\right|\right)}{\left|\mathbf{r}-\mathbf{r}'\right|}\mathrm{erfc}\left(\omega\left|\mathbf{r}-\mathbf{r}'\right|\right). \tag{8}$$

Exchange screening accelerates hybrid DFT calculations on periodic systems. Calculations with atom-centered basis functions can integrate $|\mathbf{r}-\mathbf{r}'|$ over a reduced number of replica cells [242]. Calculations with plane-wave basis functions can

downsample the k-space mesh for reciprocal space integration [226]. HSE06 is implemented in standard codes [243–245] and has been extensively applied to semiconductors [246, 247].

As for global hybrids, screened hybrids' optimum fraction of exact exchange and optimum screening parameter ω are not known a priori. The HSE06 exact exchange admixture comes from perturbation theory arguments [248], and the screening parameter is chosen empirically [235, 238]. Both parameters significantly affect the functional's computational cost in solids, and its performance for many properties [238, 249]. For example, while the HSE06 screening parameters appear nearly optimal for semiconductor bandgaps [250], they underestimate the bandgaps of large-gap insulators [218] and overestimate metals' energy bandwidths [226]. (Marques et al. [218] gives a particularly enlightening perspective on this effect, based on a relation between generalized Kohn–Sham equations with a screened hybrid functional and many-body GW calculations [251] with an effective static dielectric constant.) This has led to the exploration of several other screened hybrid forms [153, 252–257]. The "middle-range" screened hybrid HISS [152, 258] shows particular promise for surface chemistry. HISS uses a second screening function to include additional exact exchange at moderate $|\mathbf{r}-\mathbf{r}'|$. HISS accurately treats semiconductor bandgaps and lattice parameters [259], as well as some reactions on surfaces [165, 169]. Its more aggressive screening reduces its computational cost relative to HSE06 [259].

5.4 Meta-GGAs

A second route to fixing GGAs' limitations is third-rung functionals incorporating the noninteracting kinetic energy:

$$\tau(\mathbf{r}) = \frac{1}{2}\sum_i |\nabla \phi_i(\mathbf{r})|^2 = \frac{1}{2}\lim_{\mathbf{r}'\to\mathbf{r}} \nabla_{\mathbf{r}} \cdot \nabla_{\mathbf{r}'}\gamma\left(\mathbf{r}, \mathbf{r}'\right), \tag{9}$$

$$E_{\mathrm{XC}}^{\mathrm{mGGA}}[\rho] = \int d^3\mathbf{r}\, e_{\mathrm{XC}}^{\mathrm{mGGA}}(\rho(\mathbf{r}), |\nabla\rho(\mathbf{r})|, \tau(\mathbf{r})). \tag{10}$$

Meta-GGAs may also use the density Laplacian, which incorporates information similar to τ [260]. Meta-GGA calculations are not much more expensive than calculations with GGAs [209, 261]. This makes meta-GGAs particularly attractive for calculations on solids and surfaces. Early meta-GGAs [262, 263] showed promise for properties such as surface energies [264] and molecular thermochemistry [261, 265]. However, their adoption was limited by the fact that they are only comparable to GGAs for lattice parameters [263] and gas-phase kinetics [266]. Modifications improving lattice parameters [267] do not improve reaction barriers [180]. This performance was somewhat disappointing, given that Eq. (9)

should contain some information about the short-range nonlocal one-particle density matrix of Eq. (6), which clearly improves hybrid functionals' performance.

A breakthrough came with the demonstration that the M06-L meta-GGA containing 37 empirical parameters can treat many properties, from lattice parameters to molecular thermochemistry to reaction barriers, with accuracy approaching hybrid functionals [268, 269]. Reverse-engineering M06-L [270] showed that part of its success comes from the inclusion of $t^{-1} = \tau/\tau_{\mathrm{HEG}}$. Here $\tau_{\mathrm{HEG}} = (3/10)(6\pi^2)^{2/3}\rho^{5/3}$ equals τ in the HEG. t^{-1} can differentiate covalent vs non-covalent interactions [270]. The success of M06-L motivated subsequent development of minimally empirical meta-GGAs based on $\alpha = (\tau - \tau_{\mathrm{W}})/\tau_{\mathrm{HEG}}$, where $\tau_{\mathrm{W}} = |\nabla\rho|^2/(8\rho) \leq \tau$, equals τ in one-electron systems [271–273]. Viewed from this perspective, M06-L becomes a meta-GGA form fit to the exact functional (which would presumably give zero error in the fitting set), whose fitting coefficients revealed meta-GGAs' previously unrecognized possibilities. Other empirical meta-GGAs have also been explored [180, 274]. These new meta-GGAs are designed in part to overcome the oscillatory behavior of M06-L [275].

6 Beyond the GGA for Surface Chemistry

The methods introduced in Sect. 5 have begun to be applied to chemistry on periodic slab model surfaces. Results to date indicate that these new methods have a great deal of potential, and point to some remaining limitations which motivate further development.

6.1 Recent Applications

One important series of studies applies the M06-L meta-GGA to molecule-surface adsorption. Hammer and coworkers showed that the M06-L meta-GGA accurately treats "medium-range" noncovalent interactions for adsorbates on graphene [270, 276, 277], despite its lack of true long-range correlation [278]. The authors have applied M06-L in subsequent studies of hydrogenation [279] and CO intercalation [280] of supported graphene, and of RS-Au-SR "staple" motifs [281] in alkanethiol monolayers on Au(111) [282].

Another important milestone concerns treatments of CO on noble metal surfaces [132]. The HSE03 screened hybrid and the M06-L and revTPSS meta-GGAs all improve the binding site preference. The meta-GGAs correctly predict that CO preferentially adsorbs atop a single Pt atom on Pt(111), along with encouraging accuracy for lattice constants, surface formation energies, and adsorption energies [283, 284]. HSE03 provides the correct site preference for CO on Cu(111) and Rh(111) surfaces. While it still fails for Pt(111), it reduces the top-fcc energy difference relative to PBE [231]. HSE03 shows similar trends for CO adsorption on the terraces of stepped Rh(553) [59].

A third application is to adsorbates and defects on metal oxide surfaces. These studies build upon screened hybrids' successes for modeling electrons localized in bulk defects [246, 285, 286]. HSE06 and the B3LYP global hybrid were used to analyze Au adatoms on ceria defects [172], a system where GGAs' over-delocalization leads to qualitatively incorrect results [213, 287]. HSE was used to check PBE + U calculations on ceria-supported vanadia catalysts [288], to confirm conclusions about the reactivity of oxygen vacancies on titania surfaces [289], and to check the spin polarization of graphite surface defects [290]. Pacchioni [291] reviews some other relevant studies.

There have been relatively few applications of meta-GGAs and screened hybrids to reaction barriers of molecules on surfaces. However, results to date suggest that these functionals' improvements for gas-phase barriers carry over to surfaces. Unlike the PBE and RPBE GGAs [162], the HSE06 screened hybrid and PBE0 global hybrid predict a substantial barrier to O_2 dissociation on Al(111) slabs [163]. (As discussed above, this failure of GGAs was previously attributed entirely to spin selection rules for triplet O_2 dissociation [162].) M06-L predicts reasonable binding energies for molecules on zeolite catalysts [83], and has been applied to heterogeneous catalysis by zeolites and metal-oxide frameworks [292, 293]. We showed that the HSE06 and HISS screened hybrids, and the M06-L meta-GGA, improved upon GGAs' underestimate of the dissociation barrier for ammonia dissociation on a cluster model of Si(100). Similar trends were seen for calculations on Si slabs [165]. We also showed that HSE06, HISS, and M06-L improve GGAs' underestimated diffusion barriers for adatoms on graphene [169] and H_2 dissociation on gold and silver clusters [160]. Interestingly, the TPSS meta-GGA increases the too-low PBE barriers for H_2 dissociation on reconstructed Si(001) surfaces, despite the two functionals' similarity for gas-phase barriers [266].

6.2 Limitations

Some studies have identified limitations of existing beyond-GGA functionals for surface chemistry. Lousada et al. [164] shows that, similar to GGAs, M06-L predicts a qualitatively incorrect barrierless dissociation of H_2O_2 on metal oxides. Valero et al. [294, 295] shows that M06-L is problematic for the frequency shifts of CO and NO on nickel and magnesium oxides. These errors are mitigated by global hybrids [294, 295]. The M06-L and TPSS meta-GGAs do not improve upon the dispersion-corrected B97-D GGA for the aforementioned problem of methanol oxidation over Au_8^-, giving mean unsigned errors in reaction barrier heights of 10.1, 9.2, and 7.4 kcal mol^{-1}, respectively [31]. However, the B3LYP and M06 [269] global hybrids improve upon B97-D, with errors of 3.6 and 3.9 kcal mol^{-1}. HSE06's overestimated metal bandwidths [226] appear to contribute to its aforementioned problems for CO on Pt(111) [231]. The B3LYP global hybrid has other

problems for metals, which arise because its GGA for correlation [69] does not recover the correct HEG behavior [296].

6.3 Systematic Trends

Our recent extension [169] of diffusion Monte Carlo studies on adatom adsorption and diffusion over graphene [34] provides a systematic illustration of how screened exchange affects the Procrustean dilemma [62] faced by GGAs for surface chemistry. Part of that study considered systematic modification of the PBE GGA by both rescaling of the exchange enhancement factor ($F_X = \beta F_X^{PBE}$, see (3)), and admixture of a fraction α of screened exact exchange. (Put another way, changing β is similar to specific reaction parameter GGAs used for surface chemistry [181, 182]. Changing α is similar to the original specific reaction parameter global hybrids [183]). Figure 4 illustrates how changing β and α affect representative surface and molecule properties. The left panel shows calculated adsorption energies and diffusion barriers for O atom on graphene, compared to the diffusion Monte Carlo results of Hsing et al. [34]. The right panel shows errors in standard sets of gas-phase molecular thermochemistry and kinetics [297]. Computational details are in Barone et al. [169].

Figure 4 shows that GGA rescaling β simultaneously changes both adsorption energies and reaction barriers, and that *no* value of β can treat both properties. Figure 3 of Barone et al. [169] shows that a simple dispersion correction increased the GGA chemisorption energies, but did not affect reaction barriers. This is consistent with the results of Fig. 2 and with the limitations of empirical GGAs discussed in Sect. 4. In contrast, screened exchange admixture α increases both

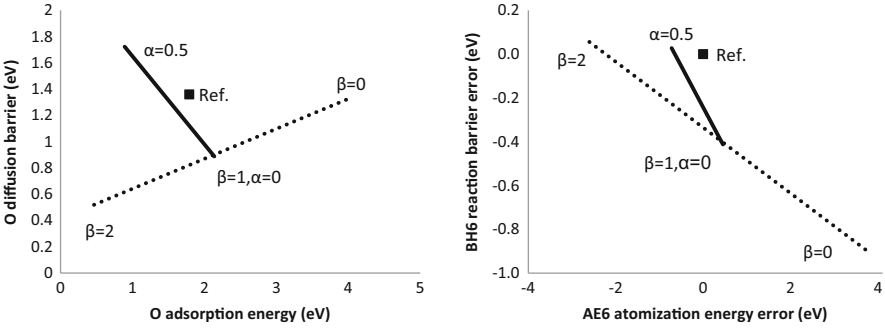


Fig. 4 Systematic variations in GGA enhancement factor $F_X = \beta F_X^{PBE}$ (*dotted lines*) and screened exchange admixture α (*solid lines*). $\beta = 1$, $\alpha = 0$ is the PBE GGA, $\beta = 1$, $\alpha = 0.25$ is the HSE06 screened hybrid. *Left*: Diffusion barrier vs adsorption energy for O on graphene. *Right*: Mean signed errors in gas-phase BH6 [297] kinetics vs AE6 thermochemistry. "Ref." are diffusion Monte Carlo from Ren et al. [34] (*left*) and zero mean signed error (*right*). Adapted with permission from Barone et al. [169]. Copyright 2013 American Chemical Society

surface diffusion barriers and molecule reaction barriers, while maintaining reasonable thermochemistry and atomization energies. The results suggest that screened hybrids have the potential to improve reaction barriers on surfaces, just as global hybrids improve gas-phase reaction barriers.

7 New Frontiers

Meta-GGAs and screened hybrids are not yet standard methods for simulating heterogeneous catalysis or surface chemistry. Additional work is needed to understand better their strengths and limitations, and to develop more accurate extensions. This section briefly reviews selected recent work along those lines, focusing on results from the author and his collaborators.

7.1 Dispersion-Corrected and Empirical Screened Hybrids

Barone et al. [169] suggests that GGAs' limitations for dispersion interactions and reaction barriers are largely orthogonal. Dispersion-corrected screened hybrids could in principle combine the aforementioned successes of dispersion corrections for adsorption energies and hybrid exchange for reaction barriers. A dispersion-corrected screened hybrid was recently benchmarked for rare-gas solids [255], and applied to C–H bond cleavage in crystalline polyethylene [298] and Au adatom adsorption on defective CeO_2(111) [172] mimicking ceria-supported gold catalysts [299]. Dispersion-corrected PBE and HSE calculations gave similar barriers to tetrachloropyrazine chemisorption on Pt(111) [300]. Dispersion-corrected global hybrids have also been applied to some molecular crystals [301], crystalline polymers [302], and surface chemistry [303].

It is interesting to consider whether empirical functional forms [173, 174, 179, 180] could benefit from screened exchange admixture. Perverati and Truhlar [257] proposed screened hybrids built upon the parameterizations of [179, 180]. These functionals improve upon HSE06 for gas-phase reaction barriers and some lattice constants [257]. They show modest promise for binding and relative energies of water clusters, properties which appear to be improved by dispersion corrections [304]. However, they have not yet been extensively tested for surface chemistry.

7.2 Rung 3.5 Functionals

We have proposed a new class of approximate functionals constructed to be intermediate between third-rung meta-GGAs and fourth-rung screened hybrids.

Rung 3.5 functionals replace one of the one-particle density matrices in (4) with a GGA model density matrix γ_{GGA}:

$$E_X^{\Pi}[\rho] = -\frac{1}{2}\int d^3\mathbf{r}\int d^3\mathbf{r}'\,\frac{\gamma(\mathbf{r},\mathbf{r}')\gamma_{GGA}\left(\rho(\mathbf{r}),\nabla\rho(\mathbf{r}),\mathbf{r}-\mathbf{r}'\right)}{|\mathbf{r}-\mathbf{r}'|}. \tag{11}$$

The integrand of (11) is symmetrized in $\mathbf{r}, \mathbf{r}'$ before use. The GGA model density matrix is implicit in the construction of the GGA exchange hole of Eq. (8):

$$h_X^{GGA}\left(\rho(\mathbf{r}),\nabla\rho(\mathbf{r}),|\mathbf{r}-\mathbf{r}'|\right) = -\frac{1}{2}\rho^{-1}(\mathbf{r})\left\langle\left|\gamma_{GGA}\left(\rho(\mathbf{r}),\nabla\rho(\mathbf{r}),\mathbf{r}-\mathbf{r}'\right)\right|^2\right\rangle_{\Omega}, \tag{12}$$

where $\langle\ldots\rangle_{\Omega}$ denotes angle averaging. Most existing exchange hole models treat the angle- and system-averaged hole [240, 241], while γ_{GGA} is explicitly not angle averaged [305]. γ_{GGA}, similar to h_X^{GGA}, decays rapidly in $|\mathbf{r}-\mathbf{r}'|$ by construction, aiding evaluation of Eq. (11) in metals. (Recall from Sect. 5.1 that this localization is central to the GGA "exchange" functionals' model of nondynamical correlation.) γ_{GGA} also tunes the amount of nonlocal information incorporated at each point, potentially providing a route to simultaneously treating several of the regions in Fig. 3. (Note that Eq. (11) cannot include 100% long-range exact exchange, and is not exact for one-electron regions.) Rung 3.5 functionals thus have the potential to address all three limitations of exact exchange admixture discussed in Sect. 5.1. Janesko [306] reviews our applications of Rung 3.5 functionals. Benchmarks for molecular thermochemistry and kinetics show that they can provide accuracy intermediate between standard GGAs and screened hybrids.

Table 1 presents previously unpublished results applying the Rung 3.5 functional *Π* 1PBE [305] to ammonia dissociation on Si(100). *Π* 1PBE admixes 25% of Eq. (11) to the PBE GGA, using a model density matrix γ_{PBE} constructed to reproduce the PBE exchange enhancement factor. Calculations use a small cluster model (nine Si atoms) from Sniatynsky et al. [165]. The Rung 3.5 reaction barrier is between the third-rung TPSS meta-GGA and the fourth-rung HSE06 screened

Table 1 Adsorption and dissociation of NH_3 on Si (100)

Method	ΔE_{ads}	$\Delta E^{\ddagger}$	ΔE_{rxn}
PBE	−1.01	0.67	−0.99
TPSS	−0.97	0.71	−1.15
Π 1PBE	−1.11	0.74	−1.03
HSE06	−1.13	0.80	−1.11
ref	−1.08	0.86	−1.19

Calculated Si–NH_3 adsorption energy ΔE_{ads}, and dissociation barrier $\Delta E^{\ddagger}$ and reaction energy ΔE_{rxn} (eV) for adsorbed NH_2–H bond dissociation. Results for NH_3 adsorbed to a cluster model for the Si(001) surface. DFT calculations use the 6-311++G (2d,2p) basis set, other computational details and "ref" complete-basis-set-extrapolated CCSD(T) benchmarks are in Sniatynsky et al. [165]

hybrid, indicating that the functional lives up to its name. (Results in Table 1 differ by ~0.02 eV from Sniatynsky et al. [165] because of a different basis set.) We are currently exploring more extensive applications of Rung 3.5 functionals to surface chemistry.

8 Conclusions

The successes of GGAs for surface chemistry are particularly remarkable, given their underlying limitations. Density functional approximations beyond the GGA, largely developed in the computational chemistry community, show promise for ameliorating these limitations in simulations of surface chemistry. Recent calculations illustrate these new methods' potential and point to remaining issues. It is hoped that these promising preliminary results motivate density functional developers to consider further the applications to surfaces, and motivate surface scientists to test beyond-GGA approximations on new systems. More accurate, computationally tractable methods including beyond-GGA DFT will help build upon GGAs' successes for surface chemistry.

Acknowledgements This work was supported by the Qatar National Research Foundation through the National Priorities Research Program (NPRP Grant No. 09-143-1-022), and by the Department of Chemistry at Texas Christian University.

References

1. Ertl G (2009) Reactions at solid surfaces. Wiley, Hoboken
2. Kneipp J, Kneipp H, Kneipp K (2008) Chem Rev 37:1052
3. Manthiram A, Murugan AV, Sarkar A, Muraliganth T (2008) Energy Environ Sci 1:621
4. Yazev O (2013) Acc Chem Res 46:2319
5. The Nobel Prize in Chemistry 2007. Nobelprize.org. Nobel Media AB 2013. Web 8 Dec 2013
6. Dry ME (2002) Catal Today 71:227
7. Schulz H (1999) Appl Catal Gen 186:3
8. Galvis HMT, Koeken AC, Bitter JH, Davidian T, Ruitenbeek M, Dugulan AI, de Jong KP (2013) J Catal 303:22
9. Mitchell WJ, Wang Y, Xie J, and Weinberg WH (1993) J Am Chem Soc 115:4381
10. Heiz U, Schneider WD (2000) J Phys D Appl Phys 33:R85
11. Zhang S, Nguyen L, Zhu Y, Zhan S, Tsung CKF, Tao FF (2013) Acc Chem Res 46:1731
12. Vang RT, Lauritsen JV, Lægsgaard E, Bensebacher F (2008) Chem Soc Rev 37:2191
13. Maitlis PM (2004) J Organomet Chem 689:4366
14. Neurock M, van Santen RA (2006) Molecular heterogeneous catalysis: a conceptual and computational approach. Wiley-VCH Verlag GmbH & Co. KGaA, Weinham
15. Schatz GC, Van Duyne RP (2002) In: Chalmers JM, Griffiths PR (eds) Handbook of vibrational spectroscopy. Wiley, New York
16. Zhao LL, Jensen L, Schatz GC (2006) J Am Chem Soc 128:2911
17. Son YW, Cohen ML, Louie SG (2006) Nature 444:347

18. Ovesen CV, Clausen BS, Hammershøi BS, Steffensen G, Askgaard T, Chorkendorff I, Nørskov JK, Rasmussen PB, Stoltze P, Taylor P (1996) J Catal 158:170
19. Callaghan C, Fishtik I, Datta R (2002) Worchester Polytechnic Institute, Fuel Chemistry Division Preprints
20. Grabow LC, Gokhale AA, Evans ST, Dumesic JA, Mavrikakis M (2008) J Phys Chem C 112:4608
21. Mirwald JR, Inderwildi OR (2012) Phys Chem Chem Phys 14:7028
22. Lanzani G, Martinazzo R, Materzanini G, Pino I, Tantardini GF (2007) Theor Chem Acc 117:805
23. Czekaj I, Wambach J, Kröcher O (2009) Int J Mol Sci 10:4310
24. Huang P, Carter EA (2008) Annu Rev Phys Chem 59:261
25. Bauschlicher CW Jr (1994) J Chem Phys 101:3250
26. Filippi C, Healy SB, Kratzer P, Pehlke E, Scheffler M (2002) Phys Rev Lett 89:166102
27. Szabo A, Ostlund NS (1989) Modern quantum chemistry: introduction to advanced electronic structure theory. Dover Publications Inc., Mineola, New York
28. de Jong GT, Bickelhaupt FM (2006) J Chem Theory Comput 2:322
29. Njegic B, Gordon MS (2008) J Chem Phys 129:124705
30. Wang Y, Qian HJ, Morokuma K, Irle S (2012) J Phys Chem A 116:7154
31. Hansen JA, Ehara M, Piecuch P (2013) J Phys Chem A 117:10416
32. Pozzo M, Alfè D (2008) Phys Rev B 78:245313
33. Kanai Y, Takeuchi N (2009) J Chem Phys 131:214708
34. Hsing CR, Wei CM, Chou MY (2012) J Phys Condens Matter 24:395002
35. Hoggan PE (2013) Int J Quant Chem 113:277
36. Ren X, Rinke P, Scheffler M (2009) Phys Rev B 80:045402
37. Göltl F, Grüneis A, Bučko T, Hafner J (2012) J Chem Phys 137:114111
38. Mittendorfer F, Garhofer A, Redinger J, Klimeš J, Harl J, Kresse G (2011) Phys Rev B 84:201401
39. Ren X, Rinke P, Joas C, Scheffler M (2012) J Mater Sci 47:7447
40. Kim HJ, Tkatchenko A, Cho JH, Scheffler M (2012) Phys Rev B 85:041403(R)
41. Karlický F, Lazar P, Dubecký M, Otyepka M (2013) J Chem Theory Comput 9:3670
42. Olsen T, Thygesen KS (2013) Phys Rev B 87:075111
43. Hayashi K, Sato S, Bai S, Higuchi Y, Ozawa N, Shimazaki T, Adachi J, Martin JM, Kubo M (2012) Faraday Discuss 156:137
44. Hohenberg P, Kohn W (1964) Phys Rev B136:864
45. Kohn W, Sham L (1965) Phys Rev 140:A1133
46. ISI Web of Science. Accessed 17 Dec 2013
47. Burke K (2012) J Chem Phys 136:150901
48. Bartlett RJ, Schweigert IV, Lotrich VF (2006) J Mol Struct 771:1
49. Perdew JP, Schmidt K (2001) In: Van Doren V, Van Alsenoy C, Geerlings P (eds) Density functional theory and its application to materials. American Institute of Physics, Melville, pp 1–20
50. Perdew JP (1986) Phys Rev A 33:8822
51. Vosko SH, Wilk L, Nusair M (1980) Can J Phys 58(8):1200
52. Gell-Mann M, Brueckner KA (1957) Phys Rev 106:364
53. Ceperley DM, Alder BJ (1980) Phys Rev Lett 45:566
54. Burke K, Cruz FG, Lam KC (1998) J Chem Phys 109(19):8161
55. Perdew JP, Burke K, Ernzerhof M (1996) Phys Rev Lett 77:3865, 78:1396(E) (1997)
56. Becke AD (1993) J Chem Phys 98(2):1372
57. Slater JC (1951) Phys Rev 81:385
58. Janthon P, Kozlov SM, Viñes F, Limtrakul J, Illas F (2013) J Chem Theory Comput 9:1631
59. Stroppa A, Kresse G (2008) New J Phys 10:063020
60. Schimka L, Harl J, Stroppa A, Grüneis A, Marsman M, Mittendorfer F, Kresse G (2010) Nat Mater 9:741

61. Goerigk L, Grimme S (2011) Phys Chem Chem Phys 13:6670
62. Perdew JP, Ruzsinszky A, Csonka GA, Vydrov OA, Scuseria GE, Constantin LI, Zhou X, Burke K (2008) Phys Rev Lett 100:136406
63. Armiento R, Mattsson AE (2005) Phys Rev B 72:085108
64. Wu Z, Cohen RE (2006) Phys Rev B 73:235116
65. Tran F, Lasowski R, Blaha P, Schwarz K (2007) Phys Rev B 75:115131
66. Zhang Y, Yang W (1998) Phys Rev Lett 80:890
67. Hammer B, Hansen LB, Nørskov JK (1999) Phys Rev B 59:7413
68. Becke AD (1988) Phys Rev A 38:3098
69. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785
70. Becke AD (1992) J Chem Phys 96:2155
71. Hamprecht FA, Cohen AJ, Tozer DJ, Handy NC (1998) J Chem Phys 109:6264
72. Perdew JP (1991) In: Ziesche P, Eschrig H (eds) Electronic structure of solids'91. Akademie Verlag, Berlin
73. Patton DC, Pederson MR (1997) Phys Rev A 56:R2495
74. Grimme S (2011) WIREs Comput Mol Sci 1:211
75. Johnson ER, Mackie ID, DiLabio GA (2009) J Phys Organ Chem 22:1127
76. Klimes J, Michaelides A (2012) J Chem Phys 137:120901
77. Ramalho JAPP, Gomes JRB, Illas F (2013) RSC Adv 3:13085
78. Thrower JD, Friis EE, Skov AL, Nilsson L, Andersen M, Ferrighi L, Jrgensen B, Baouche S, Balog R, Hammer B, Hornekr L (2013) J Phys Chem C 117(26):13520
79. Romaner L, Nabok D, Puschnig P, Zojer E, Ambrosch-Draxl C (2009) New J Phys 11:053010
80. Ruiz VG, Liu W, Zojer E, Scheffler M, Tkatchenko A (2012) Phys Rev Lett 108:146103
81. Londero E, Karlson EK, Landahl M, Ostrovskii D, Rydberg JD, Schröder E (2012) J Phys Condens Matter 24(42):424212
82. Rakow JR, Tüllmann S, Holthausen MC (2009) J Phys Chem A 113:12035
83. Zhao Y, Truhlar DG (2008) J Phys Chem C 112:6860
84. Gomes J, Zimmerman PM, Head-Gordon M, Bell AT (2012) J Phys Chem C 116:15406
85. Goerigk L, Grimme S (2010) J Chem Theory Comput 6:107
86. Zhao Y, Lynch BJ, Truhlar DG (2004) J Phys Chem A 108:2715
87. Zhao Y, Gonzáles-García N, Truhlar DG (2005) J Phys Chem A 109:2012, 110:4942 (E) (2006)
88. Gerber IC, Ángyán JG (2007) J Chem Phys 126:044103
89. Schwabe T, Grimme S (2007) Phys Chem Chem Phys 9:3397
90. Benighaus T, DiStasio RA Jr, Lochan RC, Chai JD, Head-Gordon M (2008) J Phys Chem A 112:2702
91. Chai JD, Head-Gordon M (2008) Phys Chem Chem Phys 10:6615
92. Lin YS, Li GD, Mao SP, Chai JD (2012) J Chem Theory Comput. dx.doi.org/10.1021/ct300715s
93. Svelle S, Tuma C, Rozanska X, Kerber T, Sauer J (2009) J Am Chem Soc 131(2):816
94. Hafner J (2008) J Comput Chem 29:2044
95. Greeley J, Nørskov JK, Mavrikakis M (2002) Annu Rev Phys Chem 53:319
96. Nørskov JK, Abild-Pedersen F, Studt F, Bligaard T (2011) Proc Natl Acad Sci USA 108:937
97. Hod O, Peralta JE, Scuseria GE (2007) Phys Rev B 76:233401
98. Chen Z, Lu YM, Rooks MJ, Avouris P (2007) Physica E 40:228
99. Han MY, Özyilmaz B, Zhang Y, Kim P (2007) Phys Rev Lett 98:206805
100. Kozlov SM, Viñes F, Görling A (2011) Adv Mater 23:2638
101. Balog R, Jrgensen B, Nilsson L, Andersen M, Rienks E, Bianchi M, Fanetti M, Lgs-gaard E, Baraldi A, Lizzit S, Sljivancanin Z, Besenbacher F, Hammer B, Pedersen TG, Hofmann P, Hornekr L (2010) Nat Mater 9:315
102. Lopez-Bezanilla A, Huang J, Terrones H, Sumpter BG (2011) Nano Lett 11:3267
103. Miotto R, Srivastava GP, Ferraz AC (1998) Phys Rev B 58:7944
104. Lee SH, Kang MH (1998) Phys Rev B 58:4903

105. Chung ON, Kim H, Chung S, Koo JY (2006) Phys Rev B 73:033303
106. Bowler DR, Owen JHG (2007) Phys Rev B 75:155310
107. Owen JHG (2009) J Phys Condens Matter 21:44301
108. Honkala K, Hellman A, Remediakis IN, Logadottir A, Carlsson A, Dahl S, Christiansen CH, Nørskov JK (2005) Science 307:555
109. Dahl S, Logadottir A, Egeberg RC, Larsen JH, Chorkendorff I, Törnqvist E, Nørskov JK (1999) Phys Rev Lett 83:1814
110. Ciobica IM, Van Santen RA (2003) J Phys Chem B 107:3808
111. Inderwildi OR, Jenkins SJ, King DA (2008) J Phys Chem C 112:1305
112. Andersson M, Abild-Pedersen F, Remediakis I, Bligaard T, Jones G, Engæk J, Lytken O, Horch S, Nielsen J, Sehested J, Rostrup-Nielsen J, Nørskov J, Chorkendorff I (2008) J Catal 255:6
113. Inderwildi OR, Jenkins SJ, King DA (2008) Angew Chem Int Ed 47:5253
114. Ojeda M, Nabar R, Nilekar AU, Ishikawa A, Mavrikakis M, Iglesia E (2010) J Catal 272:287
115. Shetty S, Jansen PJ, van Santen RA (2009) J Am Chem Soc 131:12874
116. Liu ZP, Hu P (2002) J Am Chem Soc 124:11568
117. Huo CF, Wu BS, Gao P, Li YYYW, Jiao H (2011) Angew Chem Int Ed 50:7403
118. Bengaard H, Nørskov J, Sehested J, Clausen B, Nielsen L, Molenbroek A, Rostrup-Nielsen J (2002) J Catal 209(2):365
119. Curulla-Ferre D, Govender A, Bromfield TC, Niemantsverdriet JWH (2006) J Phys Chem B 110:13897
120. Ge Q, Neurock M (2006) J Phys Chem B 110:15368
121. Cheng J, Hu P, Ellis P, French S, Kelly G, Lok CM (2010) Topics Catal 53:326
122. van Santen RA, Ghouri MM, Shetty S, Hensen EMH (2011) Catal Sci Technol 1:891
123. Alfonso DR (2013) J Phys Chem C 117:20562
124. Newsome DS (1980) Catal Rev Sci Eng 21:275
125. Gokhale AA, Dumesic JA, Mavrikakis M (2008) J Am Chem Soc 130:1402
126. Grabow LC, Mavrikakis M (2011) ACS Catal 1:365
127. Sanchez A, Abbet S, Heitz U, Scheider WD, Häkkinen H, Barnett RN, Landman U (1999) J Phys Chem A 103:9573
128. Yoon B, Häkkinen H, Landman U, Wörz AS, Antonetti JM, Abbet S, Judai K, Heitz U (2005) Science 307:403
129. Janssens TVW, Clausen BS, Hvolbk B, Falsig H, Christensen CH, Bligaard T, Nrskov JK (2007) Topics Catal 44:15
130. Boronat M, Concepción P, Corma A (2009) J Phys Chem C 113:16772
131. Zhang C, Michaelides A, Jenkins ST (2011) Phys Chem Chem Phys 13:22
132. Feibelman PJ, Hammer B, Nørskov JK, Wagner F, Scheffler M, Stumpf R, Watwe R, Dumesic J (2001) J Phys Chem B 105:4018
133. Hammer B, Morikawa Y, Nørskov JK (1996) Phys Rev Lett 76:2141
134. Abild-Pedersen F, Greeley J, Studt F, Rossmeisl J, Munter TR, Moses PG, Skulason E, Bligaard T, Norskov JK (2007) Phys Rev Lett 99:016105
135. Luengnaruemitchai A, Osuwan S, Gulari E (2003) Catal Commun 4:215
136. Schumachera N, Boisena A, Dahl S, Gokhalec AA, Kandoic S, Grabowc L, Dumesicc J, Mavrikakis M, Chorkendorff I (2005) J Catal 229:265
137. Bligaard T, Nørskov JK, Dahl S, Matthiesen J, Christensen CH, Sehested J (2004) J Catal 224:206
138. Cheng J, Hu P (2008) J Am Chem Soc 130:10868
139. Andersson MP, Bligaard T, Kustov A, Larsen KE, Greeley J, Johannessen T, Christensen CH, Nørskov JK (2006) J Catal 239:501
140. Sehested J, Larsen KE, Kustov AL, Frey AM, Johannessen T, Bligaard T, Andersson MP, Nørskov JK, Christensen CH (2007) Topics Catal 45:9
141. Studt F, Abild-Pedersen F, Bligaard T, Sørensen RJ, Christiansen CH, Nørskov JK (2009) Science 320:1320

142. Greeley J, Mavrikakis M (2004) Nat Mater 3:810
143. Nørskov JK, Bligaard T, Rossmeisl J, Christiansen CH (2009) Nat Chem 1:37
144. Mori-Sánchez P, Cohen AJ, Yang W (2006) J Chem Phys 125:201102
145. Perdew JP, Ruzsinszky A, Constantin LA, Sun J, Csonka GI (2009) J Chem Theory Comput 5:902
146. Csonka GI, Perdew JP, Ruzsinszky A (2010) J Chem Theory Comput 6(12):3688
147. Deng L, Branchadell V, Ziegler T (1994) J Am Chem Soc 116:10645
148. Andersson S, Grüning M (2004) J Phys Chem A 108:7621
149. Riley KE, Op't Holt BT, Merz KE Jr (2007) J Chem Theory Comput 3:407
150. Zheng J, Zhao Y, Truhlar DG (2009) J Chem Theory Comput 5:808
151. Yang K, Zheng J, Zhao Y, Truhlar DG (2010) J Chem Phys 132:164117
152. Henderson TM, Izmaylov AF, Scuseria GE, Savin A (2008) J Chem Theory Comput 4:1254
153. Janesko BG, Scuseria GE (2008) J Chem Phys 129:124110
154. Mangiatordi GF, Brémond E, Adamo C (2012) J Chem Theory Comput 8:3082
155. Nachtigall P, Jordan KD, Smith A, Jónsson H (1996) J Chem Phys 104:148
156. Dürr M, Raschke MB, Pehlke E, Höfer U (2001) Phys Rev Lett 86:123
157. de Jong GT, Bickelhaupt FM (2005) J Phys Chem A 109:9685
158. de Jong GT, Geerke DP, Diefenbach A, Bickelhaupt FM (2005) Chem Phys 313:261
159. de Jong GT, Geerke DP, Diffenbach A, Solá M, Bickelhaupt FM (2005) J Comput Chem 26:1007
160. Moncho S, Brothers EN, Janesko BG (2013) J Phys Chem C 117:7487
161. Livshits E, Baer R, Kosloff R (2009) J Phys Chem A 113:7521
162. Behler J, Delley B, Lorenz S, Reuter K, Scheffler M (2005) Phys Rev Lett 94:036104
163. Liu HR, Xiang H, Gong XG (2011) J Chem Phys 135:214702
164. Lousada CM, Johansson AJ, Brinck T, Jonsson M (2013) Phys Chem Chem Phys 15:5539
165. Sniatynsky R, Janesko BG, El-Mellouhi F, Brothers EN (2012) J Phys Chem C 116:26396
166. Hossain MZ, Yamashita Y, Mukai K, Yoshinobu J (2003) Phys Rev B 68:235322
167. Smedarchina ZK, Zgierski MZ (2003) Int J Mol Sci 4:445
168. Uthaisar C, Barone V (2010) Nano Lett 10:2838
169. Barone V, Brothers EN, Janesko BG (2013) J Chem Theory Comput 9:4853
170. Lopez N, Illas F, Rösch N, Pacchioni G (1999) J Chem Phys 110:4873
171. Di Valentin C, Pacchioni G, Selloni A (2006) Phys Rev Lett 97:166803
172. Pan Y, Nilius N, Freund HJ, Paier J, Penschke C, Sauer J (2013) Phys Rev Lett 111:206101
173. Becke AD (1997) J Chem Phys 107:8554
174. Wellendorff J, Lundgaard KT, Møgelhøj A, Petzold V, Landis DD, Nørskov JK, Bligaard T, Jacobsen KW (2012) Phys Rev B 85:235149
175. Rajasekaran S, Abild-Pedersen F, Ogasawara H, Nilsson A, Kaya S (2013) Phys Rev Lett 111:085503
176. Brogaard RY, Weckhuysen BM, Nørskov JK (2013) J Catal 300:235
177. Studt F, Abild-Pedersen F, Varley JB, Nørskov JK (2013) Catal Lett 143:71
178. Peverati R, Zhao Y, Truhlar DG (2011) J Phys Chem Lett 2:1991
179. Perverati R, Truhlar DG (2012) J Chem Theory Comput 8:2310
180. Perverati R, Truhlar DG (2012) Phys Chem Chem Phys 14:13171
181. Díaz C, Pijper E, Olsen RA, Busnengo HF, Auerbach DJ, Kroes GJ (2009) Science 326:832
182. Kroes GJ (2012) Phys Chem Chem Phys 14:14966
183. Chuang YY, Radhakrishnan ML, Fast PL, Cramer CJ, Truhlar DG (1999) J Phys Chem A 103:4893
184. Yang W (1998) J Chem Phys 109:10107
185. Kim MC, Sim E, Burke K (2013) Phys Rev Lett 111:073003
186. Vydrov OA, Scuseria GE, Perdew JP, Ruzsinszky A, Csonka GI (2006) J Chem Phys 124:094108
187. Wu Q, Kaduk B, van Voorhis T (2008) J Chem Phys 130:034109
188. Cremer D (2001) Mol Phys 99(23):1899

189. Perdew JP, Staroverov VN, Tao J, Scuseria GE (2008) Phys Rev A 78:052513
190. Yang W, Zhang Y, Ayers PW (2000) Phys Rev Lett 84:5172
191. Zhang Y, Yang W (1998) J Chem Phys 109(7):2604
192. Ruzsinszky A, Perdew JP, Csonka GI, Vydrov OA, Scuseria GE (2006) J Chem Phys 125:194112
193. Mori-Sánchez P, Cohen AJ, Yang W (2008) Phys Rev Lett 100:146401
194. Burke K, Perdew JP, Ernzerhof M (1998) J Chem Phys 109:3760
195. Garza AJ, Jimenez-Hoyos CA, Scuseria GE (2013) J Chem Phys 138:134102
196. Seidl A, Görling A, Vogl P, Majewski JA, Levy M (1996) Phys Rev B 53:3764
197. Cohen AJ, Mori-Sánchez P, Yang W (2008) Phys Rev B 77:115123
198. Yang W, Cohen AJ, Mori-Sánchez P (2012) J Chem Phys 136:204111
199. Sharp RT, Horton GK (1953) Phys Rev 90:317
200. Krieger JB, Li Y, Iafrate GJ (1992) Phys Rev A 45:101
201. Ivanov S, Hirata S, Bartlett RJ (1999) Phys Rev Lett 83:5455
202. Staroverov VN, Scuseria GE, Davidson ER (2006) J Chem Phys 124:114103
203. Becke AD (1993) J Chem Phys 98(7):5648
204. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994) J Phys Chem 98(45):11623
205. Lynch BJ, Fast PL, Harris M, Truhlar DG (2000) J Phys Chem A 104(21):4811
206. Stoll H, Savin A (1985) In: Dreizler R, da Providencia J (eds) Density functional methods in physics. Plenum, New York, p 177
207. Savin A (1996) In: Seminario JM (ed) Recent developments and applications of modern density functional theory. Elsevier, Amsterdam, p 327
208. Leininger T, Stoll H, Werner HJ, Savin A (1997) Chem Phys Lett 275:151
209. Furche F, Perdew JP (2006) J Chem Phys 124:044103
210. Schultz NE, Zhao Y, Truhlar DG (2005) J Phys Chem A 109:11127
211. Hafner J (2008) Montash Chem 139:373
212. Anisimov VI, Zaanen J, Andersen OK (1991) Phys Rev B 44:943
213. Loschen C, Carrasco J, Neyman KM, Illas F (2007) Phys Rev B 75(3):035115
214. Baer R, Neuhauser D (2005) Phys Rev Lett 94:043002
215. Livshits E, Baer R (2007) Phys Chem Chem Phys 9:2932
216. Stein T, Kronik L, Baer R (2009) J Am Chem Soc 131:2818
217. Odashima MM, Capelle K (2009) Phys Rev A 79:062515
218. Marques MAL, Vidal J, Oliveria MJT, Reining L, Botti S (2011) Phys Rev B 83:035119
219. Stein T, Eisenberg H, Kronik L, Baer R (2010) Phys Rev Lett 105:266802
220. Jaramillo J, Scuseria GE, Ernzerhof M (2003) J Chem Phys 118(3):1068
221. Arbuznikov AV, Kaupp M (2011) Int J Quant Chem 111:2625
222. Haunschild R, Odashima MM, Scuseria GE, Perdew JP, Capelle K (2012) J Chem Phys 136:184102
223. Becke AD (2003) J Chem Phys 119(6):2972
224. Becke AD (2005) J Chem Phys 122:064101
225. Arbuznikov AV, Kaupp M (2012) J Chem Phys 136:014111
226. Paier J, Marsman M, Hummer K, Kresse G, Gerber IC, Ángyán JA (2006) J Chem Phys 124:154709
227. Gygi F, Baldereschi A (1986) Phys Rev B 34:4405
228. Broqvist P, Alkauskas A, Pasquarello A (2009) Phys Rev B 80:085114
229. Zecca L, Gori-Giorgi P, Moroni S, Bachelet GB (2004) Phys Rev B 70:205127
230. Monkhorst HJ (1979) Phys Rev B 20:1504
231. Stroppa A, Termentzidis K, Paier J, Kresse G, Hafner J (2007) Phys Rev B 76:195440
232. Gil A, Clotet A, Ricart JM, Kresse G, Garciá-Hernández M, Rösch N, Sautet P (2003) Surf Sci 530:71
233. Doll K (2004) Surf Sci 573:464
234. Bylander DM, Kleinman L (1990) Phys Rev B 41:7868
235. Heyd J, Scuseria GE, Ernzerhof M (2003) J Chem Phys 118(18):8207, 124:219906 (2006)

236. Toulouse J, Colonna F, Savin A (2004) Phys Rev A 70:062505
237. Barone V, Hod O, Peralta JE, Scuseria GE (2011) Acc Chem Res 44:269
238. Krukau AV, Vydrov OA, Izmaylov AF, Scuseria GE (2006) J Chem Phys 125:224106
239. Henderson TM, Izmaylov AF, Scalmani G, Scuseria GE (2009) J Chem Phys 131:044108
240. Ernzerhof M, Perdew JP (1998) J Chem Phys 109(9):3313
241. Henderson TM, Janesko BG, Scuseria GE (2008) J Chem Phys 128:194105
242. Heyd J, Scuseria GE (2004) J Chem Phys 120:7274
243. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA Jr, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Keith T, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2010) Gaussian 09, Revision B.01. Gaussian Inc., Wallingford
244. Kresse G, Furthmüller J (1996) Phys Rev B 54:11169
245. Marsman M, Paier J, Stroppa A, Kresse G (2008) J Phys Condens Matter 20:064201
246. Janesko BG, Henderson TM, Scuseria GE (2009) Phys Chem Chem Phys 11:443
247. Henderson TM, Janesko BG, Scuseria GE (2008) J Phys Chem A 112:12530
248. Perdew JP, Ernzerhof M, Burke K (1996) J Chem Phys 105(22):9982
249. Moussa JA, Schultz PA, Chelikowsky JR (2012) J Chem Phys 136:204117
250. Heyd J, Peralta JE, Scuseria GE, Martin RL (2005) J Chem Phys 123:174101
251. Hedin L (1965) Phys Rev 139:A796
252. Iikura H, Tsuneda T, Yanai T, Hirao K (2001) J Chem Phys 115:3540
253. Krukau AV, Scuseria GE, Perdew JP, Savin A (2008) J Chem Phys 129:124103
254. Shimka L, Harl J, Kresse G (2011) J Chem Phys 134:024116
255. Yousaf KE, Brothers EN (2010) J Chem Theory Comput 6:864
256. Song JW, Yamashita K, Hirao K (2011) J Chem Phys 135:071103
257. Perverati R, Truhlar DG (2012) Phys Chem Chem Phys 14:16187
258. Henderson TM, Izmaylov AF, Scuseria GE, Savin A (2007) J Chem Phys 127:221103
259. Lucero MJ, Henderson T, Scuseria GE (2012) J Phys Condens Matter 24:145504
260. Constantin LA, Perdew JP (2006) Phys Rev B 75:155109
261. Staroverov VN, Scuseria GE, Tao J, Perdew JP (2003) J Chem Phys 119(23):12129
262. Perdew JP, Kurth S, Zupan A, Blaha P (1999) Phys Rev Lett 82:2544, 82:5179 (1999)
263. Tao J, Perdew JP, Staroverov VN, Scuseria GE (2003) Phys Rev Lett 91:146401
264. Almeida LM, Perdew JP, Fiolhais C (2002) Phys Rev B 66:075115
265. Van Voorhis T, Scuseria GE (1998) J Chem Phys 109(2):400
266. Kanai Y, Wang X, Selloni A, Car R (2006) J Chem Phys 125:234104
267. Perdew JP, Ruzsinszky A, Constantin LA, Csonka GI, Sun J (2009) Phys Rev Lett 103:026403
268. Zhao Y, Truhlar DG (2006) J Chem Phys 125:194101
269. Zhao Y, Truhlar DG (2008) Theor Chem Acc 120:215
270. Madsen GKH, Ferrighi L, Hammer B (2010) J Phys Chem Lett 1:515
271. Sun J, Xiao B, Ruzsinszky A (2012) J Chem Phys 137:051110
272. Sun J, Haunschild R, Xiao B, Bulik I, Scuseria GE, Perdew JP (2013) J Chem Phys 138:044113
273. Sun J, Xiao B, Fang Y, Haunschild R, Hao P, Ruzsinszky A, Csonka GI, Scuseria GE, Perdew JP (2013) Phys Rev Lett 111:106401
274. Peverati R, Truhlar DG (2012) J Phys Chem Lett 3:117
275. Wheeler SE, Houk KN (2010) J Chem Theory Comput 6:395

276. Ferrighi L, Madsen GKH, Hammer B (2011) J Chem Phys 135:084704
277. Andersen M, Hornekær L, Hammer B (2012) Phys Rev B 86:085405
278. Marom N, Tkatchenko A, Rossi M, Gobre VV, Hod O, Scheffler M, Kronik L (2011) J Chem Theory Comput 7(12):3944
279. Balog R, Andersen M, Jrgensen B, Sljivancanin Z, Hammer B, Baraldi A, Larciprete R, Hofmann P, Hornekr L, Lizzit S (2013) ACS Nano 7:3823
280. Grns E, Andersen M, Arman MA, Gerber T, Hammer B, Schnadt J, Andersen JN, Michely T, Knudsen J (2013) J Phys Chem C 117(32):16438
281. Zhu M, Aikens CM, Hollander FJ, Schatz GC, Jin R (2008) J Am Chem Soc 130(18):5883
282. Ferrighi L, Pan YX, Grönbek H, Hammer B (2012) J Phys Chem C 116:7374
283. Luo S, Zhao Y, Truhlar DG (2012) J Phys Chem Lett 3:2975
284. Sun J, Marsman M, Ruzsinszky A, Kresse G, Perdew JP (2011) Phys Rev B 83:121410
285. Batista ER, Heyd J, Hennig RG, Uberuaga BP, Martin RL, Scuseria GE, Umrigar CJ, Wilkins JW (2006) Phys Rev B 74:121102(R)
286. El-Mellouhi F, Brothers EN, Lucero MJ, Scuseria GE (2013) J Phys Condens Matter 25:135501
287. Paier J, Penschke C, Sauer J (2013) Chem Rev 113:3949
288. Penschke C, Paier J, Sauer J (2013) J Phys Chem C 117:5274
289. Papageorgiou AC, Beglitis NS, Pang CL, Teobaldi G, Cabailh G, Chen Q, Fisher AJ, Hofer WA, Thornton G (2010) Proc Natl Acad Sci 107(6):2391
290. Teobaldi G, Tanimura K, Shluger AL (2010) Phys Rev B 82:174104
291. Pacchioni G (2008) J Chem Phys 128:182505
292. Vilhelmsen LB, Walton KS, Sholl DS (2012) J Am Chem Soc 134:12807
293. Yadnum S, Choomwattana S, Khongpracha P, Sirijaraensre J, Limtrakul J (2013) ChemPhysChem 14:923
294. Valero R, Gomes JRB, Truhlar DG, Illas F (2008) J Chem Phys 129:124710
295. Valero R, Gomes JRB, Truhlar DG, Illas F (2010) J Chem Phys 132:104701
296. Paier J, Marsman M, Kresse G (2007) J Chem Phys 127:024103
297. Lynch BJ, Truhlar DG (2003) J Phys Chem A 107:8996, 108:1460(E) (2004)
298. Bealing CR, Ranprasad R (2013) J Chem Phys 139:174904
299. Fu Q, Saltzburg H, Flytzani-Stephanopoulos M (2003) Science 301:935
300. Lu W, Filimonov SN, Carrasco J, Tkatchenko A (2013) Nat Commun 4:2569
301. Civarelli B, Zicovich-Wilson CM, Valenzaon L, Ugliengo P (2008) Cryst Eng Comm 10:405
302. Zicovich-Wilson CM, Kirtman B, Civalleri B, Ramírez-Solís A (2010) Phys Chem Chem Phys 12:3289
303. Civarelli B, Maschio L, Ugliengo P, Zicovich-Wilson CM (2010) Phys Chem Chem Phys 12:6382
304. Leverentz HR, Qi HW, Truhlar DG (2013) J Chem Theory Comput 9(2):995
305. Aguero A, Janesko BG (2012) J Chem Phys 136:024111
306. Janesko BG (2013) In J Quant Chem 113:83

Top Curr Chem (2015) 365: 53–80
DOI: 10.1007/128_2015_625

Published online: 21 April 2015

Short-Range Cut-Off of the Summed-Up van der Waals Series: Rare-Gas Dimers

Abhirup Patra, Bing Xiao, and John P. Perdew

Abstract van der Waals interactions are important in typical van der Waals-bound systems such as noble gas, hydrocarbon, and alkaline earth dimers. The summed-up van der Waals series of Perdew et al. 2012 works well and is asymptotically correct at large separation between two atoms. However, as with the Hamaker 1937 expression, it has a strong singularity at short non-zero separation, where the two atoms touch. In this work we remove that singularity (and most of the short-range contribution) by evaluating the summed-up series at an *effective* distance between the atom centers. Only one fitting parameter is introduced for this short-range cut-off. The parameter in our model is optimized for each system, and a system-averaged value is used to make the final binding energy curves. This method is applied to different noble gas dimers such as Ar–Ar, Kr–Kr, Ar–Kr, Ar–Xe, Kr–Xe, Xe–Xe, Ne–Ne, He–He, and also to the Be_2 dimer. When this correction is added to the binding energy curve from the semilocal density functional *meta*-GGA-MS2, we get a vdW-corrected binding energy curve. These curves are compared with the results of other vdW-corrected methods such as PBE-D2 and vdW-DF2, and found to be typically better. Binding energy curves are in reasonable agreement with those from experiment.

Keywords van der Waals · Density functional · Binding energy curves · Rare-gas dimers · Summation of asymptotic series

A. Patra (✉) and J.P. Perdew
Department of Physics, Temple University, Philadelphia, PA 19122, USA
e-mail: tuf43805@temple.edu; tuf25956@temple.edu

B. Xiao
Department of Earth Sciences, University College London, London WC1E 6BT, UK
e-mail: b.xiao@ucl.ac.uk

Contents

1 Introduction

Theoretical prediction of matter and proper explanation of many physical, chemical, and biological processes require an accurate description of atomic and molecular interactions. The only way to get a clear picture of these interactions at the atomic and molecular level is to apply quantum mechanics. Much effort has been made to develop quantum mechanical methods for this purpose. As a result there are many wave-function-based ab initio quantum mechanical methods such as Configuration Integration (CI), Many Body Perturbation Theory (MBPT), and Quantum Monte Carlo (QMC) which are popular in the scientific community. However, the Kohn–Sham (KS) [1] density functional theory (DFT) [2] has become the most popular in condensed matter physics and in quantum chemistry, because of its low computational cost and reasonable accuracy. It maps a many-electron wave-function problem to a one-electron problem. The many-electron effects are in its exchange-correlation part E_{XC}. This exchange-correlation functional E_{XC} is often approximated through satisfaction of various physical constraints.

Among numerous exchange-correlation approximations, the local spin density approximation (LSDA) [1–4], the Perdew-Burke-Ernzerhof (PBE) [5] generalized gradient approximation (GGA), and the Becke-3-Lee-Yang-Parr (B3LYP) [6] and HSE03 [7, 8] hybrid GGAs are especially popular in DFT [9] calculations for

physical and chemical systems. LSDA and PBE are efficient local and semilocal functionals widely used for extended systems, whereas PBE0 and HSE06 are hybrid functionals which hybridize a GGA with the exact exchange energy. Another hybrid functional B3LYP has a more complicated mixing of LDA and GGA exchange functionals with HF exact exchange, and its correlation energy part is also a mixing of LDA and GGA. These hybrid functionals are popular for calculations in both finite and extended systems. At the semilocal level, however, the *meta*-GGA is the highest rung of the so-called Jacob's ladder of DFT [10, 11] and potentially the most accurate one [12]. *Meta*-GGA can also serve as a better base for hybridizing with the exact exchange energy.

At the semilocal level, the E_{XC} of density functional theory can be written as [3, 4, 12–14]

$$E_{\mathrm{XC}}^{\mathrm{sl}}\left[n_{\uparrow}, n_{\downarrow}\right] = \int d^3 r n \epsilon_{\mathrm{XC}}^{\mathrm{sl}}\left(n_{\uparrow}, n_{\downarrow}, \nabla n_{\uparrow}, \nabla n_{\downarrow}, \tau_{\uparrow}, \tau_{\downarrow}\right). \tag{1}$$

In this equation, $n_{\uparrow}$, $n_{\downarrow}$ are the electron densities of spin up and down $\sigma_{\uparrow}$, $\sigma_{\downarrow}$ respectively, and the $\nabla n_{\uparrow}$, $\nabla n_{\downarrow}$ are the local gradients of the spin densities. The kinetic energy densities are $\tau_{\sigma} = 1/2\sum_k \left|\nabla \psi_{k,\sigma}\right|^2$ for the occupied KS orbitals $\psi_{k\sigma}$ of spin σ, and $\epsilon_{\mathrm{XC}}^{\mathrm{sl}}$ is the approximate exchange-correlation energy per electron. LSDA uses only $n_{\uparrow}, n_{\downarrow}$ whereas GGAs use $\nabla n_{\uparrow}$, $\nabla n_{\downarrow}$ in addition. *Meta*-GGAs [15–17] also use the kinetic energy density τ_{σ} as one of its ingredients. This τ_{σ} has information about the shell structure. The τ-dependence of *meta*-GGAs has been studied by Sun et al. [18, 19]. *Meta*-GGAs can distinguish different orbital-overlap regions by "α", defined as $\alpha = \frac{\tau - \tau^{\mathrm{W}}}{\tau^{\mathrm{unif}}}$ where $\tau = \sum_{\sigma} \tau_{\sigma}, \tau^{\mathrm{W}} = \frac{1}{8}\left|\nabla n\right|^2/n$ and $\tau^{\mathrm{unif}} = \frac{3}{10}\left(3\pi^2\right)^{2/3} n^{5/3}$. Sun et al. [20] showed that different values of α recognize three different typical regions: (1) $\alpha = 0$ in the single-orbital regime with one- and two-electron densities which characterize covalent single bonds, (2) $\alpha \approx 1$ in the slowly-varying density regime that characterizes the metallic bond, and (3) $\alpha \geq 1$ in the weakly-overlapped density region which characterizes a noncovalent bond.

In principle, DFT provides exact ground-state energies and densities, but in practice there are many situations where DFT fails to give a physical result. The long-range van der Waals interaction in rare gas dimers [21, 22], hydrocarbons, and alkaline earth diatomics is one of them. There have been many tests of density functional theory in rare gas dimers and alkaline earth dimers. Tao and Perdew [23] observed that the GGA of Perdew, Burke, and Ernzerhof (PBE) [5], the *meta*-GGA of Tao, Perdew, Staroverov, and Scuseria (TPSS) [11, 12], and its hybrid version (TPSSh) [23] all give a satisfactory and reasonable description of the short-range part of the van der Waals interaction in the van der Waals bound complexes which have strong density overlap. The authors Tao and Perdew [23] concluded that these functionals predict too-long bond lengths and too-small binding energies for the rare gas dimers, which can be improved by long-range correction of the van der Waals interaction. Ruzsinszky et al. [24] have tested non-empirical GGAs and *meta*-GGAs and found that GGAs and *meta*-GGAs tend to overbind the diatomics

with valence "s" electrons such as He_2 and Be_2 while underbinding the diatomics with valence "p" electrons such as Ar_2.

van der Waals interactions [25] are important for many material properties. The source of this weak, long-range interaction between two objects is instantaneous charge fluctuation. The van der Waals interaction is important in atomic and molecular systems, where it has many implications such as heat of sublimation of hydrocarbon molecules, chemical reaction precursor complexes, energy transfer intermediates, protein folding, stacking of nucleobases, crystal packing, and self-assembly of organic molecules. Long-range van der Waals interaction between two distinct objects requires full density-functional non-locality. There are many long-range correction methods [26–29] developed in the last few years which are good for predicting van der Waals (vdW) interaction. These also include many post Hartree–Fock (HF) methods [30–33]; see Klimeš and Michaelides [34] for an overall review of DFT-based dispersion methods. Tao et al. [35] developed a reliable approach to evaluate accurately the dynamic multipole polarizabilities and higher order vdW coefficients from electron densities and static multipole polarizabilities for spherical atoms or objects, without using any empirical fitting parameter.

Perdew et al. [36, 37] have discussed the vdWinteraction. For two spherical objects of radius R separated by a distance d, second-order perturbation theory gives the attractive long-range van der Waals interaction [38]

$$E_{\mathrm{vdW}} = -\frac{C_6}{d^6} - \frac{C_8}{d^8} - \frac{C_{10}}{d^{10}} - \ldots \tag{2}$$

The above expression is valid for $d \to \infty$. Here C_6 describes the dipole–dipole interaction, C_8 the dipole–quadrupole interaction, and C_{10} the quadrupole–quadrupole interaction and the dipole–octupole interaction. In Ruzsinszky et al. [39] these coefficients are modeled accurately and analytically for classical solid-spheres (nanoclusters) and shells (fullerenes) using the static dipole polarizability.

Furthermore, Perdew et al. [38] proved that the above asymptotic expansion can be summed to all orders for two identical spherical shells. In that work a possible two-parameter solution to the divergence problem in van der Waals interactions at very short atomic separation has also been discussed. Such divergences (which occur at $d = 0$ for any finite-order series but at $d > 0$ for the summed-up infinite series) are normally removed by a damping function [40]. Inspired by Perdew et al. [38], we suggest that a physical summation of the vdW series [35] can be used for long-range correction of semilocal density functionals, which by themselves do not have any long-range vdW interaction correction.

In this chapter we discuss a simplified cut-off approach based on the summed-up asymptotic series. This method uses only one parameter, optimized here for different systems and averaged for all systems, which can usefully provide the long-range part of the van der Waals interaction when added to the calculated binding energy curves from the *meta*-GGA-MS2 [18, 20].

2 Methods

2.1 Asymptotic van der Waals Series

The van der Waals interaction between two spherical-shell objects A and B (with thickness "t", radius "R", and electron density "ρ") can be found from (10) of Perdew et al. [38] as the infinite series

$$E(d) = -\sqrt{4\pi\rho}\sum_{k=3}^{\infty} c_k(t/R)z^k, \tag{3}$$

where d is the distance between the centers of the two objects, $z = \left(\frac{2R}{d}\right)^2$, and all equations are in atomic units. The reduced coefficient $c_k(t/R)$ is defined by $C_{2k} = c_k(t/R)\sqrt{4\pi\rho}\left[(2R)^2\right]^k$. When the geometric series $\sum_{k=1}^{\infty} z^k = (1-z)^{-1}$ for $0 \le z < 1$ is introduced and the approximation $c_k \to c_\infty$ is used for $k > 5$, we find that (3) can be summed up as [38]

$$E^{\text{geo}}(d) = -\sqrt{4\pi\rho}\left[c_3\left(\frac{t}{R}\right)z^3 + c_4\left(\frac{t}{R}\right)z^4 + c_5\left(\frac{t}{R}\right)z^5 + c_\infty\left\{(1-z)^{-1} - \sum_{k=0}^{5} z^k\right\}\right. \tag{4}$$

(See Appendices 1 and 2 for details; also see Table 1 for the values of $c_k(t/R = 1)$.)

The important message from Table 1 is that the reduced coefficients tend to a constant value as $k \to \infty$. This means that the higher-order terms of (3) can be summed as a geometric series $\sim(1-z)^{-1}$, leading to an analytic closed-form expression which sums up this asymptotic ($d \to \infty$) series. The resulting expression for $E^{\text{geo}}(d)$ (4) approximately sums the asymptotic van der Waals series to all orders in d^{-1}, but diverges at very short atomic separation when the two spheres touch

Table 1 Values of reduced van der Waals coefficients at $t/R = 1$ (solid spheres)

Values of $c_k(t/R = 1)$	
Coefficient	Values for $t/R = 1$
c_3	0.006766 (0.006766)
c_4	0.008842 (0.0101015)
c_5	0.009599 (0.01217)
c_6	0.009946
c_{10}	0.010447
c_{20}	0.010761
c_{40}	0.010904
c_{80}	0.010979
c_∞	0.011 (0.020)

Values from the Hamaker [41] expression are in parentheses. The Hamaker expression has one adjustable parameter, chosen here to make the lowest-order coefficients agree (from Perdew et al. [38])

each other. The corresponding separation is $d = R_A + R_B$ for two non-identical spheres and $d = 2R$ for two identical spheres, but the true van der Waals interaction energy remains finite there. Thus, the summed-up expression must be cut off at small d while remaining unchanged to all orders in d^{-1} at large d. The simplest way to do this is to replace d by d', where the difference between d' and d vanishes exponentially as $d \to \infty$. Perdew et al. [38] suggested a possible choice for d':

$$d' = d + g\exp[-(d - 2R)/h] \tag{5}$$

which ensures $d' > 2R$ for $g > 0$ and $h > 0$. In that work, Perdew et al. [38] have also shown that reasonable parameters for C_{60}–C_{60} are $g = 10$ and $h = 2$ atomic units. These parameters are of course system-dependent. We assume that the summed-up series is only valid when d is sufficiently greater than $2R$. Each choice of $g > 0$ and $h > 0$ defines a different asymptotic summation of (3) to all orders in d^{-1}.

In this work we have chosen instead the simplest possible form:

$$y' = y + a\exp\left(-\frac{y}{a}\right). \tag{6}$$

Here $y' = \frac{d'}{R}$, and $y = \frac{d}{R}$. $a > 2$ is a parameter. It should be noted that y' starts out as $a = \frac{y^2}{2a}$ at small y, then increases monotonically, approaching y exponentially at large y. Moreover, $a > 2$ guarantees that the singularity is removed. (To establish that y' is monotonic in y for non-negative y, just compute $\frac{dy'}{dy} = 1 - \exp\left(-\frac{y}{a}\right) \geq 0$.) We have considered $E(d')$ as an additive correction to the binding energy curve of a semilocal density functional such as GGA or *meta*-GGA. For each functional, we adjust the parameter "a" to obtain the best overall fit to known accurate reference binding energy curves. PBE GGA needs a smaller value of "a", which more significantly shortens the equilibrium bond length and strengthens the binding. The *meta*-GGA Made Simple (MGGA-MS2), which captures more of the intermediate-range vdW interaction, needs a larger value of "a", providing a correction which is less short-ranged and has less effect on the equilibrium bond length and binding energy. It should be noted that the fitting is done only for the range of d greater than or equal to the reference equilibrium bond length, because we cannot expect any useful correction of short-range errors in the functional from this approach. Here we present the vdW-corrected binding energy curves calculated by MGGA-MS2, which should give us a proper insight into the summed-up series expansion of the van der Waals interaction.

2.2 *Physical Explanation of* R

Let $R = (R_A + R_B)/2$ be the arithmetic average of the radii of the two spherical objects A and B. If the objects are classical metallic spheres with uniform density

inside and zero density outside a cutoff radius, then the radius of a sphere is clearly the only relevant length scale for that sphere. One could find polarizability radii $R_A = \left[\alpha_1^A(0)\right]^{\frac{1}{3}}$ and $R_B = \left[\alpha_1^B(0)\right]^{\frac{1}{3}}$, where $\alpha_1^A(0)$ and $\alpha_1^B(0)$ are the static dipole polarizabilities of A and B, or one could find the same radius R from $R = R^{\mathrm{hbl}}$ = half the equilibrium bond length of a dimer, because two such spheres would be attracted to one another right up to the point where they touched. In this case, there is no decay length for the density of a sphere (as in (6)), but the summed-up van der Waals interaction must still be cut off to avoid a spurious divergence to minus infinity when the spheres touch. This shows that R_A is indeed the radius of object A.

This gives us an insight: the infinite van der Waals series (2) is an asymptotic series, valid only when the "aspect ratio" R/d is small enough. This series can be summed to all orders, but that summation misses contributions which are important when the aspect ratio is not small. Only when the objects are so far apart that each "looks small" to the other is the summed-up van der Waals series accurate.

When we extend these ideas from pairs of classical metallic spheres to pairs of atoms, we can no longer expect that $R^{\mathrm{pol}} = R^{\mathrm{hbl}}$. We need to choose between these two alternatives. Simple density functionals for R are not expected to work because the static polarizability of a classical metallic sphere depends only on the radius, and not on the density inside that sphere (e.g., Perdew et al. [38] and Ruzsinszky et al. [39]). We have found empirically, for rare gas and alkaline earth dimers (Be_2), that we can obtain a reasonable long-range correction to the binding energy curve of MGGA-MS2 [19] using a system-averaged value of the fitting parameter "a" (see Table 2) along with

$$R = \max\left[R^{\mathrm{hbl}}, 1.37R^{\mathrm{pol}}\right]. \tag{7}$$

This is R^{hbl} for all considered rare gas atoms, but $1.37R^{\mathrm{pol}}$ for Be. This choice guarantees that the singularity of the summed-up vdW series is removed for any pair of spherical objects. It should be noted that this R^{hbl} is basically the half-bond length of any homo-dimer (A_2) and hetero-dimer (AB) which can be defined as

$$R_A^{\mathrm{hbl}} = R_{AA}/2 \tag{8}$$

for homo- and

Table 2 Optimized values of the fitting parameter "a" for different rare-gas dimers

Optimized values of "a" from MGGA-MS2 + vdW[d']							
Fitting parameter	Ar–Ar	Kr–Kr	Xe–Xe	Ar–Kr	Ar–Xe	Kr–Xe	System-averaged "a".
a	2.19	2.14	2.00	2.00	2.125	2.09	2.09
Optimized values of "a" from MGGA-MS2 + H[d']							
a	2.86	2.52	2.14	2.46	2.45	2.33	2.46

$$R_{AB}^{hbl} = (R_{AA} + R_{BB})/2 \tag{9}$$

for hetero-dimers. We take R_{AA} from the experimental binding energy curve.

3 Computational Details

3.1 Calculation of Dimer Binding Energy

The binding energies of different pairs of atom have been calculated in the projector augmented wave approach (PAW) [42] implemented in the Vienna ab initio simulation package (VASP) [43–45] within *meta* generalized gradient approximation (MGGA-MS2) for the exchange-correlation functional. In the calculations, the kinetic energy cutoff is first set to be 600 eV, and a Γ centered $1 \times 1 \times 1$ *k*-point mesh in the BZ is used for the *k*-space integration. The total energy of the atomic pairs has been calculated using a $10 \times 20 \times 10$ rectangular super-cell. The PAW scheme is utilized with the potentials taken from the VASP PBE library. The energy of an isolated atom is calculated by a $10 \times 10 \times 10$ cubic super-cell. The binding energy is calculated in the following way:

$$\Delta E = -[E(R_{dimer}) - 2 \times E(R_{Atom})]. \tag{10}$$

We have also used other van der Waals methods, PBE-D2 [27] and vdW-DF2 [46], to perform this calculation. All these DFT methods are available in VASP. Our long-range-corrected *meta*-GGA-MS2 seems to give better results than these two vdW-corrected functionals.

3.2 Optimization of the Fitting Parameter

The average error (AE) calculation and the additive correction to the DFT results have been made in computer codes. The inputs to these codes are the binding energy curves calculated using DFT ($E_{MGGA\text{-}MS2}$), static dipole polarizabilities, and electron densities for each atom of the pair, and van der Waals coefficients C_6, C_8, C_{10}. Reference values for comparison are the experimental binding energy curves [47, 48].

There is a strong singularity near $d = 2R$ in the summed-up van der Waals series, as discussed before. To remove it, we replace y and y' in (6) by d/R and d'/R, respectively, to get

$$d' = d + aR\exp\left[-\frac{d}{aR}\right] \tag{11}$$

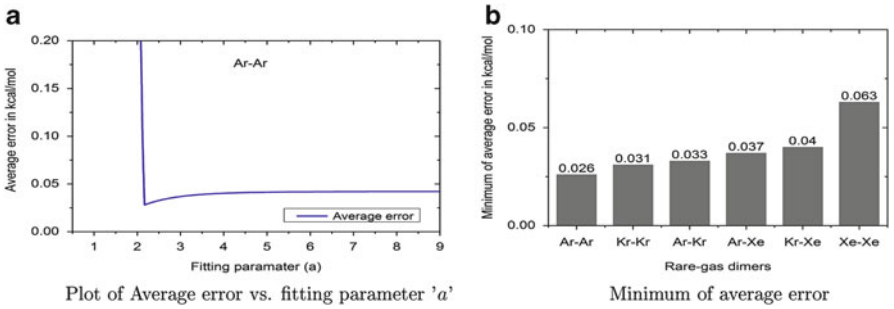


Fig. 1 (**a**) Plot of average error vs fitting parameter "a" for Ar–Ar. (**b**) Plot of minimum average error for different rare gas dimers

with R^{hbl} for rare-gas dimers and $R = 1.37R^{\text{pol}}$ for Be_2. This d' is used to calculate the summed-up $E(d')$. In this section, we use $E(d')$ from (14) below as the long-range van der Waals correction to be added to the MGGA-MS2 binding energy curve.

We have optimized the fitting parameter "a" by minimizing the error between the binding energy curve from the vdW-corrected method and the experimental binding energy curve over the set of d's in the range $d_{\min}$ Å $< d < 9$ Å in steps of 0.1 Å for each dimer. Here $d_{\min}$ Å is the separation at the minimum in the experimental binding energy curve. The fitting parameter "a" in (11) is varied in the range $0.0 < a < 10.0$ in steps of 0.001 to maintain accuracy. The average error (AE) at each value of "a" is calculated using

$$\text{AE}|_a = \left[\sum_{d'=d_{\min}}^{9} \left|E_{\text{MGGA-MS2+vdW}}\left(d'\right) - E_{\text{Exp}}\right|\right] / \left(\text{Total number of } d' \text{ points}\right). \tag{12}$$

Figure 1a shows the average error vs fitting parameter "a" plotted for the Ar_2 dimer. Figure 1b shows a histogram plot of minimum average error (MAE) for different rare gas dimers. MAE is the minimum of the average error at the optimized value of "a" for each system. So this is a discrete-column graph where each column gives the error at the minimizing value of "a".

3.3 Calculation of the van der Waals Interaction Corrected Binding Energy

To avoid the system dependence of the fitting parameter, the system-averaged value of the fitting parameters "a" was used to recalculate all the binding energy curves.

We found that the system-averaged fitting parameter gives us even better correction in most cases. In this work, all the binding energy curves are calculated using the system-averaged fitting parameter "a_{sys}" (see Table 2). Replacing "a" with "a_{sys}", (11) becomes

$$d' = d + a_{sys} R \exp\left(-\frac{d}{a_{sys} R}\right). \tag{13}$$

Finally, the vdW-corrected energy is calculated by replacing $E(d)$ by $E(d')$. The correction part is calculated using Hamaker's expression $E^{H}(d)$ as modified in Perdew et al. [38]:

$$E_{vdW}\left(d'\right) = \left(-\frac{C_6}{d'^6} - \frac{C_8}{d'^8} - \frac{C_{10}}{d'^{10}}\right) + \left(\frac{C_\infty}{C_\infty^H}\right)\left[E^H\left(d'\right) + \frac{C_6^H}{d'^6} + \frac{C_8^H}{d'^8} + \frac{C_{10}^H}{d'^{10}}\right] \tag{14}$$

and then added to the MGGA-MS2 result

$$\begin{aligned} E_{\text{MGGA-MS2+vdW}}\left(d'\right) &= E_{vdW}\left(d'\right) + E_{\text{MGGA-MS2}}(d) \\ &= \left(-\frac{C_6}{d'^6} - \frac{C_8}{d'^8} - \frac{C_{10}}{d'^{10}}\right) + \left(\frac{C_\infty}{C_\infty^H}\right)\left[E^H\left(d'\right) + \frac{C_6^H}{d'^6} + \frac{C_8^H}{d'^8} + \frac{C_{10}^H}{d'^{10}}\right] \\ &\quad + E_{\text{MGGA-MS2}}(d). \end{aligned} \tag{15}$$

It should be noted that by $E^H(d)$ we mean Hamaker's [41] classical expression for the van der Waals interaction energy $E^H(R_A, R_B, d)$ between two spherical objects of radii R_A, and R_B when the two centers are separated by a distance "d". This $E^H(R_A, R_B, d)$ is given by the following equation:

$$E^H(R_A, R_B, d) = -\frac{\pi^2 \beta}{6}\left\{\frac{2R_A R_B}{d^2 - (R_A + R_B)^2} + \frac{2R_A R_B}{d^2 - (R_A - R_B)^2} + \ln\left[\frac{d^2 - (R_A + R_B)^2}{d^2 - (R_A - R_B)^2}\right]\right\} \tag{16}$$

where $\beta = c_3(1)\sqrt{4\pi\rho}\left(\frac{3}{4\pi}\right)^2 2^6 = 0.006766\sqrt{4\pi\rho}\left(\frac{3}{4\pi}\right)^2 2^6$ is evaluated using the value of $c_3(1)$ from Table 1. The electron density is $\rho = \frac{N}{\left(\frac{4}{3}\pi R^3\right)}$ for a sphere with radius R. N is the total number of valence electrons (2 for He and 8 for the other rare gas atoms). For non-identical spheres, we replace $\sqrt{\rho}$ by $2\sqrt{\rho_A}\sqrt{\rho_B}/\left(\sqrt{\rho_A} + \sqrt{\rho_B}\right)$.

The van der Waals coefficients (C_6, C_8, C_{10}) used in the first part of the right-hand side of (15) are taken from their tabulated values in the supporting information of [35]. All these standard van der Waals coefficients are calculated by time-dependent Hartree-Fock theory. The van der Waals coefficients from

Hamaker's formula ($C_6^{\mathrm{H}}, C_8^{\mathrm{H}}, C_{10}^{\mathrm{H}}$) in the second part of the right-hand side of the same equation are evaluated by extracting the coefficients of d^{-6}, d^{-8}, and d^{-10} from the Taylor series expansion of the expression of $E^{\mathrm{H}}(R_{\mathrm{A}}, R_{\mathrm{B}}, d)$ of (16) in power of d^{-1}:

$$E^{\mathrm{H}}(R_{\mathrm{A}} = R_{\mathrm{B}} = R, d) = -\frac{\pi^2\beta}{6}\left(\frac{32R_{\mathrm{A}}^3R_{\mathrm{B}}^3}{3d^6} + \frac{32R_{\mathrm{A}}^3R_{\mathrm{B}}^3\left(R_{\mathrm{A}}^2 + R_{\mathrm{B}}^2\right)}{d^8} + \frac{64\left(5R_{\mathrm{A}}^7R_{\mathrm{B}}^3 + 14R_{\mathrm{A}}^5R_{\mathrm{B}}^5 + 5R_{\mathrm{A}}^3R_{\mathrm{B}}^7\right)}{5d^{10}}\right) + \ldots \tag{17}$$

For two identical spheres we can easily obtain the values of $C_6^{\mathrm{H}}, C_8^{\mathrm{H}}, C_{10}^{\mathrm{H}}$ from the simpler expression $E^{\mathrm{H}}(R, d) = -\frac{\pi^2\beta}{6}\left(\frac{32R^6}{3d^6} + \frac{64R^8}{d^8} + \frac{1536R^{10}}{5d^{10}}\right) + \ldots$.

3.4 Calculation of Binding Energy Using the Unmodified Hamaker Expression

Simpler than in (15), Hamaker's [41] expression can be used without modification as a long-range vdW additive correction to the DFT results. We have used

$$\begin{aligned} E_{\text{MGGA-MS2+H}}\left(d'\right) &= E^{\mathrm{H}}\left(d'\right) + E_{\text{MGGA-MS2}}(d) \\ &= \left[-\frac{\pi^2\beta}{6}\left\{\frac{2R_{\mathrm{A}}R_{\mathrm{B}}}{d'^2 - \left(R_{\mathrm{A}} + R_{\mathrm{B}}\right)^2} + \frac{2R_{\mathrm{A}}R_{\mathrm{B}}}{d'^2 - \left(R_{\mathrm{A}} - R_{\mathrm{B}}\right)^2}\right.\right. \\ &\quad \left.\left. + \ln\left[\frac{d'^2 - \left(R_{\mathrm{A}} + R_{\mathrm{B}}\right)^2}{d'^2 - \left(R_{\mathrm{A}} - R_{\mathrm{B}}\right)^2}\right]\right\}\right] + E_{\text{MGGA-MS2}}(d). \end{aligned} \tag{18}$$

Here β is calculated in the same way as is done for $E^{\mathrm{H}}(R_{\mathrm{A}}, R_{\mathrm{B}}, d)$ as discussed in the above section. (See Appendix 1 for an explanation of $E^{\mathrm{H}}(d)$.) In this version, no input vdW coefficient is needed.

4 Results and Discussion

We believe MGGA-MS2 gives better result than other semilocal GGA and *meta*-GGA functionals (PBE,TPSS) in our calculation because it uses "α" [18–20] which can recognize and describe the intermediate-range vdW interaction [20].

4.1 Ar–Ar Dimer

In Fig. 2 for Ar_2, binding energy curves are calculated using MGGA-MS2, MGGA-MS2 + vdW[d], MGGA-MS2 + vdW[d'], MGGA-MS2 + H[d], MGGA-MS2 + H[d'], vdW-DF2, and PBE-D2, and are compared with the experimental results [47]. The binding energy (ΔE in kcal/mol) and the equilibrium distance (R_e in Å) for different methods are tabulated in Table 3 along with the experimental results. Figure 2a shows that the MGGA-MS2 + vdW[d'] is in very good agreement with the experimental curve, whereas MGGA-MS2 binds slightly less than the experimental curve. In this figure the MGGA-MS2 + vdW[d] curve is calculated by adding the van der Waals correction part to the MGGA-MS2 result, but without using the cut-off distance d', and thus it clearly depicts the divergence near $d = 2R$. The interesting thing about this graph is that MGGA-MS2 gives a comparatively good intermediate-range part of the binding energy curve as expected [19, 20], slightly overestimating the equilibrium bond length ~3.75 Å compared to the experimental one at 3.76 Å. Figure 2a also shows that at larger atomic separation MGGA-MS2 significantly underbinds the experimental binding energy curve. Figure 2b shows

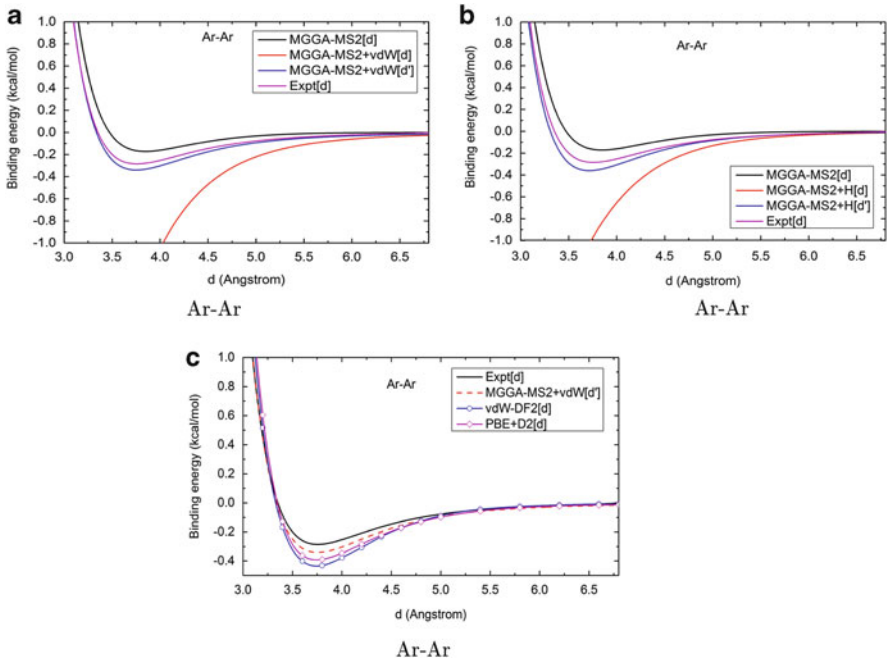


Fig. 2 Binding energy curves for the Ar–Ar dimer calculated from the vdW method combined with MGGA-MS2 using the system-averaged $a = 2.09$. (**a**) Binding energy curves for MGGA-MS2, MGGA-MS2 + vdW[d], MGGA-MS2 + vdW[d'], and experiment. (**b**) Binding energy curves for MGGA-MS2, MGGA-MS2 + H[d], MGGA-MS2 + H[d'], and the experimental curve. (**c**) Comparison of the vdW-DF2 [46] and PBE-D2 [27] curves with the experimental one

Table 3 Binding energy (ΔE in kcal/mol) and equilibrium bond length (R_e in Å) for the Ar–Ar dimer

Quantity	MGGA-MS2	MGGA-MS2 + vdW(d')	MGGA-MS2 + H(d')	PBE-D2	vdW-DF2	Expt.
ΔE	0.17206	0.34025	0.36057	0.38889	0.43207	0.28500
R_e	3.85	3.75	3.72	3.80	3.80	3.76

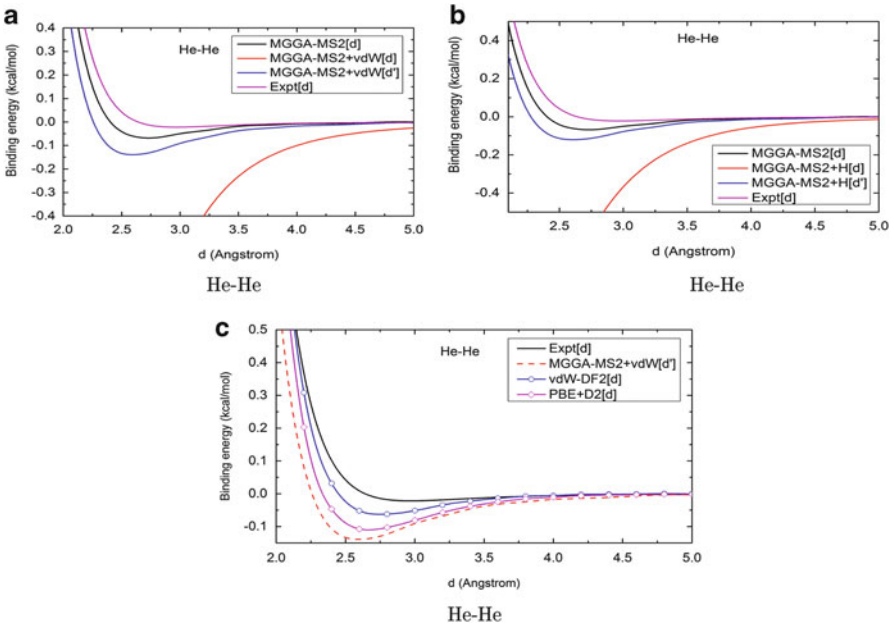


Fig. 3 Binding energy curves of the He–He dimer. Here we have also used the cut-off distance d' for long-range correction with the system-averaged $a = 2.09$. (**a**, **b**) Binding energy curves for MGGA-MS2, MGGA-MS2 + vdW[d], MGGA-MS2 + vdW[d'], MGGA-MS2 + H[d], and MGGA-MS2 + H[d']. (**c**) Comparison of the curves for vdW-DF2 [46] and PBE-D2 [27] with the experimental curve

that when the unmodified Hamaker (E^{H}) expression is used, it also gives a very good long-range part of the interaction potential, although MGGA-MS2 + H[d'] slightly overbinds the experimental curve. The divergence in the additive correction part (see MGGA-MS2 + H[d] curve) can be seen as for the vdW correction part in Fig. 2a. We also calculated the binding energy curves of the Ar–Ar dimer with the two popular van der Waals corrected methods vdW-DF2 [46] and PBE-D2 [27]. The binding energy curves from these calculations along with the MGGA-MS2 + vdW[d'] and experimental results can be seen in Fig. 2c. Both vdW-DF2 and PBE-D2 overbind the experimental binding energy curve in the range 3.3 Å $\leq R_e \leq$ 5.4 Å. Figure 3c also shows that MGGA-MS2 + vdW[d'], PBE-D2, and vdW-DF2 almost overlap with the experimental curve beyond 5.10 Å.

Table 4 Binding energy (ΔE in kcal/mol) and equilibrium bond length (R_e in Å) for the He–He dimer

Quantity	MGGA-MS2	MGGA-MS2 + vdW(d')	MGGA-MS2 + H(d')	PBE-D2	vdW-DF2	Expt.
ΔE	0.06849	0.13913	0.12038	0.10742	0.06214	0.02201
R_e	2.73	2.59	2.61	2.60	2.80	2.97

4.2 He–He Dimer

Figure 3 shows the binding energy curves for He_2 from MGGA-MS2, MGGA-MS2 + vdW[d], MGGA-MS2 + vdW[d'], MGGA-MS2 + H[d], MGGA-MS2 + H[d'], vdW-DF2, and PBE-D2 calculations, along with the experimental results [47]. As predicted by Ruzsinszky et al. [24], nonempirical GGAs and *meta*-GGAs tend to overbind van der Waals-bound diatomics that have valence "s" electrons, such as He_2. Figure 3a shows that MGGA-MS2 and MGGA-MS2 + vdW[d'] overbind the experimental curve. When we plot MGGA-MS2 + vdW[d] we observe a divergence similar to the one for the Ar dimer because, in this case, the cut-off d' has not been used to remove the singularity. This graph also shows very good long-range behavior of MGGA-MS2 + vdW [d']. Figure 3b shows an almost similar picture where MGGA-MS2 + H [d'] also overbinds the experimental curve. In this case, a divergence in the Hamaker (E^H) expression can be seen. From Table 4 we note that the strong attractive nature of MGGA-MS2 yields a minimum at 2.73, 2.59, and 2.61 Å for MGGA + MS2, MGGA + MS2 + vdW[d'], and MGGA-MS2 + H[d'], respectively, which is quite different from the experimental one at 2.97 Å. From Fig. 3c we can see that two other popular methods, PBE-D2 [27] and vdW-DF2 [46], overbind the experimental curve in the same way as MGGA-MS2 + vdW(d'), but all these three curves almost overlap with the experimental curve in the very long-range part beyond 4.00 Å, a significant success for the method.

4.3 Xe–Xe Dimer

In the Xe–Xe dimer, MGGA-MS2, MGGA-MS2 + vdW[d'], and MGGA-MS2 + H [d'] underbind the experimental curve, which can be seen in Fig. 4a, b. These two graphs also show similar divergences for MGGA-MS2 + vdW[d] and MGGA-MS2 + H[d] because of the singularity near $d = 2R$ in both the vdW correction part and in the Hamaker expression. Figure 4c compares PBE-D2 [27], vdW-DF2 [46], and MGGA-MS2 + vdW[d'] with the experimental curve. From these three figures, and from Table 5, it can be said that, where PBE-D2 and vdW-DF2 overestimate the binding energy, MGGA-MS2 + vdW[d'] and MGGA-MS2 + H[d'] underestimate the same, but all these four methods give very different minima than the experimental one.

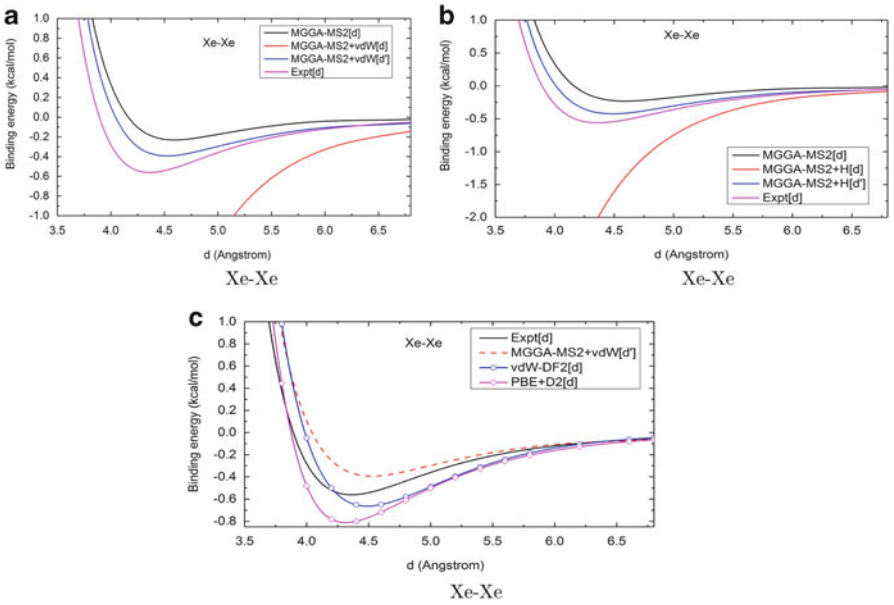


Fig. 4 (**a**) Binding energy curves for MGGA-MS2, MGGA-MS2 + vdW[d], and MGGA-MS2 + vdW[d']. (**b**) Binding energy curves for MGGA-MS2, MGGA-MS2 + H[d], and MGGA-MS2 + H[d']. (**c**) Binding energy curves using vdW-DF2 [46] and PBE-D2 [27] plotted with the experimental one

Table 5 Binding energy (ΔE in kcal/mol) and equilibrium bond length (R_e in Å) for the Xe–Xe dimer

Quantity	MGGA-MS2	MGGA-MS2 + vdW(d')	MGGA-MS2 + H(d')	PBE-D2	vdW-DF2	Expt.
ΔE	0.23100	0.39302	0.42232	0.80012	0.65163	0.56057
R_e	4.59	4.53	4.54	4.40	4.40	4.35

4.4 Kr–Kr Dimer

Figure 5a–c shows the binding energy curves for the Kr–Kr dimer for MGGA-MS2, MGGA-MS2 + vdW[d], MGGA-MS2 + vdW[d'], MGGA-MS2 + H[d], MGGA-MS2 + H[d'], vdW-DF2, and PBE-D2 approaches. These curves show that both MGGA-MS2 + vdW[d'] and MGGA-MS2 + H[d'] give a very satisfactory long-range part of the van der Waals interaction. It should be noted that these two curves cross the experimental curve at R(Kr–Kr) ~4.1 Å and at ~3.77 Å. One can readily note the divergence of MGGA-MS2 + H[d] at small atomic separation because of the obvious singularity in the expression of E^{H}. Table 6 gives the binding energy

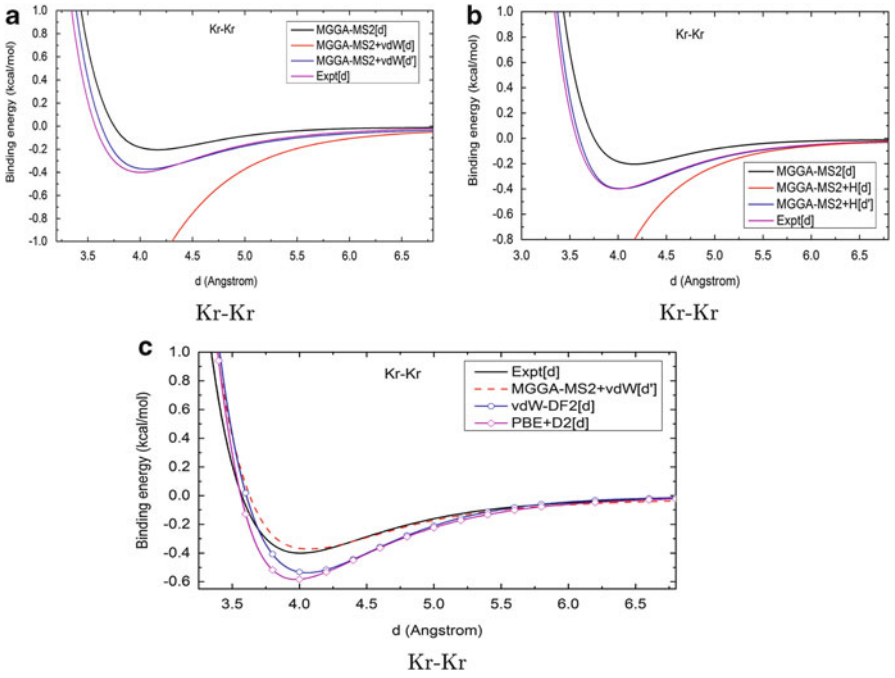


Fig. 5 (**a**) Binding energy curves for MGGA-MS2, MGGA-MS2 + vdW[d], and MGGA-MS2 + vdW[d']. (**b**) Binding energy curves for MGGA-MS2, MGGA-MS2 + H[d], and MGGA-MS2 + H[d']. (**c**) Binding energy curves using vdW-DF2 [46] and PBE-D2 [27] plotted with the experimental one. The vdW corrected curves are calculated using the cut-off distance d' and system-averaged $a = 2.09$

Table 6 Binding energy (ΔE in kcal/mol) and equilibrium bond length (R_e in Å) for the Kr–Kr dimer

Quantity	MGGA-MS2	MGGA-MS2 + vdW(d')	MGGA-MS2 + H(d')	PBE-D2	vdW-DF2	Expt.
ΔE	0.20379	0.37102	0.39750	0.53832	0.53242	0.40005
R_e	4.17	4.11	4.03	4.00	4.00	4.01

(ΔE in kcal/mol) and the equilibrium distance (R_e in Å) for different methods. We notice that MGGA-MS2 + H[d'] gives very good equilibrium bond length when compared to the experimental result. Figure 5c confirms that another two methods, PBE-D2 [27] and vdW-DF2 [46], significantly overbind the experimental curve. We can see the same trend in MGGA-MS2 + vdW[d'] along with PBE-D2 and vdW-DF2 in the very long-range part, where they tend to overlap with the experimental result.

4.5 *Ne–Ne Dimer*

In the Ne–Ne dimer, MGGA-MS2, MGGA-MS2 + vdW [d'], and MGGA-MS2 + H [d'] overbind the experimental curve. MGGA-MS2 + H[d] and MGGA-MS2 + vdW [d] show the divergence at small atomic separation. However, Fig. 6a, b shows the similar kind of long-range correction by both MGGA-MS2 + vdW[d'] and MGGA-MS2 + H[d']. We plot the PBE-D2 and vdW-DF2 result with the experimental and MGGA-MS2 + vdW[d'] results in Fig. 6c. We conclude that, for the Ne dimer, semilocal functionals normally overbind the experimental result. From Fig. 6c we find that the long-range part of the MGGA-MS2 + vdW[d'] and vdW-DF2 curves almost overlap with the experimental curve in the range 3.9 Å $\leq R_e \leq$ 5.0 Å. Table 7 shows numerical values.

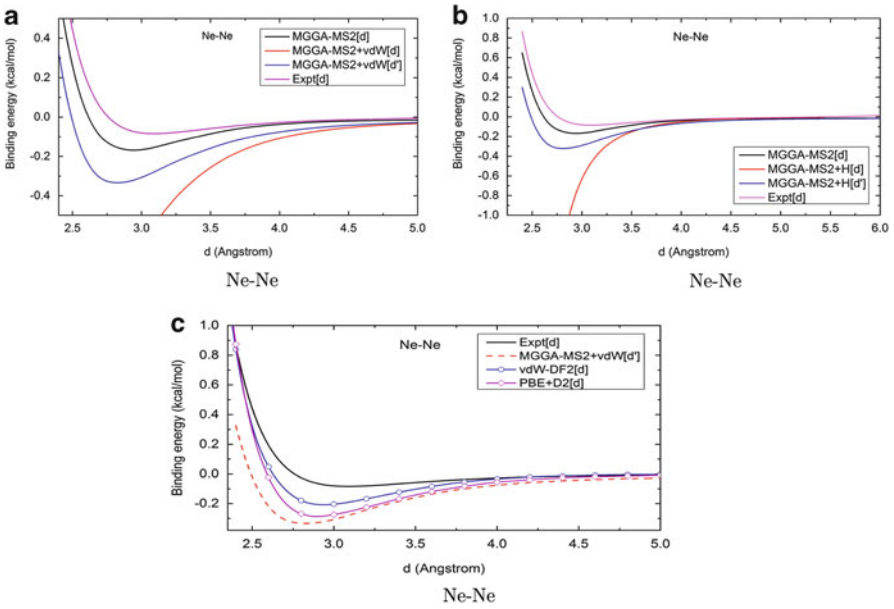


Fig. 6 Binding energy curves for the Ne–Ne dimer: (**a**) for MGGA-MS2, MGGA-MS2 + vdW[d], and MGGA-MS2 + vdW[d']; (**b**) for MGGA-MS2, MGGA-MS2 + H[d], and MGGA-MS2 + H[d']. (**c**) Comparison of the calculated binding energy curves using vdW-DF2 [46] and PBE-D2 [27] with the experimental curve

Table 7 Binding energy (ΔE in kcal/mol) and equilibrium bond length (R_e in Å) for the Ne–Ne dimer

Quantity	MGGA-MS2	MGGA-MS2 + vdW(d')	MGGA-MS2 + H(d')	PBE-D2	vdW-DF2	Expt.
ΔE	0.16811	0.33323	0.32208	0.27369	0.20395	0.08401
R_e	2.94	2.83	2.81	3.00	3.00	3.09

4.6 Ar–Kr Dimer

MGGA-MS2 underbinds the experimental curve [48] in the Ar–Kr dimer, as can be seen in Fig. 7a, b. This tendency is removed and we get very satisfactory binding energy curves when the cut-off distance d' is introduced in the MGGA-MS2 + vdW[d'] and MGGA-MS2 + H[d'] methods. This removes the singularity as expected and gives significant improvement in the long-range part of the van der Waals interaction. The strong divergence in van der Waals interaction series and in Hamaker's expression can be seen in MGGA-MS2 + vdW[d] and in MGGA-MS2 + H[d] in Fig. 7a, b. PBE-D2 [27] and vdW-DF2 [46] results are not so satisfactory compared to the experimental result. Table 8 shows the binding energy (ΔE in kcal/mol) and the equilibrium distance (R_e in Å) for different methods.

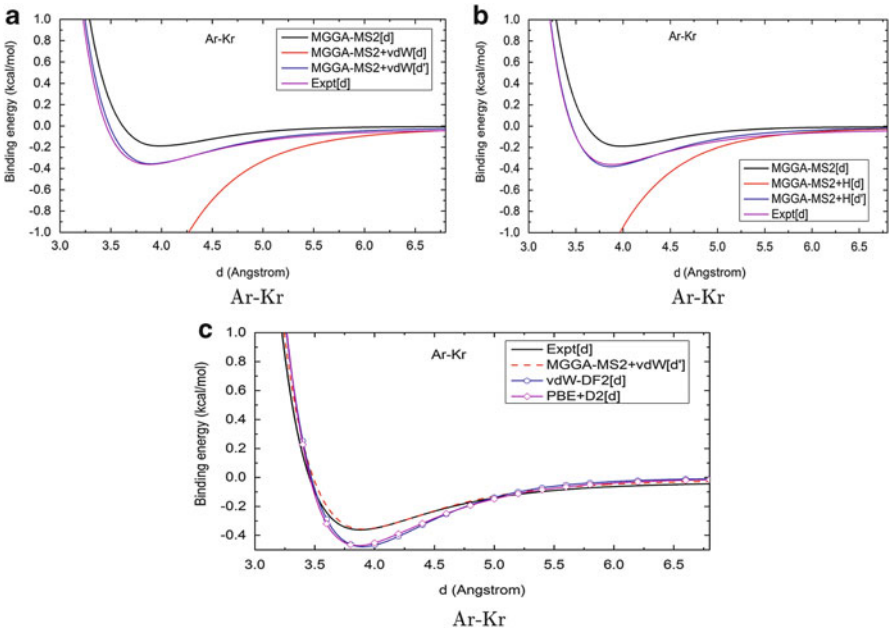


Fig. 7 Binding energy curves of the Ar–Kr dimer calculated: (**a**) using MGGA-MS2, MGGA-MS2 + vdW[d], and MGGA-MS2 + vdW[d']; (**b**) using MGGA-MS2, MGGA-MS2 + H[d], MGGA-MS2 + H[d'], and the experimental result. (**c**) Comparison of vdW-DF2 [46] and PBE-D2 [27] with experiment. The vdW corrected binding energy curves are calculated using the cut-off distance d' and system-averaged $a = 2.09$

Table 8 Binding energy (ΔE in kcal/mol) and equilibrium bond length (R_e in Å) for the Ar–Kr dimer

Quantity	MGGA-MS2	MGGA-MS2 + vdW(d')	MGGA-MS2 + H(d')	PBE-D2	vdW-DF2	Expt.
ΔE	0.18772	0.35712	0.38053	0.46469	0.46481	0.36123
R_e	3.98	3.89	3.88	3.80	4.00	3.88

4.7 Ar–Xe Dimer

We plot the binding energy curves of the Ar–Xe dimer in Fig. 8a, b. These figures show outstanding performance of the method in the long-range part of the energy when either of the additive corrections $E_{\mathrm{vdW}}(d')$ or $E^{\mathrm{H}}(d')$ is used. Table 9 gives a qualitative picture of different methods for the estimation of equilibrium bond length and binding energy. It can be inferred from Table 9 that both MGGA-MS2 + vdW [d'] and MGGA-MS2 + H[d'] produce almost correct binding energy but slightly overestimate the equilibrium bond length.

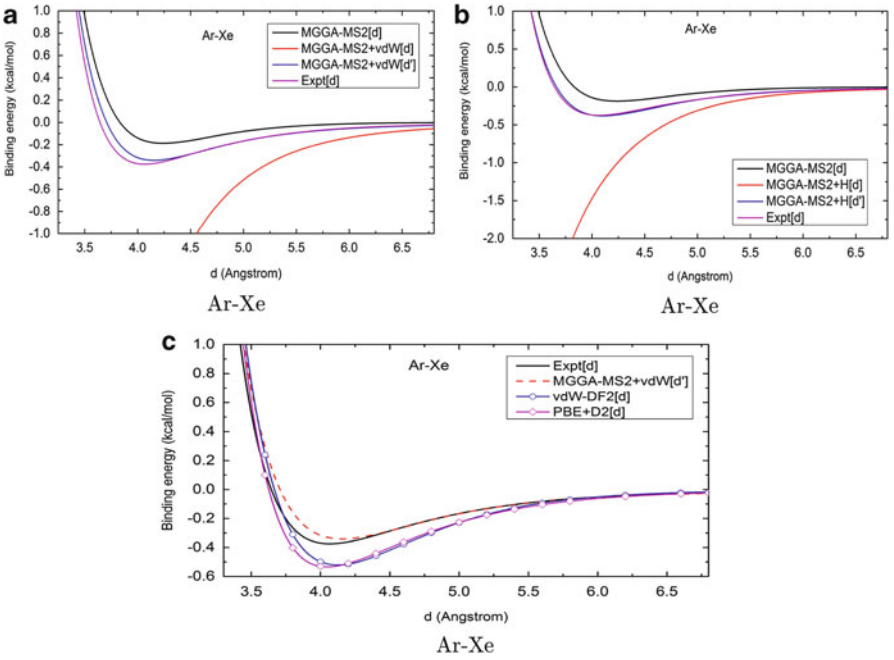


Fig. 8 Binding energy curves of the Ar–Xe dimer calculated from the vdW method combined with MGGA-MS2 using the system-averaged a. (**a**) Binding energy curves calculated using MGGA-MS2, MGGA-MS2 + vdW[d], and MGGA-MS2 + vdW[d'], and the experimental one. The vdW-corrected binding energy curve is calculated using the cut-off distance d' and system-averaged $a = 2.09$. (**b**) Binding energy curves for MGGA-MS2, MGGA-MS2 + H[d], and MGGA-MS2 + H[d'], and the experimental results. (**c**) Binding energy curves from other vdW-corrected methods

Table 9 Binding energy (ΔE in kcal/mol) and equilibrium bond length (R_{e} in Å) for the Ar–Xe dimer

Quantity	MGGA-MS2	MGGA-MS2 + vdW(d')	MGGA-MS2 + H(d')	PBE-D2	vdW-DF2	Expt.
ΔE	0.18677	0.34032	0.38191	0.53029	0.51532	0.37505
R_{e}	4.24	4.10	4.10	4.00	4.20	4.07

4.8 *Kr–Xe Dimer*

As in the Ar–Kr dimer, in the Kr–Xe dimer MGGA-MS2 underbinds the experimental curve. Figure 9a, b shows that introduction of the cut-off distance d' not only removes the singularity but also improves the underbinding of MGGA-MS2 in the long-range part of the van der Waals interaction energy, although it remarkably overestimates the equilibrium bond length. The binding energy estimated by MGGA-MS2 + H [d'] is in very good agreement with the experimental result. The PBE-D2, vdW-DF2, MGGA-MS2 + vdW[d'], and experimental curves are shown in Fig. 9c. PBE-D2 and vdW-DF2 seem to be too attractive in the short-range part and overbind the experimental curve, although the equilibrium bond lengths from these two curves are in very good agreement with experiment. Table 10 shows numerical values.

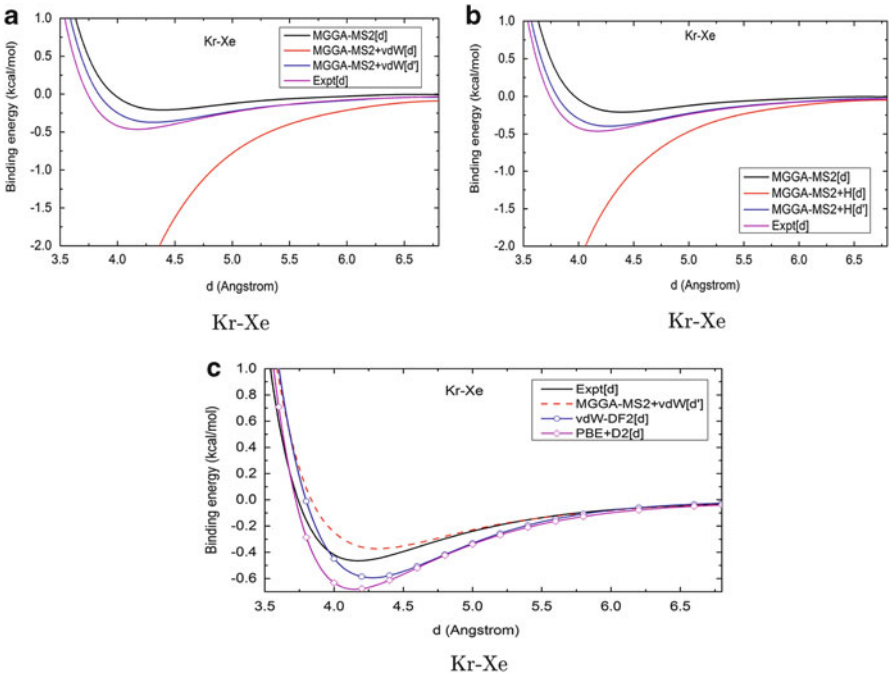


Fig. 9 Binding energy curves of the Kr–Xe dimer calculated from the vdW method combined with MGGA-MS2 using the system-averaged $a = 2.09$. (**a**) Binding energy curves for MGGA-MS2, MGGA-MS2 + vdW[d], and MGGA-MS2 + vdW[d'], and the experimental one. (**b**) Binding energy curves for MGGA-MS2, MGGA-MS2 + H[d], and MGGA-MS2 + H[d'], and the experimental results. (**c**) Comparison of vdW-DF2 [46] and PBE-D2 [27] with experiment

Table 10 Binding energy (ΔE in kcal/mol) and equilibrium bond length (R_e in Å) for the Kr–Xe dimer

Quantity	MGGA-MS2	MGGA-MS2 + vdW(d')	MGGA-MS2 + H(d')	PBE-D2	vdW-DF2	Expt.
ΔE	0.20900	0.37242	0.39787	0.67692	0.58528	0.46422
R_e	4.39	4.31	4.27	4.20	4.20	4.18

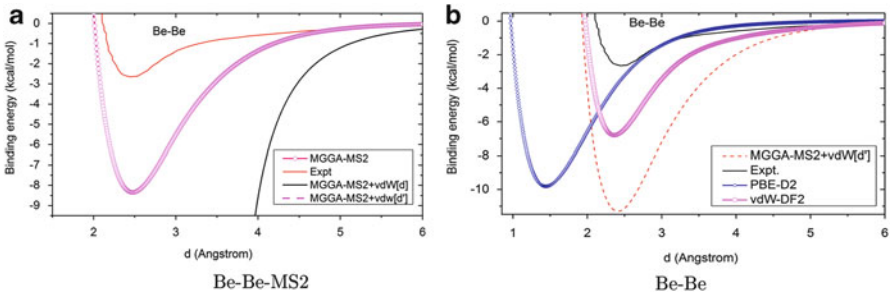


Fig. 10 Binding energy curves for the Be–Be dimer calculated (**a**) from *meta*-GGA-MS2 and (**b**) from other vdW-corrected methods

4.9 *Be–Be Dimer*

Be_2 is chosen because it is a van der Waals-bound diatomic from the alkaline earth group. The characteristic difference of Be_2 from the rare gas dimers is that it shows much more density overlap. Figure 10a gives us a qualitative picture of how the strong divergence at $d = 2R$ in the $E^{\mathrm{H}}(d)$ part of the van der Waals interactions energy expression of (15) is successfully removed upon introduction of the cut-off distance d'. Figure 10b shows that MGGA-MS2 + vdW [d'] gives a satisfactory long-range correction compared to PBE-D2. In the same figure the similar performance of vdW-DF2 can be seen. The common tendency of overbinding the experimental result [49] by semilocal functionals MGGA-MS2 and PBE [24] can be seen in Fig. 10b. From this result we conclude that more accurate correction is needed for semilocal functionals for He_2, Ne_2, and Be_2. Moreover, Fig. 10b shows that vdW-DF2 correctly estimates the equilibrium bond length ~2:5 Å, where MGGA-MS2 + vdW[d'] fails to do so. It overestimates the bond length at ~2.7 Å. The experimental equilibrium bond length is 2.45 Å. Careful observation of Fig. 10b reveals that the too-attractive nature of PBE yields a deep minimum at very short distance ~1.5 Å for the PBE-D2 curve.

4.10 *Comparison Between Hamaker and Geometric Series for Two Identical Solid Spheres*

An interesting feature of the Hamaker expression [41] and the geometric expression [38] (see (25) in Appendix 1) is that both have divergences at two different values of d, one at $d = 2R$ and the other at $d = 0$. To investigate further the divergence of these two expressions, we have plotted $c_\infty \left\{ \left(1 - \left(\frac{2R}{d}\right)^2\right)\right)^{-1} - \sum_{k=0}^{2} \left(\frac{2R}{d}\right)^{2k} \right\}$ and $\left(\frac{c_\infty}{c_\infty^{\mathrm{H}}}\right) E^{\mathrm{H}}(R_{\mathrm{A}} = R_{\mathrm{B}} = R, d)$. These two quantities have the same singularity at $d = 2R$ but they are rather different in the range of typical vdW energies, which can be seen in Fig. 11. This could explain why the Hamaker expression works better in the fits.

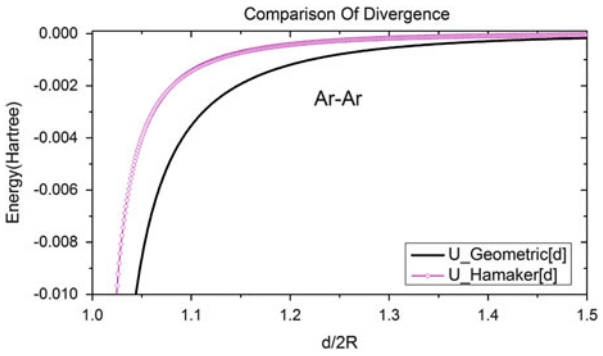


Fig. 11 Divergence of Hamaker's [41] expression and the geometric expression of Perdew et al. [38] for Ar–Ar dimer

5 Conclusions

In summary, our work is a reasonable and physical way to remove the nonzero d singularity in the summed-up van der Waals interaction series. We also conclude that, although the fitting parameter "a" is system-dependent and empirical in nature, averaging it and using $R = \max(R^{\text{hbl}}, 1.37R^{\text{po1}})$ as an input to our calculation gives our formulas some predictive value. In this work we have presented a simplified version with only one fitting parameter a, replacing the earlier model [38] with two parameters (g, h), to get the long-range van der Waals correction to a density-functional binding energy curve for objects with spherical densities. We have reached similar or better accuracy when compared with the PBE-D2 [27] and vdW-DF2 [46] correction schemes.

An interesting outcome of our work is that the complicated (15), as proposed in Perdew et al. [38], can be replaced without loss of accuracy by the simpler (18). This means in particular that no input vdW coefficient is needed.

One extension of this work could be to use it to obtain a semi-local density functional with an embedded long-range vdW correction using this summed-up series method. The positive outcome of this method motivates us to use it for other strongly dispersion-driven systems such as layered-materials. Furthermore, MGGA-MS2 often provides a useful description of intermediate-range vdW interaction. It underestimates this in most rare gas dimers, but overestimates it in He_2, Ne_2, and Be_2. Appreciable performance by MGGA-MS2 [19] in predicting stacking energies between nucleobases of DNA and RNA confirms that improved density functionals can give a better description of different chemical and physical properties. It remains to be seen whether further improvements in *meta*-GGAs yield a more consistent description of intermediate-range vdW interactions.

Acknowledgments This work was supported in part by the National Science Foundation under Grant No. DMR-1305135. We thank Prof. Adrienn Ruzsinszky and Dr. Jianwei Sun for their useful suggestions.

Appendix 1 Summed-Up Series Expression

The Casimir–Polder [50, 51] formula for the van der Waals coefficients between two objects A and B to the second-order in electron–electron interaction is

$$C_{2k}^{\mathrm{AB}} = \frac{(2k-2)!}{2\pi}\sum_{l_1=1}^{k-2}\frac{1}{(2l_1)!(2l_2)!}\int_0^\infty du\alpha_{l_1}^{\mathrm{A}}(iu)\alpha_{l_2}^{\mathrm{B}}(iu), \tag{19}$$

where $l_2 = k - l_1 - 1$ and $\alpha_{l_1}^{\mathrm{A}} \to 2^{l_1}$ pole dynamic polarizability of A at imaginary frequency $\omega = iu$. It should be noted that $l = 1$, $l = 2$, $l = 3$ are for dipole, quadrupole, and octupole interactions, respectively.

The dynamic multipole polarizabilities for a classical conducting spherical shell can be found from the following expressions:

$$\alpha_l(iu) = R^{2l+1}\frac{\omega_l^2}{{\omega_l}^2 + u^2}\frac{1-\theta_l}{1-\beta_l\theta_l}, \tag{20}$$

where

$$\beta_l = \frac{{\omega_l}^2\widetilde{\omega}_l^2}{({\omega_l}^2 + u^2)(\widetilde{\omega}_l^2 + u^2)} \tag{21}$$

and

$$\theta_l = \left(\frac{R-t}{R}\right)^{2l+1} = (1 - t/R)^{2l+1} \tag{22}$$

from the work of Lucas et al. [52]. Here $\omega_l = \omega_{\mathrm{p}}\sqrt{l/(2l+1)}$ and $\widetilde{\omega}_l = \omega_{\mathrm{p}}\sqrt{(l+1)/(2l+1)}$. The plasma frequency of the system is $\omega_{\mathrm{p}} = \sqrt{4\pi\rho}$ with $\rho = \frac{N}{\left[\frac{4}{3}\pi\left\{R^3-(R-t)^3\right\}\right]}$ for the spherical shell (with radius R and thickness t) and $\rho = \frac{N}{\left(\frac{4}{3}\pi R^3\right)}$ for the sphere (with radius R). N is the total number of valence electrons, equal to 2 for He and 8 for other rare gas atoms.

In Perdew et al. [37] it is shown that, for a classical conducting spherical shell of radius R, thickness t, and uniform density ρ, the above integration at (18) can be performed to get all the higher order vdW coefficients. For two identical spheres, i.e., when A = B, one can get

$$C_{2k}^{\mathrm{AA}} = \omega_{\mathrm{p}}(2R)^{2k}\frac{1}{2^{2k}}\frac{(2k-2)!}{4}\sum_{l=1}^{k-2}\frac{1}{(2l)!(2k-2l-1)!} \times\frac{1}{\sqrt{(2l+1)/l} + \sqrt{(2k-2l-1)/(k-l-1)}}. \tag{23}$$

Then the van der Waals interaction of (2) can be written as

$$E(d) = -\sqrt{4\pi\rho}\sum_{k=3}^{\infty} c_k(t/R) z^k, \tag{24}$$

where $c_k(t/R)$ is related to C_{2k} by (4) and $z = \left(\frac{2R}{d}\right)^2$.

Now, by the introduction of the geometric series of $\sum_{k=1}^{\infty} z^k = (1-z)^{-1}$ for $0 \le z < 1$ and approximating $c_k \to c_\infty$ for k > 5, we find

$$E^{\text{geo}}(d) = -\sqrt{4\pi\rho}\left[c_3\left(\frac{t}{R}\right)z^3 + c_4\left(\frac{t}{R}\right)z^4 + c_5\left(\frac{t}{R}\right)z^5 + c_\infty\left\{(1-z)^{-1} - \sum_{k=0}^{5} z^k\right\}\right]. \tag{25}$$

This expression interpolates between the very large d and $d \to 2R$ limits. The above expression for $E^{\text{geo}}(d)$ has an unphysical divergence at $z = 1$ or $d = 2R$ where the two spheres touch each other. This divergence appears because we sum up all the terms. However, in reality there is no divergence in (2) because it is an asymptotic expansion for large value of d.

This is true because at large d the exponential density overlap between the two real quantum-mechanical objects may be neglected. This divergence in the expression of $E^{\text{geo}}(d)$ can be removed by replacing z by z' where $z' = (2R/d')^2$ with a proper choice of d'. The expression for $E^{\text{geo}}(d)$ is true for the interaction between identical spheres but it can be generalized to non-identical spheres $2R \to R_A + R_B$, which leads to an equation such as (14) for the expression of $E_{\text{vdW}}(d')$.

In the pair interaction picture, Hamaker's [41] expression of the van der Waals interaction between two solid spheres of uniform density ρ is

$$E(d) = -\beta\int_A d^3r \int_B d^3r' \frac{1}{|r - r'|^6}, \tag{26}$$

where $\beta = c_3(1)\sqrt{4\pi\rho}\left(\frac{3}{4\pi}\right)^2 2^6 = 0.006766\sqrt{4\pi\rho}\left(\frac{3}{4\pi}\right)^2 2^6$ can be evaluated using the value of $c_3(1)$ from Table 1.

Appendix 2 Binding Energy Curves from Geometric Series

The summed-up van der Waals series expression of (4) can also be used to obtain the binding energy curves for the rare-gas dimers if we use our short-range cut-off idea. Reduced van der Waals coefficients $c_3(t/R)$, $c_4(t/R)$, $c_5(t/R)$ and $c_\infty(t/R)$ are taken from Table 1 for $t/R = 1$, e.g., for solid spheres. For identical solid-spheres the electron density is $\rho = N/\left(4\pi R^3/3\right)$ for a sphere with radius R and number of total valence electrons N (N is 2 for He and 8 for the other rare-gas atoms). The electron density for non-identical spheres can be evaluated using $2\sqrt{\rho_A}\sqrt{\rho_B}/\left(\sqrt{\rho_A} + \sqrt{\rho_B}\right)$.

We could not optimize the fitting parameter for every dimer. We have used $a = 4.09$, the average of the optimum values for Ar–Ar and Kr–Kr. Figure 12 shows the results.

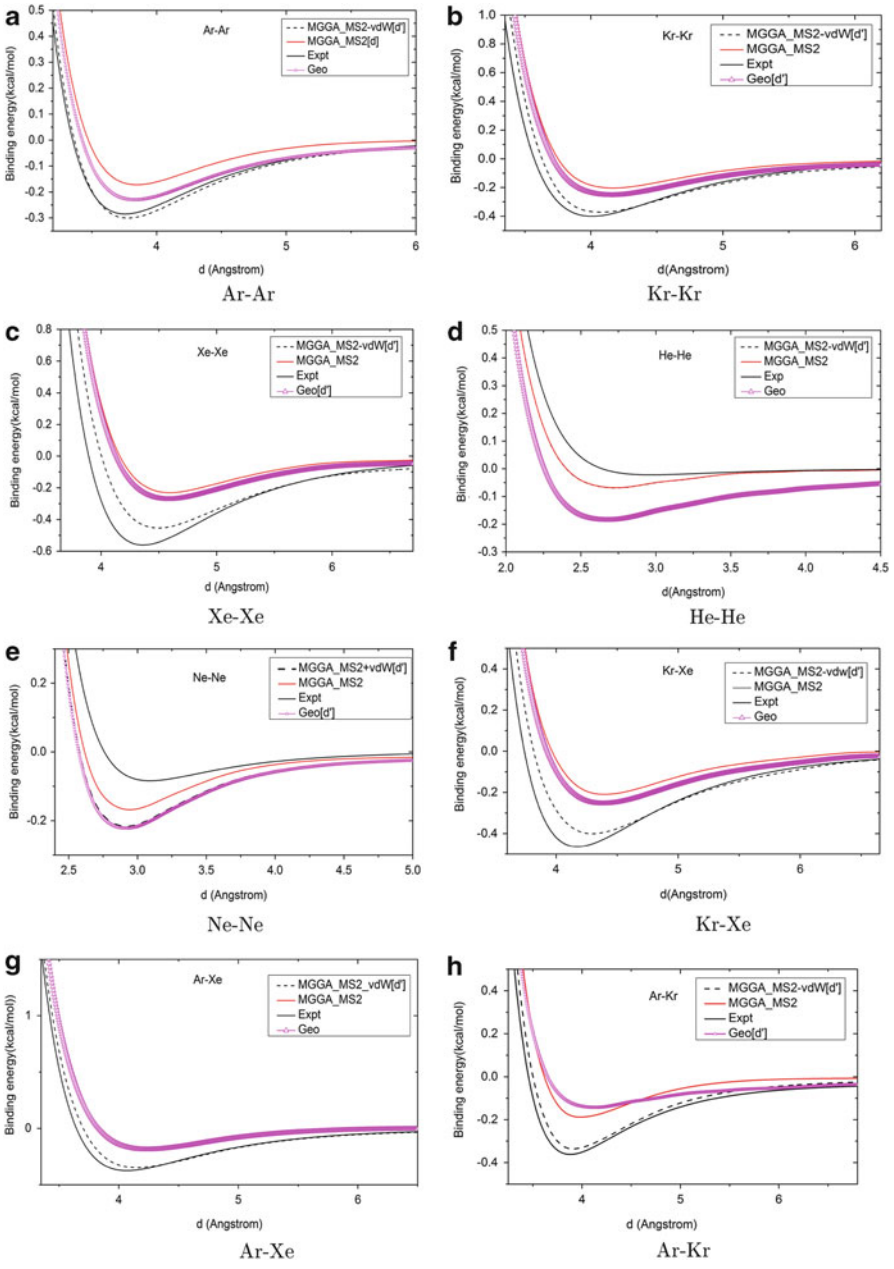


Fig. 12 Binding energy curves for different dimers using the geometric-series expression of (25) $a = 4.09$

References

1. Kohn W, Sham LJ (1965) Self-consistent equations including exchange and correlation effects. Phys Rev 140:A1133–A1138
2. Perdew JP, Kurth S (2003) A primer in density functional theory, Lecture notes in physics. Springer, Berlin
3. Perdew JP, Wang Y (1992) Accurate and simple analytic representation of the electron-gas correlation energy. Phys Rev B 45:13244–13249
4. Sun J, Perdew JP, Seidl M (2010) Correlation energy of the uniform electron gas from an interpolation between high- and low-density limits. Phys Rev B 81(8):085123
5. Perdew JP, Burke K, Ernzerhof M (1996) Generalized gradient approximation made simple. Phys Rev Lett 77:3865–3868
6. Becke AD (1993) Density functional thermochemistry, III. The role of exact exchange. J Chem Phys 98(7):5648–5652
7. Heyd J, Scuseria GE, Ernzerhof M (2003) Hybrid functionals based on a screened Coulomb potential. J Chem Phys 118(18):8207–8215
8. Paier J, Marsman M, Hummer K, Kresse G, Gerber IC, Ángyán JG (2006) Screened hybrid density functionals applied to solids. J Chem Phys 124(15):154709
9. Burke K (2012) Perspective on density functional theory. J Chem Phys 136(139):150901
10. Perdew JP, Schmidt K (2001) Jacob's ladder of density functional approximations for the exchange-correlation energy. AIP Conf Proc 577(1):1–20
11. Tao J, Perdew JP, Staroverov VN, Scuseria GE (2003) Climbing the density functional ladder: nonempirical meta-generalized gradient approximation designed for molecules and solids. Phys Rev Lett 91:146401
12. Perdew JP, Ruzsinszky A, Csonka GI, Vydrov OA, Scuseria GE, Constantin LA, Zhou X, Burke K (2008) Restoring the density-gradient expansion for exchange in solids and surfaces. Phys Rev Lett 100:136406
13. Perdew JP, Ruzsinszky A, Csonka GI, Constantin LA, Sun J (2009) Workhorse semilocal density functional for condensed matter physics and quantum chemistry. Phys Rev Lett 103:026403
14. Staroverov VN, Scuseria GE, Tao J, Perdew JP (2003) Comparative assessment of a new nonempirical density functional: molecules and hydrogen-bonded complexes. J Chem Phys 119:12129
15. Perdew JP, Ruzsinszky A, Tao J, Csonka GI, Scuseria GE (2007) One-parameter optimization of a nonempirical meta-generalized-gradient-approximation for the exchange-correlation energy. Phys Rev A 76(4):042506
16. Tao J, Perdew JP, Ruzsinszky A, Scuseria GE, Csonka GI, Staroverov VN (2007) Meta-generalized gradient approximation: non-empirical construction and performance of a density functional. Philos Mag 87(7):1071–1084
17. Constantin LA, Perdew JP, Tao J (2006) Meta-generalized gradient approximation for the exchange-correlation hole with an application to the jellium surface energy. Phys Rev B 73 (20):205104
18. Sun J, Xiao B, Ruzsinszky A (2012) Communication: effect of the orbital-overlap dependence in the meta generalized gradient approximation. J Chem Phys 137(5):051101
19. Sun J, Xiao B, Fang Y, Haunschild R, Hao P, Ruzsinszky A, Csonka GI, Scuseria GE, Perdew JP (2013) Density functionals that recognize covalent, metallic, and weak bonds. Phys Rev Lett 111(10):106401
20. Sun J, Haunschild R, Xiao B, Bulik IW, Scuseria GE, Perdew JP (2013) Semilocal and hybrid *meta*-generalized gradient approximations based on the understanding of the kinetic-energy-density dependence. J Chem Phys 138(4):044113
21. Zhao Y, Truhlar DG (2006) Comparative DFT study of van der Waals complexes: rare-gas dimers, alkaline-earth dimers, zinc dimer, and zinc-rare-gas dimers. J Phys Chem A 110 (15):5121–5129

22. Tran F, Hutter J (2013) Nonlocal van der Waals functionals: the case of rare-gas dimers and solids. J Chem Phys 138(20):204103
23. Tao J, Perdew JP (2005) Test of a non-empirical density functional: short-range part of the van der Waals interaction in rare-gas dimers. J Chem Phys 122(11):114102
24. Ruzsinszky A, Perdew JP, Csonka GI (2005) Binding energy curves from nonempirical density functionals II. van der Waals bonds in rare-gas and alkaline-earth diatomics. J Phys Chem A 109(48):11015–11102
25. Stone AJ (1996) The theory of intermolecular forces, vol 32, International series of monographs on chemistry. Oxford University Press, Oxford
26. Wu Q, Ayers PW, Yang W (2003) Density-functional theory calculations with correct long-range potentials. J Chem Phys 119:2978
27. Grimme S (2006) Semiempirical GGA-type density functional constructed with a long-range dispersion correction. J Comput Chem 27:1787
28. Sato T, Tsuneda T, Hirao K (2007) Long-range corrected density functional study on weakly bound systems: balanced descriptions of various types of molecular interactions. J Chem Phys 126:234114/1
29. Vydrov OA, Scuseria GE (2006) Assessment of a long-range corrected hybrid functional. J Chem Phys 125:234109/1
30. Grimme S, Ehrlich S, Goerigk L (2011) Effect of the damping function in dispersion corrected density functional theory. J Comput Chem 32(7):1456–1465
31. Johnson ER, Becke AD (2005) A post-Hartree–Fock model of intermolecular interactions. J Chem Phys 123(2):024101
32. Tkatchenko A, Scheffler M (2009) Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. Phys Rev Lett 102(7):073005
33. Vydrov OA, Van Voorhis T (2010) Dispersion interactions from a local polarizability model. Phys Rev A 81(6):062708
34. Klimeš J, Michaelides A (2012) Perspective: advances and challenges in treating van der Waals dispersion forces in density functional theory. J Chem Phys 137(12):120901
35. Tao J, Perdew JP, Ruzsinszky A (2012) Accurate van der Waals coefficients from density functional theory. Proc Natl Acad Sci 109(1):18–21
36. Tao J, Perdew JP, Ruzsinszky A (2013) Long-range van der Waals interaction. Int J Modern Phys B 27(18):30011
37. Perdew JP, Tao J, Hao P, Ruzsinszky A, Csonka GI, Pitarke JM (2012) Spherical-shell model for the van der Waals coefficients between fullerenes and/or nearly spherical nanoclusters. J Phys Condens Matter 24(42):424207
38. Perdew JP, Ruzsinszky A, Sun J, Glindmeyer S, Csonka GI (2012) van der Waals interaction as a summable asymptotic series. Phys Rev A 86(6):062714
39. Ruzsinszky A, Perdew JP, Tao J, Csonka GI, Pitarke JM (2012) van der Waals coefficients for nanostructures: fullerenes defy conventional wisdom. Phys Rev Lett 109(23):233203
40. Wu Q, Yang W (2002) Empirical correction to density functional theory for van der Waals interactions. J Chem Phys 116(2):515–524
41. Hamaker H (1937) The London-van der Waals attraction between spherical particles. Physica 4(10):1058–1072
42. Blöchl PE (1994) Projector augmented-wave method. Phys Rev B 50(24):17953
43. Kresse G, Hafner J (1993) *Ab initio* molecular dynamics for liquid metals. Phys Rev B 47:558–561
44. Kresse G, Hafner J (1994) *Ab initio* molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium. Phys Rev B 49:14251–14269
45. Kresse G, Joubert D (1999) From ultrasoft pseudopotentials to the projector augmented-wave method. Phys Rev B 59:1758–1775
46. Lee K, Murray ED, Kong L, Lundqvist B, Langreth D (2010) Higher-accuracy van der Waals density functional. Phys Rev B 82:081101

47. Ogilvie JF, Wang FY (1992) Potential-energy functions of diatomic molecules of the noble gases I. Like nuclear species. J Mol Struct 273:277–290
48. Ogilvie JF, Wang FY (1993) Potential-energy functions of diatomic molecules of the noble gases: II. Unlike nuclear species. J Mol Struct 291(2):313–322
49. Schmidt MW, Ivanic J, Ruedenberg K (2010) Electronic structure analysis of the ground-state potential energy curve of Be2. J Phys Chem A 114(33):8687–8696
50. Casimir HBG, Polder D (1948) The influence of retardation on the London-van der Waals forces. Phys Rev 73:360–372
51. Patil SH, Tang KT (1997) Multipolar polarizabilities and two- and three-body dispersion coefficients for alkali isoelectronic sequences. J Chem Phys 106(6):2298–2305
52. Lucas AA, Henrard L, Lambin P (1994) Computation of the ultraviolet absorption and electron inelastic scattering cross section of multishell fullerenes. Phys Rev B 49:2888–2896

Top Curr Chem (2015) 365: 81–96
DOI: 10.1007/128_2014_600

Published online: 24 February 2015

Judging Density-Functional Approximations: Some Pitfalls of Statistics

Andreas Savin and Erin R. Johnson

Abstract Density-functional theory (DFT) methods have achieved widespread popularity for thermochemical predictions, which has lead to extensive benchmarking of functionals. While the use of statistics to judge the quality of various density-functional approximations is valuable and even seems unavoidable, the present chapter suggests some pitfalls of statistical analyses. Several illustrative examples, focusing on analysis of thermochemistry and intermolecular interactions, are presented to show that conclusions can be heavily influenced by both the data-set size and the choice of the criterion used to assess an approximation's quality. Even with reliable approximations, the risk of publishing inaccurate results naturally increases with the number of calculations reported.

Keywords Density-functional theory · Intermolecular interactions · Statistics · Thermochemistry

Contents

A. Savin
CNRS, UPMC Univ Paris 06, UMR7616, Laboratoire de Chimie Théorique, 75005 Paris, France

E.R. Johnson (✉)
Chemistry and Chemical Biology, School of Natural Sciences, University of California, Merced, 5200 North Lake Road, Merced, CA 95343, USA
e-mail: ejohnson29@ucmerced.edu

1 Introduction

A combination of the high performance of modern computers, fast algorithms, and good accuracy have permitted widespread use of density-functional approximations (DFAs) across chemistry and physics. However, the high number of DFAs often causes users of the theory to pose the question: "Which functional should I use?" Nowadays, large sets of reference data are available to provide valuable help in answering this question. Consequently, there has been a recent surge of benchmarking studies where readers can take their pick of statistical measures to justify use of their chosen functional in a particular application. The rapid growth in the numbers and citations of benchmarking studies is illustrated in Fig. 1.

Habitually, statistical measures such as the mean absolute error are used to indicate the quality of a density-functional approximation. Unfortunately, statistics can also deform reality, and using statistics to judge the quality of an approximation is no exception. A recently published paper [1] sought to answer several questions regarding benchmarking of DFAs:

1. Is the approximation the only source of error and would an exact treatment give the right result?
2. Do the approximations provide the necessary quantitative accuracy?
3. Are we interested in obtaining absolute values or in reproducing trends?

In general, the user is provided with a selection of DFAs and must decide which to choose. However, there is no single objective criterion to determine which DFA is best for a particular problem and a somewhat subjective choice is made regarding, for example, the statistical criteria used to rank the functionals. Which specific functional is then qualified as the best depends on this choice [1].

In the spirit of reference [2], we consider the performance of selected DFAs for thermochemistry and intermolecular interactions. In these cases, accurate data

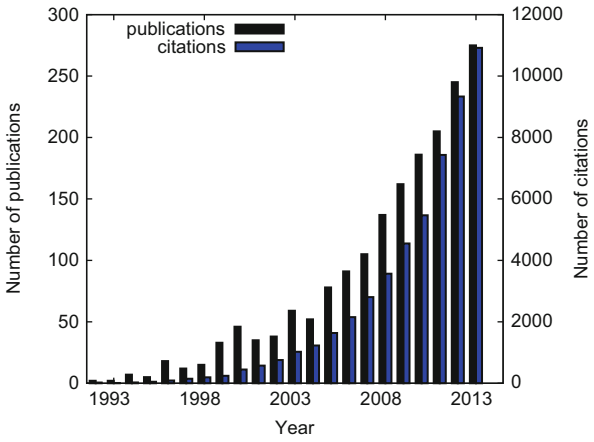


Fig. 1 Growth in the number of published density-functional benchmarking studies and the total citations, as determined from the Thomson Reuters Web of Science Core Database

should be accessible via coupled-cluster theory with large basis sets and extrapolation. We address the following questions:

1. How strong are the effects caused by the finite sample size?
2. Can the same data produce opposite interpretations?
3. What is the probability that all of the results published in an article have sufficient quantitative accuracy?

In the process, we discuss some potential pitfalls of statistics that rank the quality of density-functional approximations.

2 Methodology

In a previous study [3], various dispersion-corrected density functionals were benchmarked against either experimental or high-level ab initio reference data for intermolecular complexes, thermochemistry, and reaction barrier heights. All data sets used in the present work are taken from this prior study. The types and sources of the reference data are summarized in Table 1. These data sets are by no means a comprehensive collection of DFT benchmarks and this chapter focuses on the illustrative examples of thermochemistry and intermolecular interaction data sets only.

The density-functional approximations considered herein are also the same as in our earlier work [3]. The acronyms follow the usual notation related to the names of the authors: BLYP [13, 14], B3LYP [14, 15], BH&HLYP [14, 16], B97-1 [17], CAM-B3LYP [18], LC-ωPBE [19, 20], PBE [21], PBE0 [22], and PW86PBE [21, 23]. To ensure that the conclusions are not adversely impacted by the basis set, aug-cc-pVTZ bases were used as they give results close to the complete basis set limit for conventional density-functional calculations. In all cases, the exchange-hole dipole moment (XDM) dispersion correction was applied [3, 24], but this does *not* appear in the acronyms used below. Note that the two empirical parameters used in the dispersion damping function were fitted separately for each of the functionals, with the aug-cc-pVTZ basis set, which also tends to correct deficiencies of the base DFA.

To assess the quality of the various DFAs for each data set, three statistical error measures are used: the mean absolute errors (MAE):

$$\mathrm{MAE} = \overline{x} = \frac{1}{n}\sum_{i=1,n} x_i, \tag{1}$$

the mean absolute percent errors (MAPE):

Table 1 List of reference data, showing the abbreviation of the data-set name, the source of the data (either calculated or experimental), the number of data points, a brief description of the set, and the relevant literature reference

Name	Type	No.	Description	References
KB49	calc	49	Intermolecular interactions	[3, 4]
S22	calc	22	Intermolecular interactions	[5, 6]
S66	calc	66	Intermolecular interactions	[7, 8]
HSG	calc	21	Intermolecular interactions	[6, 9]
G1	expt	56	Atomization energies	[10]
G2	expt	149	Atomization energies	[11]
G3	expt	222	Atomization energies	[24]

$$\mathrm{MAPE} = \frac{100}{n} \sum_{i=1,n} \frac{x_i}{|x_{\mathrm{ref},i}|} \tag{2}$$

and the root-mean-square errors (RMSE):

$$\mathrm{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1,n} x_i^2} \tag{3}$$

where n is the number of data points in the set and $x_i = |x_{\mathrm{calc},i} - x_{\mathrm{ref},i}|$ is the absolute error for each point. Finally, the sample variance is

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1,n} (x_i - \bar{x})^2 \tag{4}$$

and the mean absolute error has a variance equal to the sample variance divided by the size of the sample, or σ^2/n. The uncertainty of the MAE is thus $\sigma/\sqrt{n}$.

3 Using Statistical Measures to Judge Density Functional Approximations

When judging the performance of DFAs, the mean error does not tell the full story and further statistical measures are available to describe the distribution of errors for a particular benchmark set. As an example, let us consider the G3/99 data set [12] which gauges the ability of the functionals to describe atomization energies. Histograms of the error distribution for the G3 set are shown in Fig. 2 for selected functionals. From the figure, we see that BLYP and B3LYP have narrow error distributions, while the BHandHLYP distribution is much broader. The maximum errors are 27.3 kcal/mol for BLYP, 34.4 kcal/mol for B3LYP, and 79.7 kcal/mol for

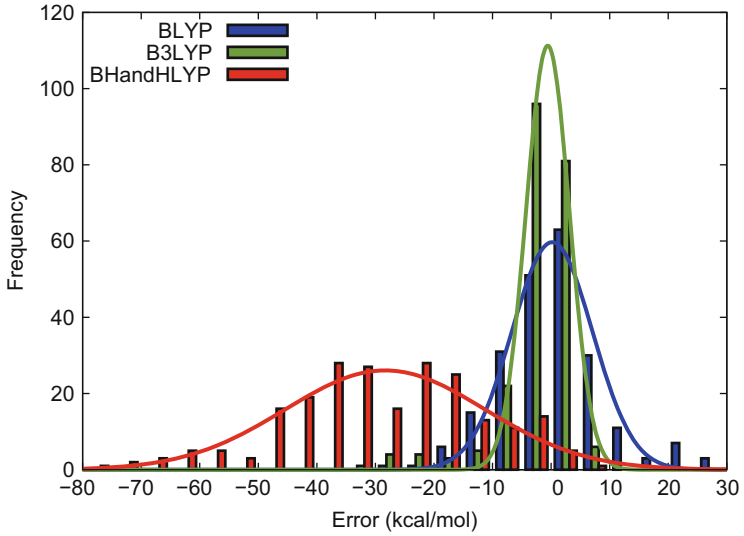


Fig. 2 Histograms of the signed errors for the G3 set with selected functionals. The lines show fits of the data to Gaussian distributions and are included to guide the eye

BHandHYLP. Thus, even for a relatively narrow error distribution, the maximum errors are far from the desired chemical accuracy.

Analogous results can be presented more compactly via "box-and-whisker" plots, where the boxes span the interquartile range of the data (i.e., the range that spans the middle half of the data), the ends of the whiskers show the minimum and maximum errors, and the lines show the median errors. Error distributions are plotted in this fashion for all the DFAs considered in Fig. 3. The figure also shows an analogous plot for the errors expressed on a per-atom basis, which eliminates any bias of the error distribution based on molecular size.

When benchmarking density functionals, mean absolute errors are most often presented to indicate the quality of an approximation. However, the average can be greatly inflated by a few outliers in a data set. The root-mean-square error (RMSE) is even more strongly affected, while the median absolute error reduces this bias. Table 2 collects these statistics for the G3 set. B3LYP gives the lowest MAE and median absolute error, while CAM-B3LYP gives the lowest RMSE and B97-1 the lowest maximum error. As a further alternative, one can also consider the MAE, RMSE, and median or maximum absolute error per atom, to account for differences in molecular size within the benchmark. With the per-atom statistics, B3LYP, CAM-B3LYP, LC-ωPBE, B97-1, and PBE0 all give MAEs of near 1 kcal/mol per atom. B97-1 gives the lowest RMSE, B3LYP gives the lowest median absolute error, and CAM-B3LYP the lowest maximum error. Thus, the choice of which statistical indicator is used to judge the DFAs determines which is ultimately selected as the best functional.

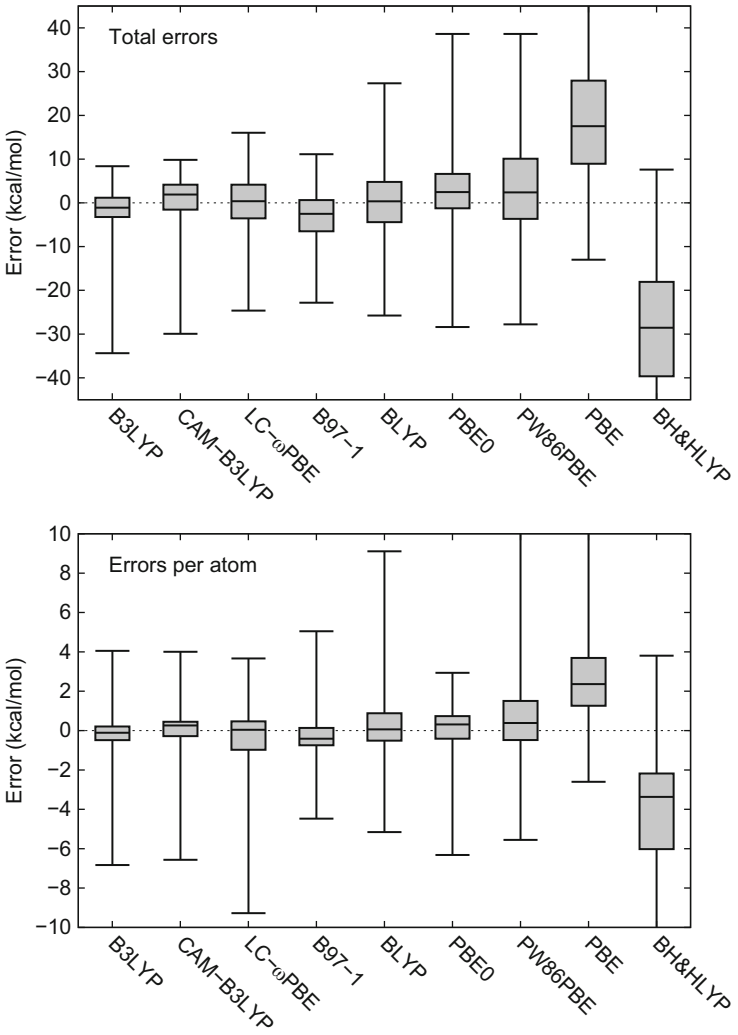


Fig. 3 Box-and-whisker plots of the atomization-energy errors for the G3 set with selected functionals. The *boxes* span the interquartile range, the ends of the *whiskers* show the minimum and maximum errors, and the *lines* show the median errors. The *upper panel* shows the total errors and the *lower panel* shows the errors per atom

With each DFA, the distribution of errors for the G3 set is very broad and the decay of the absolute errors is very slow, so that the mean is not well defined. To demonstrate what is meant by this, consider a linear molecule of $n+1$ atoms, forming n chemical bonds. For each bond energy, we let the error with a particular DFA be x kcal/mol. Then the mean error for chains of size $n = 1, 2, \ldots, m$ is

Table 2 MAEs, RMSEs, median absolute errors (Med), and maximum absolute errors (Max) for the G3 set, in kcal/mol. The same quantities, expressed per atom, are also shown

	Total				Per atom			
Functional	MAE	RMSE	Med	Max	MAE	RMSE	Med	Max
B3LYP	4.0	6.7	2.2	34.4	0.8	1.5	0.3	6.8
CAM-B3LYP	4.6	6.5	3.5	29.9	0.9	1.5	0.6	5.0
LC-ωPBE	5.2	7.0	3.9	24.6	1.1	1.7	0.6	9.3
B97-1	5.3	7.1	4.0	22.8	0.8	1.1	0.4	6.6
BLYP	6.5	8.8	4.7	27.3	1.3	2.1	0.6	9.1
PBE0	6.7	9.5	4.2	38.6	1.0	1.5	0.6	6.3
PW86PBE	9.4	12.6	6.7	38.6	1.7	2.6	1.0	11.5
PBE	20.8	25.8	17.5	82.0	3.0	3.8	2.4	13.7
BH&HLYP	29.2	33.5	28.6	79.7	4.5	5.7	3.4	20.4

$$\frac{1}{m}\sum_{n=1}^{m} nx = \frac{m+1}{2}x \tag{5}$$

and this diverges as $m \to \infty$. Thus, the atomization energy errors increase steadily with molecular size and, as larger systems are added to the benchmark set, the errors increase and the mean is inflated. Conversely, the atomization energy per atom is well defined and approaches x when $m \to \infty$. The divergence of the MAE with increasing system size can be seen by comparing the error distributions for the G1, G2, and G3 test sets, as shown in Fig. 4. The G1 set was the first set of atomization energy data, compiled for small molecules only. This set was later expanded to the G2, and subsequently the G3, by adding progressively larger organic molecules to the benchmark. From the figure, the errors in total atomization energies increase going from the G1 to the G3 set. On the other hand, the errors per atom remain roughly constant or even decrease slightly going from the G1 to the G3 set. The errors are reduced in some cases since the DFAs tend to perform better for organic molecules, which constitute a larger fraction of the G3 set. Ultimately, the MAE per atom should be a favored statistic over the total MAE when comparing performance of DFAs for atomization energies.

Another consequence of the breadth of the error distribution is that the MAE may have a large associated uncertainty, particularly for small sample sizes. Therefore, the variance of the mean may be larger than the difference in MAEs between two or more functionals, prohibiting use of this metric to make an informed ranking of DFA quality. For the G3 set, MAEs and their uncertainties for each functional are collected in Table 3. The uncertainties are fairly large, ranging from 0.3 to 1.1 kcal/mol, illustrating that the MAEs are not certain beyond the first decimal point. Additionally, we cannot definitively conclude that B3LYP is the optimum functional, despite giving the lowest MAE, because the difference between the MAEs from B3LYP and CAM-B3LYP is smaller than the sum of the uncertainties.

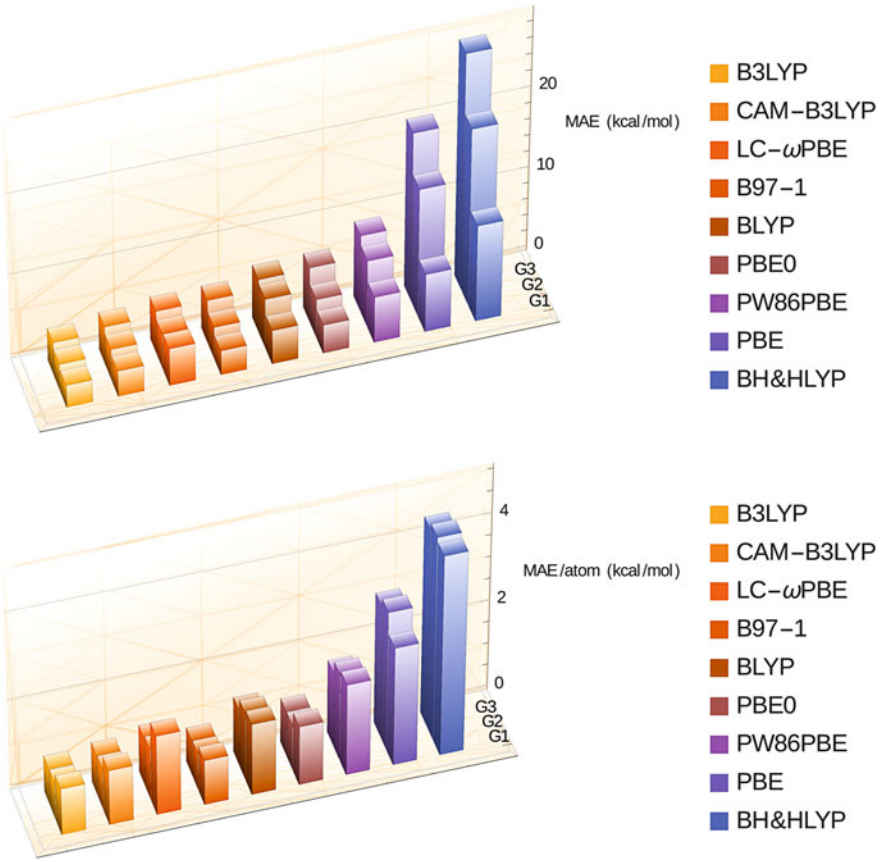


Fig. 4 Histograms of the MAE and MAE per atom with selected functionals for the G1, G2, and G3 sets of atomization energies

Table 3 MAEs and their uncertainties, measured using the square-root of the sample variance, for the full G3 set and three randomly-chosen subsets, in kcal/mol

Functional	G3	Subset 1	Subset 2	Subset 3
B3LYP	4.0 ± 0.4	2.7 ± 0.5	2.4 ± 0.4	3.6 ± 0.8
CAM-B3LYP	4.6 ± 0.3	3.5 ± 0.5	4.0 ± 0.7	4.8 ± 0.8
LC-ωPBE	5.2 ± 0.3	4.6 ± 0.7	5.1 ± 0.7	4.6 ± 0.8
B97-1	5.3 ± 0.3	4.6 ± 0.8	4.4 ± 0.8	5.0 ± 1.0
BLYP	6.5 ± 0.4	5.1 ± 1.1	5.5 ± 1.2	6.3 ± 1.1
PBE0	6.7 ± 0.5	6.6 ± 1.9	4.8 ± 0.8	5.6 ± 1.0
PW86PBE	9.4 ± 0.6	8.7 ± 2.1	7.2 ± 1.4	7.1 ± 1.2
PBE	20.8 ± 1.0	20.6 ± 4.2	20.2 ± 3.8	20.7 ± 2.0
BH&HLYP	29.2 ± 1.1	26.1 ± 4.3	27.4 ± 2.6	30.3 ± 2.2

In general, because of the finite number of systems considered, the mean of a data set is uncertain and has a variance which is inversely dependent on the sample size. To investigate the dependence of the statistics on sample size, we consider the effect of taking three randomly-chosen subsets, each consisting of 22 molecules, from the G3 and evaluating the MAE and its variance for each subset. This procedure demonstrates the danger of using small data sets when benchmarking functionals. For example, using subset 2, the errors for CAM-B3LYP, B97-1, and PBE0 are all equivalent, to within the uncertainty. However, for the full G3 set, CAM-B3LYP is more accurate to within one standard deviation, with an MAE 2 kcal/mol lower than that of PBE0. Additionally, the MAEs obtained with B3LYP for each of the subsets are much lower than the MAE for the full G3 set, showing that it is quite probable that larger errors appear for increasingly larger sets. Indeed, the means for the full set are often worse, because some molecules for which the errors are particularly large are not included in the subsets.

We now turn our focus away from atomization energies and towards intermolecular interactions. Since small sets can give misleading results, are these data sets large enough for us to trust our conclusions? The MAEs for the KB49, S22, S66, and HSG data sets, obtained with the various DFAs, are shown in Table 4, together with their uncertainties, calculated using the square-root of the sample variance. We also consider a superset of 115 intermolecular complexes by combining the KB49 (which already includes the S22 set) with the S66 data set (which is entirely separate from the KB49 set). The MAEs and uncertainties are also shown graphically in Fig. 5. Because of the small size of the benchmarks, particularly for the S22 and HSG sets, definitive statements about the relative quality of the DFAs cannot be made because the uncertainties are so large that the MAEs for many of the functionals are not distinguishable. The statistics for the combined S115 set show, more definitively than for any of the smaller constituent data sets, the separation in performance between the three best-performing functionals (B3LYP, BLYP, and LC-ωPBE) and the rest of the DFAs.

Because of the small values of the interaction energies for dispersion-bound complexes, the mean absolute percent error (MAPE) is often preferred over the MAE. The MAPEs for the intermolecular data sets are shown in Table 5. However, as seen previously for the G3 set, different conclusions regarding which DFA is preferable may be drawn depending on which statistic is used as the selection criterion. The difference between MAE and MAPE is particularly important in sets where the absolute values cover a wide range. For example, considering the HSG set, B3LYP gives the lowest MAE while CAM-B3LYP gives the lowest MAPE. This occurs because, when the binding energies are large in magnitude (typically for H-bonding), CAM-B3LYP generally gives larger errors than B3LYP, while it performs quite well for small binding energies (BEs). The errors for large BEs have less weight in the MAPE than the MAE, so the apparent accuracy of CAM-B3LYP

Table 4 MAEs and their uncertainties for the KB49, S22, S66, and HSG sets, together with the S115 superset (combining KB49 and S66), in kcal/mol

Functional	KB49	S22	S66	HSG	S115
LC-ωPBE	0.27 ± 0.03	0.31 ± 0.06	0.21 ± 0.02	0.23 ± 0.04	0.23 ± 0.02
BLYP	0.29 ± 0.04	0.22 ± 0.04	0.19 ± 0.02	0.20 ± 0.04	0.23 ± 0.02
B3LYP	0.26 ± 0.03	0.31 ± 0.06	0.22 ± 0.02	0.12 ± 0.03	0.24 ± 0.02
PW86PBE	0.39 ± 0.06	0.35 ± 0.09	0.26 ± 0.03	0.17 ± 0.02	0.32 ± 0.03
BH&HLYP	0.35 ± 0.06	0.47 ± 0.11	0.31 ± 0.04	0.18 ± 0.05	0.33 ± 0.04
CAM-B3LYP	0.37 ± 0.06	0.50 ± 0.11	0.35 ± 0.04	0.16 ± 0.05	0.36 ± 0.03
PBE0	0.39 ± 0.06	0.53 ± 0.13	0.36 ± 0.04	0.15 ± 0.03	0.37 ± 0.04
B97-1	0.42 ± 0.09	0.62 ± 0.18	0.38 ± 0.05	0.21 ± 0.05	0.40 ± 0.05
PBE	0.48 ± 0.08	0.57 ± 0.15	0.39 ± 0.04	0.16 ± 0.02	0.43 ± 0.04

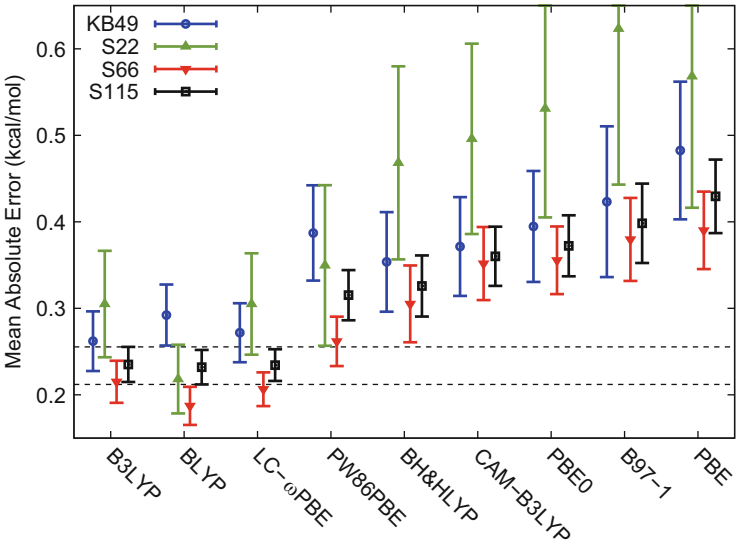


Fig. 5 Mean absolute errors for intermolecular interaction data sets, with the square root of the variance indicated by error bars, for selected DFAs. The numbers in the benchmark names refer the size of each data set

improves. Thus, the choice of statistical indicator determines which would be selected as the best functional and this is an example where the same data can produce opposite interpretations. It is also interesting that B3LYP, BLYP, and LC-ωPBE all give equivalent MAEs for the S115 set, but B3LYP gives a significantly lower MAPE.

As we have demonstrated, using a sufficiently large data set to ensure that the differences between functionals are less than the sum of the uncertainties is of key importance. However, other hidden problems may be encountered when enlarging the test set. For example, a method can perform well for one part of the set, but not

Table 5 MAPEs for the KB49, S22, S66, and HSG sets, together with the S115 superset

Functional	KB49	S22	S66	HSG	S115
B3LYP	6.3	5.0	3.9	10.0	4.9
LC-ωPBE	7.6	5.0	4.3	23.7	5.7
BLYP	9.4	4.8	3.9	9.9	6.2
BH&HLYP	8.1	6.4	5.2	9.5	6.4
CAM-B3LYP	8.2	7.2	6.1	8.4	7.0
PW86PBE	11.3	5.9	6.0	10.5	8.3
PBE0	9.8	8.3	7.4	13.5	8.4
B97-1	11.9	11.7	8.8	14.9	10.1
PBE	13.8	10.5	8.5	10.7	10.8

Table 6 MAPEs for selected methods, for the combined S115 set, divided into two subsets of 35 H-bonded (HB) complexes and 80 other weakly-interacting (WI) complexes

Functional	HB	WI
B97-1	3.0	13.3
B3LYP	3.4	5.6
BLYP	3.4	7.5
LC-ωPBE	3.5	6.7
PW86PBE	3.7	10.3
PBE0	3.9	10.4
PBE	4.2	13.7
BH&HLYP	5.6	6.8
CAM-B3LYP	6.5	7.2

for another. This is the case for intermolecular complexes, where the functionals behave differently depending on whether a dimer is bound by dispersion interactions or hydrogen-bonding. To illustrate this, consider the combined KB49 and S66 sets (called S115 above), which can then be divided into two new subsets: a subset of 35 H-bonded (HB) complexes and a subset of 80 weakly-interacting (WI) complexes. The MAPEs for each subset are given in Table 6 for the selected DFAs. For the WI subset, the lowest MAPE is obtained with B3LYP, and it also gives the second-lowest MAPE for the HB subset, so it performs best for the combined set. However, many of the other functionals do not provide such balanced performance for both interaction types. For the hydrogen-bonding complexes, the best result is obtained with B97-1, although it performs much worse for dispersion-bound complexes. Conversely, CAM-B3LYP and BH&HLYP give the largest MAPEs for H-bonding, but perform much better than B97-1 for dispersion-bound complexes.

When combining data sets, we often wish to determine which functional gives the best balance of errors. As an example of a potential pitfall in such an assessment, let us attempt to judge whether B97-1 or BH&HLYP is most accurate for the union of the HB and WI sets, using ratios of the MAPEs.

On one hand, the defender of B97-1 makes the following argument. True, the ratio of MAPEs for the WI subset ($6.8/13.3 \approx 0.51$) is <1 and favors BH&HLYP. However, for the HB subset, the ratio is >1 ($5.6/3.0 \approx 1.87$) and BH&HLYP is

worse for this set. Thus, on average, we get $(1.87 + 0.51)/2 \approx 1.19$ and conclude that the errors of BH&HLYP are larger than those of B97-1 so the latter should be preferred.

On the other hand, the proponent of BH&HLYP makes the following, analogous argument. True, the ratio of MAPEs for the HB subset ($3.0/5.6 \approx 0.54$) is <1 and favors B97-1. However, for the WI subset, the ratio is >1 ($13.3/6.8 \approx 1.96$) and B97-1 is worse for this set. Thus, on average, we get $(1.96 + 0.54)/2 \approx 1.24$ and conclude that the errors of B97-1 are larger than those of BH&HLYP so the latter should be preferred.

This illustrates how the same data can support two different conclusions. To understand how it is possible to reach two contradictory conclusions in this fashion, consider the more general case, where the two functionals to be compared are labeled methods 1 and 2 and the data sets are A and B. We denote the average error (MAPE in this case) for functional 1 on set A as $\overline{A_1}$ and follow an analogous notation for the other functional and data set. To compare the errors of the two functionals we average the ratios for both data sets: $\overline{A_1}/\overline{A_2}$ and $\overline{B_1}/\overline{B_2}$. If the result is <1, then functional 1 is judged to perform better. However, this analysis can also be performed using the *inverse* ratios: $\overline{A_2}/\overline{A_1}$ and $\overline{B_2}/\overline{B_1}$. The result depends on which ratio was used because of the different order of operations:

$$\frac{1}{2}\left(\frac{\overline{A}_1}{\overline{A}_2} + \frac{\overline{B}_1}{\overline{B}_2}\right) \neq \frac{1}{2}\left(\frac{\overline{A}_2}{\overline{A}_1} + \frac{\overline{B}_2}{\overline{B}_1}\right) \tag{6}$$

Thus, we see that the way the mean has been produced (arithmetic vs harmonic) can sometimes lead to different conclusions. In particular, using the harmonic mean can favor a method that gives near zero error for one data set.

Finally, we consider the probability that all the results published in an article are significant. As we have seen, based on statistics for the G3 set, density-functional approximations work very well for general thermochemistry. However, there is always the possibility of obtaining an error larger than one wants to accept for useful chemical predictions and this risk increases as more results are generated.

For example, assume that the distribution of the errors for the G3 is representative for the chemical systems under study and that the required accuracy is 0.5 kcal/mol per atom. Then the number of cases where this accuracy is reached in the G3 data set, divided by the size of the set, gives the probability that a single additional calculation yields an accurate result. The highest probability to obtain the desired accuracy is provided by B3LYP, but it is only 0.64. Thus, if ten new B3LYP calculations are performed, for systems having the same error distribution as the G3 data set, the probability that all results possess the needed accuracy is extremely small ($0.64^{10} = 0.01$).

One can also impose less strict requirements to judge a result as accurate. Figure 6 shows how the probability of obtaining one, two, five, or ten accurate results for the G3 set evolves when gradually increasing the acceptance bar. For

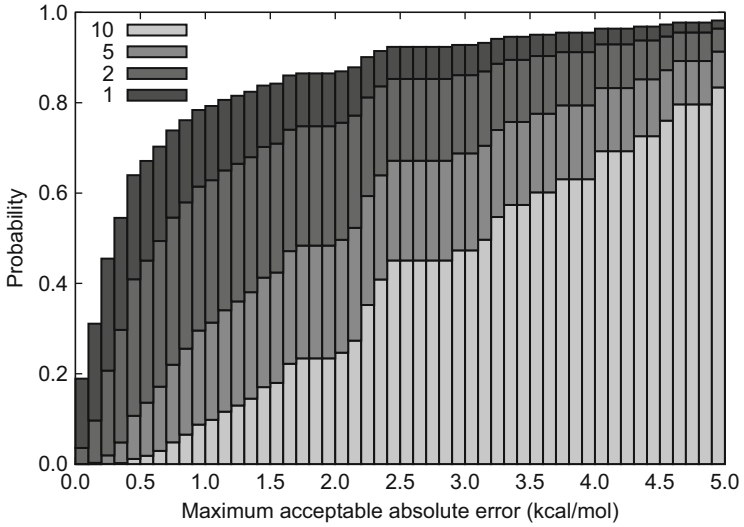


Fig. 6 The probability to obtain B3LYP atomization energies with absolute errors per atom less than a chosen maximum acceptable value, for a set of n systems and assuming the same error distribution as the G3 data set

Table 7 Probabilities of a single calculated result having an error less than a particular accuracy threshold (in kcal/mol) for either the G3 or KB49 data sets, with selected functionals. The values for the G3 set refer to errors per atom

	G3		KB49
Functional	<0.5	<1	<0.5
B3LYP	0.64	0.78	0.86
CAM-B3LYP	0.57	0.77	0.73
B97-1	0.43	0.80	0.78
LC-ωPBE	0.45	0.68	0.84
BLYP	0.41	0.68	0.80
PBE0	0.38	0.65	0.73
PW86PBE	0.29	0.48	0.67
PBE	0.05	0.15	0.63
BH&HLYP	0.03	0.08	0.71

example, if one wants to accept errors within ± 2 kcal/mol per atom, the probability of reaching the desired accuracy for ten calculations is now 0.23.

Similar effects are observed for other data sets and functionals, as shown in Table 7. For example, if we repeat this analysis for the KB49 set, the probability of obtaining a required accuracy of 0.5 kcal/mol from a single B3LYP calculation is very high at 0.86. However, the probability of having ten sufficiently accurate results decreases to 0.22, and to 0.05 for 20. Even if the probability of obtaining a reliable result is high, the probability that all future calculations are accurate becomes low as the number of published results increases.

4 Summary

There is no doubt that density-functional approximations have enhanced the field of computational chemistry. This would not have been the case without the ability of DFAs to produce interesting and reliable data. As the use of statistical measures to assess the quality of DFAs is valuable and necessary, the number of benchmarking studies has been growing rapidly. It should be pointed out, however, that such statistics-based judgments are subject to several potential pitfalls. For atomization energies, commonly used in parameterization of new functionals, the mean absolute error is not well defined in the limit of large molecular size and errors per atom are a preferable statistic. Large data sets are critical for ranking of functionals to minimize the variance, although they can include a variety of effects and can thus blur one's judgment of the functionals. It is even possible to reach opposite conclusions using the same data, for example depending on the choice of statistical measure. Finally, publishing more data naturally augments the risk of including some data with unsatisfactory accuracy. Statistical data used to judge the quality of density-functional approximations must be carefully analyzed and understood in advance of drawing conclusions.

Acknowledgements The authors would like to thank Prof. J. Percus and Prof. O. Percus (both NYU, USA), Dr. Alberto Otero-da-la-Roza (NINT, Canada), and Stephen Dale (UC Merced) for stimulating discussions.

References

1. Civalleri B, Presti D, Dovesi R, Savin A (2012) On choosing the best density functional approximation, specialist periodical reports. Chem Model Appl Theory 9:168–185
2. Hao P, Sun J, Xiao B, Ruzsinszky A, Csonka G, Tao J, Glindmeyer S, Perdew JP (2013) J Chem Theory Comput 9:355
3. Otero-de-la Roza A, Johnson ER (2013) Non-covalent interactions and thermochemistry using XDM-corrected hybrid and range-separated hybrid density functionals. J Chem Phys 138:204109
4. Kannemann FO, Becke AD (2010) van der Waals interactions in density-functional theory: intermolecular complexes. J Chem Theory Comput 6:1081–1088
5. Jurečka P, Šponer J, Černý J, Hobza P (2006) Benchmark database of accurate (MP2 and CCSD (T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. Phys Chem Chem Phys 8:1985–1993
6. Marshall MS, Burns LA, Sherrill CD (2011) Basis set convergence of the coupled-cluster correction, $\Delta_{MP2}^{CCSD(T)}$: best practices for benchmarking noncovalent interactions and the attendant revision of the S22, NBC10, HBC6, and HSG databases. J Chem Phys 135:194102
7. Rezac J, Riley KE, Hobza P (2011) S66: a well-balanced database of benchmark interaction energies relevant to biomolecular structures. J Chem Theory Comput 7:2427–2438
8. DiLabio GA, Johnson ER, Otero-de-la Roza A (2013) An evaluation of the performance of conventional and dispersion-corrected density-functional theory methods for hydrogen bonding interaction energies. Phys Chem Chem Phys 15:12821

9. Faver JC, Benson ML, He X, Roberts BP, Wang B, Marshall MS, Kennedy MR, Sherrill CD, Merz KM (2011) Formal estimation of errors in computed absolute interaction energies of protein-ligand complexes. J Chem Theory Comput 7:790–797
10. Curtiss LA, Jones C, Trucks GW, Raghavachari K, Pople JA (1990) Gaussian-1 theory of molecular energies for second-row compounds. J Chem Phys 93:2537–2545
11. Curtiss LA, Raghavachari K, Redfern PC, Pople JA (1997) Assessment of Gaussian-2 and density functional theories for the computation of enthalpies of formation. J Chem Phys 106:1063–1079
12. Curtiss LA, Raghavachari K, Redfern PC, Pople JA (2000) Assessment of Gaussian-3 and density functional theories for a larger experimental test set. J Chem Phys 112:7374–7383
13. Becke AD (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. Phys Rev A 38:3098–3100
14. Lee C, Yang W, Parr RG (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys Rev B 37:785–789
15. Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. J Chem Phys 98:5648–5652
16. Becke A (1993) A new mixing of Hartree–Fock and local density-functional theories. J Chem Phys 98:1372
17. Hamprecht F, Cohen A, Tozer D, Handy N (1998) Development and assessment of new exchange-correlation functionals. J Chem Phys 109:6264
18. Yanai T, Tew DP, Handy NC (2004) A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP). Chem Phys Lett 393:51–57
19. Vydrov OA, Scuseria GE (2006) Assessment of a long-range corrected hybrid functional. J Chem Phys 125:234109
20. Vydrov OA, Heyd J, Krukau AV, Scuseria GE (2006) Importance of shortrange versus long-range Hartree-Fock exchange for the performance of hybrid density functionals. J Chem Phys 125:074106
21. Perdew J, Burke K, Ernzerhof M (1996) Generalized gradient approximation made simple. Phys Rev Lett 77:3865–3868
22. Adamo C, Barone V (1999) Toward reliable density functional methods without adjustable parameters: the PBE0 model. J Chem Phys 110:6158–6170
23. Perdew JP, Wang Y (1986) Accurate and simple density functional for the electronic exchange energy: generalized gradient approximation. Phys Rev B 33:8800
24. Becke AD, Johnson ER (2007) Exchange-hole dipole moment and the dispersion interaction revisited. J Chem Phys 127:154108

Top Curr Chem (2015) 365: 97–144
DOI: 10.1007/128_2014_557

Published online: 5 December 2014

The Ring and Exchange-Ring Approximations Based on Kohn–Sham Reference States

Andreas Heßelmann

Abstract The ring or random-phase approximation (RPA) method combined with Kohn–Sham reference states has become established as an alternative method to common ab initio wave function methods for the description of the electronic structure of molecules and solids. The reason for this lies in the fact that the RPA possesses, in contrast to, for example, configuration interaction or coupled-cluster methods, a favourable scaling behaviour of $N^4 \cdot \log(N)$ with the system size and describes a number of thermodynamic and electronic properties with a higher accuracy than standard density-functional theory methods. Moreover, the RPA method is able to describe not only dynamic but also strong static electron correlation effects, in contrast to conventional single-reference methods. The latter also include large systems with a small or vanishing band gap. In this work, the performance of the RPA and some extensions to the RPA, including exchange correlations, are tested for the description of thermochemical properties.

Keywords Random-phase approximation, Exchange interactions, Feynman diagrams, Correlation energy, Brueckner orbitals, Atomisation energies, Reaction energies, Reaction barriers

Contents

A. Heßelmann (✉)
Lehrstuhl für Theoretische Chemie, Universität Erlangen-Nürnberg, Egerlandstr. 3,
D-91054 Erlangen, Germany
e-mail: andreas.hesselmann@fau.de

1 Introduction

Consider an infinitely large many-electron system in which the electrons are not confined in an external potential but in which a positive background charge serves for a neutral charge of the whole system (physicists term this model 'homogeneous electron gas'). Without the restriction of generality, one can further assume that the system is one dimensional, that is, all electrons are constrained on a single line. There are then three cases which would have to be considered in order to describe the motion of the electrons, displayed in Fig. 1:

1. Low density limit: the electrons are sufficiently far apart from each other so that the movement of a single electron within the system is almost free, because the interactions between the electrons decay with the inverse distance $1/x$ (Coulombic decay). In this case, the whole system can very well be described in terms of a noninteracting many-electron system obeying solely the Pauli exclusion principle. The latter means that the wave function of the system may be represented by a Slater determinant of the individual single-particle states.
2. Mean density case: collisions between the electrons become more frequent and the system can no longer be well described in terms of a noninteracting system as in

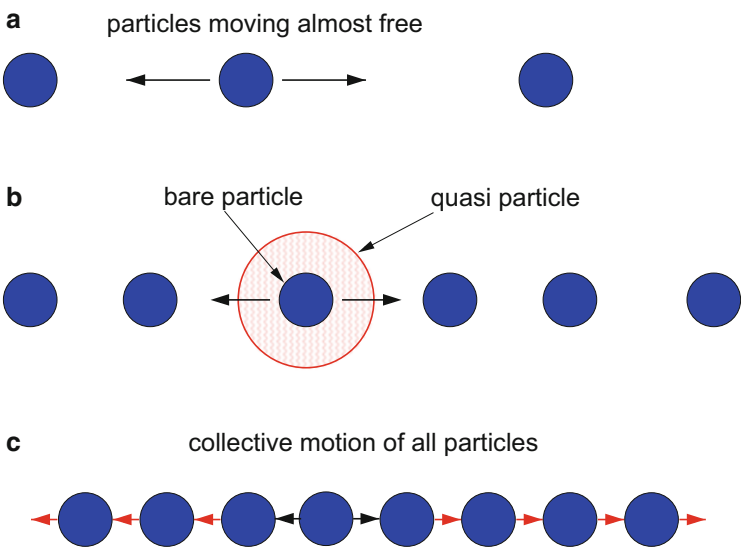


Fig. 1 One-dimensional many-electron system at (**a**) low density, (**b**) mean density and (**c**) high density

case 1. However, the repulsion of the electrons from each other now leads to a charge depletion around each electron, where it is then unlikely to find another electron. Identifying this region of space as so-called holes or regions with an increased positive charge (because of the lack of negative charges around) means that, for each electron within the system, its negative charge is screened because of this positive 'cloud' around it. One can then introduce a new fictitious particle, the quasi particle, which is just the combination of the bare particles (the electrons) with their positive charge cloud: see Fig. 1. In this picture, mapping the bare particle case on a set of quasi particles, again the system can be well described in terms of movements of these quasi particles as their interactions are much reduced because of the screening effects compared to the bare particles. One can then set up another single particle Hamiltonian describing the interactions of the quasi particles by an effective external potential produced by the 'quasi' particle system. All interactions which go beyond this effective single-particle picture can then be accurately described in terms of (low-order) perturbation theory methods. This holds true because of the weak interactions of the quasi particles in comparison to the true interactions of the bare electrons.

3. High density limit: the electrons are now that close to each other that the above quasi-particle model breaks down and the system can no longer be well described solely in terms of a superposition of single-particle states. Indeed, the random oscillations of the electrons now also affect the motion of all other electrons, as is shown in Fig. 1. This means that, in order to describe this situation properly, one has to account for the collective motion of all electrons. It goes without saying that a (truncated) perturbation theory approach is unlikely to be successful in this case because of the strong coupling of the motions of the electrons. Interestingly, analogous to case 2, the collective motions of the electrons can also be associated with a quasi-particle termed the plasmon and the behaviour of electrons in the high-density case is referred to as plasmon oscillation.

For chemists, the mean density case 2 is certainly the most interesting as this is most akin to a molecular system. In the following, however, we now look more in detail at case 3, the high-density limit. Experimental evidence for the mentioned plasma waves in such a system was found when a thin metal foil (being a real representative system for the homogeneous electron gas) is bombarded with high energy electrons; see, e.g. [1]. In a theoretical study of this behaviour, Bohm and Pines have set up the equations for the time behaviour of these density fluctuations [2], leading to two distinct terms: a term which arises from the random thermal motions of the electrons which also appear in a noninteracting system, and a second term which describes the collective motions of the electrons caused by their interactions with each other. This second term was split up into two further terms by Bohm and Pines, yielding one term which describes the oscillations all in phase with each other, and a second term which describes the coupling of oscillations in different modes [2]. Since this coupling term depends explicitly on the particle coordinates, distributed nearly in random positions, it can be argued that their average contribution tends to zero. Furthermore, it was shown, with the aid of numerical arguments by Bohm and Pines, that these anharmonic coupling

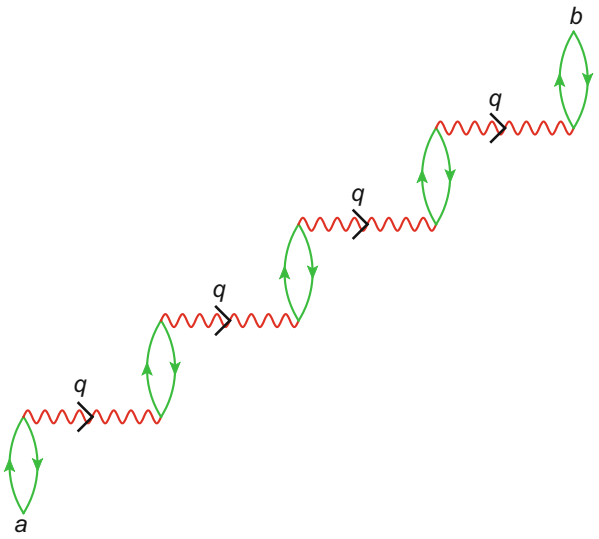


Fig. 2 Polarisation propagator in the random-phase approximation

contributions are negligible in typical electron gas systems with a number of about 10^{12} electrons per cm^3 [2]. The omission of the 'out of phase' terms in the equations was defined as 'random-phase approximation' (RPA) in [2].

Using the model system of Fig. 1c, the RPA can be interpreted as follows. Consider the linear many-electron model as a system of coupled oscillators. In the high-density case, if an electron is displaced from its position, it is likely that it transfers its momentum completely to its neighbouring electron. In total, this leads to a motion of the whole system with all oscillations in phase, and any 'out of phase' oscillations are small in the limit of a highly dense electronic system. Clearly this would be different in the case in Fig. 1b where the transfer of the momentum of one electron to another is largely reduced because of the lower density within the system.

The propagation of the density fluctuation of the system can be represented graphically by the diagram in Fig. 2 [3]. In this figure a density fluctuation arising in point a is transferred to point b through a series of electron–electron collisions represented by the red wiggly lines. The upward pointing arrows in the figure represent the movement of the electron which is excited from its single-particle ground-state while the accompanying downward pointing arrows represent a corresponding unoccupied hole state. Thus the two-arrow units in Fig. 2 can be interpreted as an electron which leaves its mean position and oscillates back in a cycle. The separation of the particle (up) and hole (down) arrows (so-called particle-hole pair) therefore mark a polarisation within the system, i.e. a separation of negative and positive charges.

The propagator shown in Fig. 2 represents the most important contributions in the high-density case, i.e. in the diagram in Fig. 2 each particle–hole pair transfers its momentum (marked by q in the diagram) completely to another particle–hole pair in concordance with the above considerations. A more unlikely situation in the

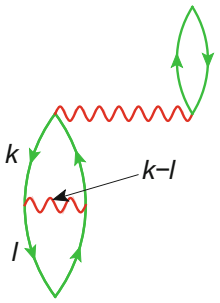


Fig. 3 Polarisation propagator with a negligible contribution in the high-density case

high-density case is shown in Fig. 3, since here an electron transfers only a fraction of its momentum $(k - l)$ to another electron. Such a case is omitted within the RPA.

After the RPA was introduced by Bohm and Pines [2, 4, 5], the theoretical predictions of the ground-state energy and the quasiparticle properties of the homogeneous electron gas led to a considerable improvement over the commonly used Hartree–Fock results [6]. However, its validity was questioned by a number of people because the collective modes were treated on a quite different basis compared to quasiparticle excitations. Soon after the works of Bohm and Pines, it was found that a perturbational technique by Gell-Mann and Brueckner [7] which was used by them to calculate the ground state energy and specific heat of the homogeneous electron gas turned out to be identical to the RPA [8]. In this method, the ground-state correlation energy is approximated by summing over all most divergent terms in each order in the electron–electron interaction (the divergency originating from the long-range character of the Coulomb interaction). Interestingly, while each term then in itself diverges, the total sum of all these terms leads to a finite number which is identical to the exact energy of the homogeneous electron gas at high density [7]. Independently, Hubbard also established a connection between the collective coordinate approach and the perturbation-theoretic summation using a field-theoretic method [9, 10].

It should be noted that this perturbative technique, the summation over the most divergent terms in each order, is closely connected to the form of the propagator shown in Fig. 2. This is because each red wiggly photon exchange line in the diagram produces a term that contributes a factor of $1/q^2$ in the propagator expression (cf. [3]). Therefore, the larger the total momentum transfer of a given propagation process, the larger its degree of divergence for $q \to 0$. Collecting all diagrams contributing to the correlation energy with the highest degree of divergence means that only fragments of the type of the one shown in Fig. 2 but not of the type of the one shown in Fig. 3 have to be summed. The resulting diagrams, as shown in Sect. 2, all possess a ring structure. This is why the RPA is sometimes also referred to as *ring-approximation*.

The question which arises now is why the RPA has also gained an increasing popularity for describing the electronic structure in molecular systems, cf. [11–15]? As mentioned above, apart from the fact that a molecular system is finite in size, in

contrast to the homogeneous electron gas model, it certainly is much better represented by the mean density case in Fig. 1b rather than the high-density case in Fig. 1c. A first point which should be noted and which was already taken into account by Hubbard is the fact that the RPA is not accurate anymore, even in the electron gas case for large momentum transfers [10, 16]. The reason is that in the RPA there is no distinction between the correlations of electrons with a parallel and antiparallel spin, and thus the method does not take into account the Pauli principle (in second and higher orders). Clearly, a remedy for this problem is obtained by also including exchange contributions in the correlation energy. In the high-density electron gas the only nonvanishing contribution to the exchange correlations is the second-order one that affects the total correlation energy by more than 10% [17]. In chemistry, the first applications of exchange-RPA methods were carried out by McLachlan and Ball, who studied both the ground-state correlations as well as the excited states of the π electrons of a double bond [18, 19]. At that time, thanks to the work of Ehrenreich and Cohen, it was already clear that the RPA with (a complete account of) exchange effects is identical to the time-dependent generalisation of Hartree–Fock theory (TDHF) [20]. Since then, the TDHF method has been used much more often to study excited states or response properties of chemical systems rather than ground-state correlations [21–24]. Two significant contributions on the application of exchange-RPA methods to describe the ground state are those by Szabo and Ostlund [25, 26] and Oddershede [24]. In the exchange-RPA method by Szabo and Ostlund, the RPA energy expression is modified such that the long-range correlation energy between two closed-shell molecules resembles the dispersion interaction energy in the TDHF formalism [25–27], in contrast to the method by McLachlan and Ball [18, 19]. The RPA method by Oddershede takes into account that the TDHF method is often affected by triplet instabilities caused by an incompatibility of the wave function ansatz and the excitation operators [24] and expresses the ground-state correlation energy solely in terms of the singlet excitations. Only recently, however, it was observed that these exchange-RPA variants based on Hartree–Fock reference states yield rather unsatisfactory results for chemical reaction energies or intermolecular interaction energies if compared to low-order Møller–Plesset perturbation theory or coupled-cluster results [28].

To that end, it should be noted that yet another approach for calculating the (exchange-) correlation energy of metallic systems was developed in the 1970s by Langreth and Perdew, based on the fluctuation-dissipation theorem [29] and termed the adiabatic-connection fluctuation dissipation theorem (AC-FDT) [30, 31]. While the fuctuation-dissipation theorem itself relates the internal random fluctuations of a many-particle system to the response the system undergoes caused by a small applied external perturbation, the AC-FDT method generalises this theorem to enable the calculation of the correlation energy of the system. This is achieved by connecting the noninteracting and interacting many-electron systems in such way that their density along the connection path is identical [30, 31]. Surprisingly, it can be shown that this approach and the above sketched perturbative approach are identical within the RPA [32, 33]. From that perspective it is, however, clear that

using the Hartree–Fock reference wave function to represent the noninteracting system might not be the optimal choice for use within the AC-FDT method. The reason for this is that, while the Hartree–Fock determinant minimises the energy expectation value, its density differs from that of the correlated many-electron system. A much more natural choice therefore would be to use a reference determinant in conjunction with the AC-FDT method, which yields the exact density, i.e. the Kohn–Sham determinant.

Recent works applying Kohn–Sham-based RPA (KS-RPA) methods to molecular systems, starting with the initial work by Furche [11], have indeed shown that KS-RPA not only describes dynamic electron correlation effects with a reasonable accuracy, important, for example, for the description of non-covalent interactions of molecules [14, 34] or molecular crystals [35], but also that this method is able to describe static correlation effects qualitatively correctly [11, 14, 36–38]. Such effects become important in systems with a small HOMO-LUMO gap, generally for transition metal complexes, metallic systems, or for the description of bond dissociations of single- and multiple-bonds. These cases can usually only be described by computationally expensive multi-reference methods; standard single-reference methods (including standard density-functional theory (DFT) methods) usually fail severely for systems with a small HOMO-LUMO gap. Certainly this flexibility of the KS-RPA method in the description of various kinds of correlation interactions is one of the reasons for its increased popularity. It has to be noted, however, that the KS-RPA method fails dramatically in describing the bond dissociation of systems with an uneven number of electrons, e.g. in case of the H_2^+ molecule [14, 37, 39]. The reason again lies in the fact that the RPA does not account for exchange interactions, and therefore strongly overestimates electron correlation energies caused by a self-correlation error (see also Sect. 2). The most striking example for this error is a single-electron system for which the RPA correlation energy is unequal to zero [36, 40]. As a consequence of this, it was observed that the RPA also systematically overestimates the strength of ionic and covalent bonds [14, 15].

A remedy for this problem is again obtained by including exchange contributions within the RPA correlation energy. As has already been mentioned above in the context of Hartree–Fock-based RPA methods, this can be done in many various ways, and therefore a number of different exchange KS-RPA methods have been developed in recent years [41–46]; see [12, 13] for an overview. While these methods certainly are computationally more expensive than the corresponding direct RPA methods, they often yield a higher accuracy, e.g. for intermolecular interaction energies [15, 45], chemical reaction energies [45] or atomisation energies [15]. Furthermore, certain variants, such as the second-order screened exchange (SOSEX) method [43] or the RPAX2 method [45], still possess a moderate scaling behaviour of N^5 with the molecular size N, which is only slightly larger than with RPA (having a scaling behaviour of $N^4 \cdot \log(N)$ [47]). An interesting alternative approach for treating exchange effects indirectly within the RPA is the RXH method proposed by Gould which uses a model of the exchange-hole

[46]. This method possesses the same scaling behaviour as RPA, although it gives improved total correlation energies and C_6 atom–atom dispersion coefficients for a number of atoms [46]. Whether this also holds true for molecular systems, however, has yet to be proven.

An important aspect of using RPA methods for larger molecular systems in favour of, e.g. more popular standard DFT methods is its scaling behaviour with the molecular size. As was mentioned above, the scaling behaviour of the RPA method with the molecular size N is relatively favourable in comparison to standard perturbation-theory or coupled-cluster methods, namely it increases by $N^4 \cdot \log(N)$. Currently there exist a number of efficient implementations of the RPA method in various computer programs for molecules [14, 45] and solids [15, 48, 116]. For example, the resolution-of-identity implementation of the RPA by Eshuis et al. allows the computation of a system of 120 electrons (tetracene) within a quarter of an hour on a 2.80-GHz machine and a moderate memory requirement of 900 MB [47]. This is good enough to also be able to perform geometry optimisations for large molecular systems.

Apart from the scaling behaviour of the RPA method with respect to the size of the molecule, however, the convergence of the RPA correlation energy with respect to the number of basis functions used to expand the molecular orbitals also has to be taken into account. Since in RPA methods, similar to wave-function methods, the interelectronic cusp of the pair correlation function has to be modelled explicitly [40, 49] by a finite basis set, the basis set convergence is de facto identical in both cases and can be described by an inverse cubic function of the cardinal number of the underlying basis set [11, 50, 51]. This means that generally large basis sets [34] or basis set extrapolation techniques [45, 50, 51] have to be used in conjunction with the RPA, otherwise the results are more or less affected by a basis set error.

A different way of tackling this problem is to use so-called range-separated methods [52]. In these the Coulomb operator is split up into a short- and long-range part with the idea that then only the long-range interactions are described on the RPA (or generally any other wave-function theory) level while the short-range interactions are described by standard DFT methods instead [42, 53, 54]. In this way a large fraction of the dynamic electron correlation effects are described on the conventional DFT level, leading to a much faster basis set convergence compared to the RPA [42]. Different variants of these range-separated methods describe intermolecular interactions [42, 54, 55] and atomisation energies [37] with good accuracy.

The above overview could hopefully shed some light on the origin of the RPA method, developed primarily in order to describe the homogeneous electron gas, and also explain its recent success in the molecular world. This work continues with a derivation of the RPA correlation energy using Feynman diagrams (Sect. 2). It should be noted that the purpose is to sketch briefly how diagram expansions of wave functions or correlation energies can be set up, rather than to give a thorough and detailed discussion of the underlying concepts of drawing Feynman diagrams. The presentation is therefore incomplete (if not inaccurate) and the reader is referred to a number of good reference texts on the subject; see [3, 17, 56–59].

Section 3 explains why Kohn–Sham RPA methods are favoured compared to Hartree–Fock-based RPA methods for describing electron correlation effects. In Sect. 4 it is analysed how accurate Kohn–Sham RPA methods are in describing thermochemical properties, namely atomisation energies, ionisation energies, electron and proton affinities (Sect. 4.1), chemical reaction energies (Sect. 4.2) and reaction barrier heights (Sect. 4.3). Section 5 gives a brief summary.

2 Derivation of the RPA Correlation Energy Using Feynman Diagrams

In order to describe a quantum many-body system, a sum over all possible interactions within the system has to be made. Practically this is what is different compared to a classical many-body system where no interference properties of the particles need to be accounted for and for which particles follow definite and physically observable paths. Nevertheless, it has become convenient to describe a quantum many-body system in terms of quasi-physical processes which can be represented by graphical cartoons derived by Feynman [60]. Such a process is exemplified in Fig. 4.

In Fig. 4, two electrons which are in the states a and b, marked by the green arrows, move towards the points X and Y. They then interact with each other as marked by a red wiggly photon-exchange line in the diagram, and thereby scatter into two new states c and d (note that the exchanged photon is a virtual one, so the uncertainty principle is not violated). As in a classical scattering process, the initial and final momenta of the electrons have to be such that the total momentum is conserved. While the Feynman diagram in Fig. 4 now appears to describe an actual physical process, its interpretation differs from how one usually calculates the time propagation of particles in the classical world. Namely, rather than describing directly the probability that if two electrons enter in states a and b and are scattered into the states c and d, the diagram represents the *probability amplitude* for this process. The physical significance of a probability amplitude is that it needs to be squared in order to obtain the (physically observable) probability for the underlying

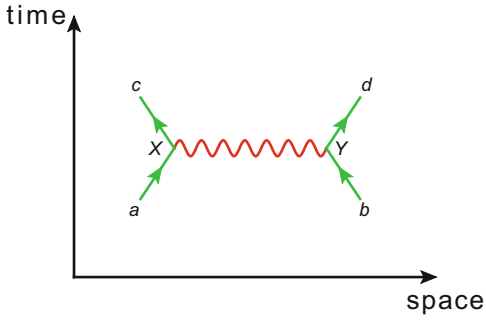


Fig. 4 Interaction of two electrons in a space-time coordinate system

process. In this way, Fig. 4 may be interpreted as a quasi-physical process which can be used to describe true physical observables.

As a matter of fact, however, Fig. 4 illustrates just one possibility of how two electrons can propagate from a, b to c, d. Practically, as is the nature of electrons being quantum particles, they follow all possible paths on their way from the states a, b to c, d. As an example, they could first scatter into two different states e and f and then interact again, being pushed in their final states, which would be a second-order process. Thus, in order to describe accurately a many-electron system, all different possibilities need to be accounted for. For systems already of the size of the water molecule containing ten electrons, as an example, this is not possible anymore in praxis, because the effective calculation of the sum of amplitudes becomes too expensive.

Therefore, in order to describe a many-electron system, approximations have to be made. The aim of quantum chemistry methods derived from perturbation theory methods is to make partial sums over the most important contributions to an amplitude, while neglecting other terms, either because they make only small contributions to the total sum or because they cancel to some extent with other terms not accounted for. One such approximate method, which includes only a subset of all possible interactions within a quantum many-electron system, is the RPA, which is derived here in terms of a diagrammatical expansion.

In order to do this, note that quantum chemical methods usually describe a many-electron system in terms of an amplitude termed the many-electron wave function. The ground state wave function $|\Psi_0\rangle$ itself may be referred to as an unconditional amplitude for finding simultaneously electrons in all possible states i, j,. . ., a, b,. . . below (hole-states) or above (particle-states) the Fermi level (the mixing of the excited states into the noninteracting 'vacuum' arising from the mutual interactions). If the electronic ground state is considered, one calculates the probability to find the system in an unperturbed (non-interacting) ground state Φ_0 (which usually is a single-determinant wave function and referred to as Fermi vacuum) at a time t_1 when it started from the same state Φ_0 at a previous time t_0 and with the external potential and the interactions between the electrons switched on within the time interval $[t_0, t_1]$. The quantity of interest is also termed vacuum amplitude as there are no electrons added or removed from the system. Corresponding Feynman diagrams for the vacuum amplitude possess, in contrast to the diagram shown in Fig. 4, no open lines (see below for some examples of vacuum amplitude diagrams).

Certainly, if the system is in a state Φ_0 at a time t_0 and ends up in this state at a later time t_1, then in between the system may undergo all possible transitions from its noninteracting ground-state. Thus the interacting ground-state many-electron wave function Ψ_0, in addition to the noninteracting ground-state configuration Φ_0, contains a linear combination of all possible excitations as shown schematically in Fig. 5.

The second term on the right-hand side of Fig. 5 represents all singly excited configurations from the reference state Φ_0, the third term all doubly excited configurations, and so on. As an example, note that the doubly excited configuration

Fig. 5 Diagrammatic representation of the many-electron wave function

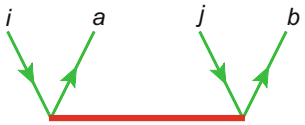


Fig. 6 Two-particle excitation involving hole states i, j and particle states a, b

involves a sum over all amplitude probabilities that two electrons are simultaneously excited from hole states i, j into two particle states a, b; see Fig. 6.

Since all possible double excitations occur simultaneously, a summation over all distinct pairs of hole and particle states in Fig. 5 is implied, and thus the explicit labels shown in Fig. 6 are left out. The same is also true for the other excitation amplitudes of Fig. 5.

In order now to calculate the vacuum amplitude from the many-electron wave function, the system has to go back to its noninteracting ground-state configuration Φ_0 which, in terms of diagrams, means that no open lines may appear above the final vertex of a given diagram. Obviously, the simplest possibility for the probability that the system is in state Φ_0 at a time t_1 is that it remains in this state within the time period $[t_0, t_1]$. This would mean that the system is unperturbed and can be described solely in terms of a single Slater determinant Φ_0. In all other cases, however, the system has to be described by two successive processes: (1) the system is excited from its noninteracting ground state; (2) the system is de-excited into its noninteracting ground state. According to the rules for two successive Feynman paths, the amplitudes for these two steps have to be multiplied. This is exemplified in Fig. 7 for the singly, doubly and triply excited cases. The upper dashed line in the figure represents the corresponding de-excitation and corresponds to the complex conjugate of the wave function. In the case of the double excitation, it stands for the probability amplitude that if two electrons enter in the particle states a, b whether, after interacting with each other, they scatter into two hole states i, j. The blue zigzag lines in Fig. 7 denote possible electron–electron interactions.

From the vacuum amplitude diagrams in Fig. 7 one can directly calculate the ground-state correlation energy of the many-electron system using corresponding rules for evaluating Feynman diagrams [3, 56, 58, 59, 61, 62]. In order to see this, note that the vacuum amplitude diagrams as drawn in Fig. 7 just represent the expectation value of the electron–electron interaction operator $\hat{V}$ given as $\langle \Psi_0 | \hat{V} | \Psi_0 \rangle$ [3, 56]. In the following we are now concerned with the approximate evaluation of the ground-state energy since, as stated above, a summation of all possible vacuum amplitudes is already no longer feasible for small molecular systems.

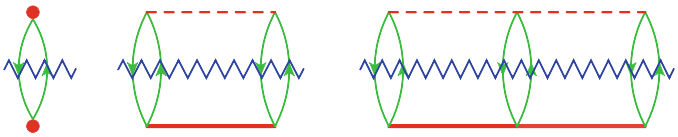


Fig. 7 Examples of vacuum amplitudes. The *bottom* and *upper red fragments* represent the wave function and the *blue zigzag lines* the electron–electron interaction operator

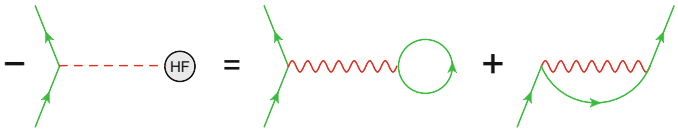


Fig. 8 A diagram containing the Hartree–Fock effective interaction potential cancels with the same diagrams containing a bubble vertex and an open oyster vertex instead

As a first restriction, which is abolished later on, the reference state Φ_0 is assumed to be the Hartree–Fock determinant. With this, the number of diagrams which have to be accounted for in a given order of the electron–electron interactions is significantly reduced because all diagram fragments containing the Hartree–Fock one-electron effective interaction operator cancel with corresponding diagrams which contain a bubble vertex and an open-oyster vertex instead [56]; see Fig. 8.

The interpretation of the diagrams in Fig. 8 is as follows: the diagram on the left-hand side describes an interaction of an electron with the effective Hartree–Fock potential while the two diagrams on the right-hand side describe the actual interaction. In the first one the bubble vertex means that the second electron which interacts with the electron coming on the left-hand side is instantaneously pushed back in its hole state after the interaction. The second diagram on the right-hand side is a first-order exchange diagram and practically describes the same process as in the first diagram, but in which the electrons after the interaction are exchanged, that is, the electron in the particle state coming from the left is pushed in a hole state from which another electron is excited in the same particle state.

It has to be noted that the specification of the noninteracting system does not affect the total sum of vacuum amplitudes as long as *all* contributions are summed in Fig. 7 (also including quadruple, pentuple, etc. contributions not shown in the figure). It *does*, however, affect the sum in an *incomplete* expansion of the vacuum amplitudes. This is an important point in this work and is discussed below when the Hartree–Fock reference state is replaced by a noninteracting Kohn–Sham reference state.

Having specified the reference state, it should be noted in passing that the first-order vacuum amplitudes, i.e. the amplitudes involving only one photon exchange between the electrons, in RPA methods are always calculated as in Hartree–Fock-based RPA methods, i.e. by the sum of the double-bubble (direct Coulomb interaction) and oyster (exchange interaction) diagrams shown in Fig. 9.

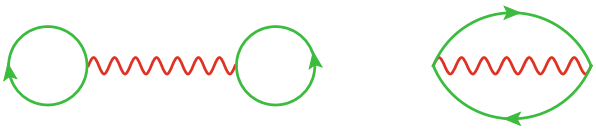


Fig. 9 In RPA methods the first-order vacuum amplitudes are always calculated by the sum of the double-bubble and oyster diagrams shown in the figure, irrespective of the chosen reference state

Actually, the second first-order exchange diagram in Fig. 9 represents an important contribution to the total interaction energy, since it cancels the unphysical Coulomb self-interaction contributions from the first diagram in Fig. 9. This self-interaction contribution arises from the fact that, in unlabeled diagrams, a summation over all hole- and particle-lines is made, and, thus, the double-bubble diagram of Fig. 9 partially describes unphysical interactions of an electron with itself. This contribution, however, is cancelled by the corresponding same-labeled terms arising from the second diagram in Fig. 9, showing the importance of the description of electron exchange in the sum of vacuum amplitudes.

This problem of equal labeled diagrams at the same horizontal position holds true in general, and a corresponding diagram is termed an exclusion-principle violating (EPV) diagram [56, 58, 63]. While such diagrams would cancel completely in an unlinked expansion of the vacuum amplitude (which also means that disconnected diagrams are accounted for in the sum), in praxis it is much more convenient to use the simpler linked expansion (with only connected diagram contributions), using unconstrained summations, and keep the EPV diagrams [56, 58].

The basic first approximation which is used to define the RPA is that the wave function at most contains double excitations from the reference state. This means that the triply and all higher order excitations from the reference state displayed in Fig. 5 are neglected. It has to be noted that this is also the lowest possible level at which the expansion of the wave function may be truncated, because single excitations from the reference state cannot describe electron correlation effects beyond the single-determinant approximation. The reason is that singly excited configurations can be written in terms of unitary transformations of the reference state, and can thus be absorbed in the single-determinant reference state (Thouless theorem [64]). As a second approximation, single-excitations are also omitted in the RPA wave function. This appears to be a good approximation if the Hartree–Fock determinant is used as the reference determinant, since in a Hartree–Fock basis singly-excited vacuum amplitudes at first appear in fourth order [59]. Since below, however, the RPA is combined with Kohn–Sham reference states, this point needs to be justified by a different argument to be given later on. In summary, the RPA wave function can be written by the sum of a noninteracting single-determinant reference state and two-electron excited configurations; see Fig. 10.

By this means, one can now start to derive explicitly the RPA vacuum amplitudes (defining the RPA correlation energy, i.e. omitting the first-order interaction contributions of Fig. 8) in an order-by-order expansion. This can be done by

Fig. 10 Diagrammatic representation of the RPA wave function

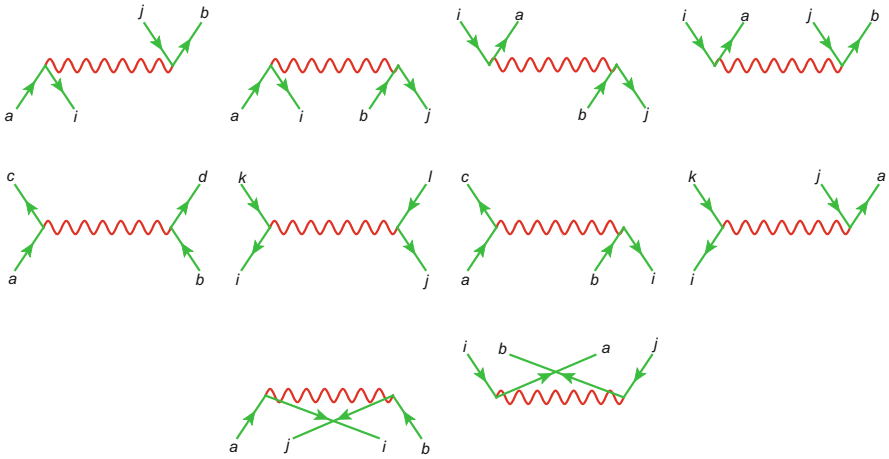


Fig. 11 Possible types of electron–electron interactions. The *first line* displays different particle-hole interactions which are described by the RPA. The *second line* displays particle–particle, hole–hole (not described in any RPA methods) and mixed type interactions which are described by some exchange RPA variants. The *third line* displays two different exchange particle-hole interactions which are accounted for in exchange RPA methods

creating double excitations from the reference state in all possible ways, but using the following confinements which further reduce the number of contributions to the RPA correlation energy:

- In RPA only particle-hole interactions are included in the electron–electron interactions
- In the particular direct RPA method exchange interactions (beyond first order) are omitted

The first line in Fig. 11 shows the different particle-hole interactions which can occur in RPA methods. As a rule of thumb, an electron that interacts with another electron in an RPA description of the many-electron system is either excited from a hole state or de-excited from a particle state. In contrast, the diagrams in the second row in Fig. 11 present processes in which a particle state is transferred to another particle state (particle–particle interaction) or a hole state is transferred to another hole state (hole–hole interaction). The first two types of interactions do not occur in any RPA methods, while the last two types can occur in some exchange RPA methods.

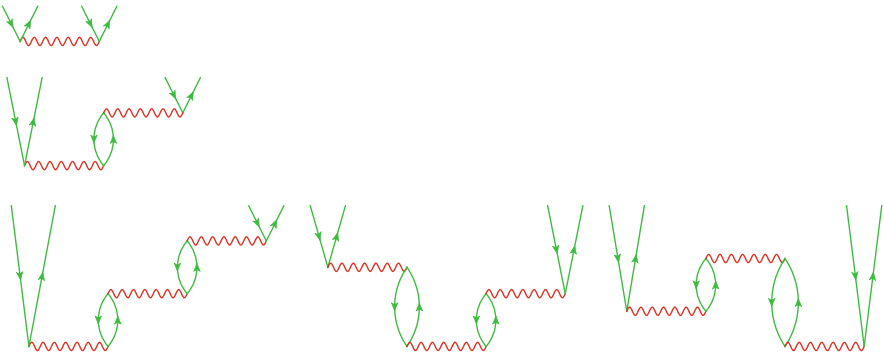


Fig. 12 Double excitation amplitudes occurring in the direct RPA method in first (1st line), second (2nd line) and third (3rd line) order

The generation of RPA vacuum amplitude diagrams now involves the steps to create all possible distinct (connected, see above) double excitations using the diagram fragments of the first line of Fig. 11 and close them by an electron–electron interaction (operator) contribution, so as to leave no open lines in the resulting diagram. This has to be done such that only hole lines are attached to hole lines and particle lines to particle lines. The resulting double excitation amplitudes which can be constructed with the constraints given above are shown in Fig. 12 up to third-order (i.e. involving up to three interaction processes).

As can be seen, in first- and second-order there exists only one distinct way to connect particle–hole interactions to a corresponding double excitation amplitude. A set of distinct diagrams is one in which it is impossible to obtain one diagram from another without cutting any line, i.e. by distorting the diagram in the horizontal direction but leaving the vertical ordering of the interaction lines. Thus, in third order there are three different ways to create a double excitation amplitude; see the third row in Fig. 12. Here the involved particle–hole interactions in the three cases possess different time orderings and the resulting amplitudes cannot be obtained from one another without violating the rule given above. All other possibilities to draw a third order amplitude results, however, in a diagram which is topologically equivalent to one shown in Fig. 12. As an example, the first third-order diagram reflected around the vertical axis is topologically equivalent to the one given in Fig. 12, since it can be obtained by distorting this diagram in the horizontal direction.

The vacuum amplitude diagrams can now be obtained from the diagrams of Fig. 12 by closing each diagram with a particle–hole interaction operator, explicitly, using the second diagram in the first line in Fig. 11. Since no exchange-type interactions are allowed in the direct RPA method, this can be done in only one possible way, so up to fourth order the vacuum amplitudes of Fig. 13 contributing to the RPA correlation interaction energy result.

It goes without saying that the total RPA correlation energy is obtained by an infinite sum over all diagrams of the type shown in Fig. 13. One can observe that

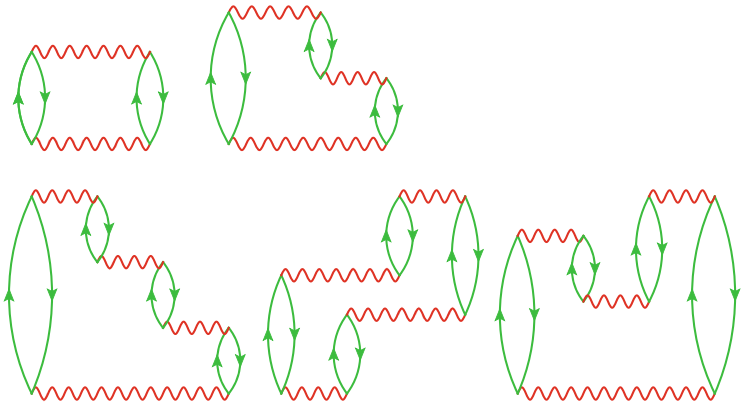


Fig. 13 Vacuum amplitudes of second, third and fourth order contributing to the RPA correlation energy

they all possess a characteristic ring form, because of which the RPA is sometimes also referred to as ring-approximation. Of course, in an actual calculation of the RPA correlation energy it is not possible to calculate the infinite sum of all ring-diagrams. Instead, an iterative algorithm is used to calculate the doubles amplitudes contributions of Fig. 12 up to a given order [32, 45]. Fortunately, the high order ring-diagrams become less and less important in the summation and can then be neglected. This and the fact that actually the order-by-order summation of the ring-diagram correlation energy $E_{\text{corr}}^{\text{rings}} = E^{(2)} + E^{(3)} + E^{(4)} + \ldots$ is an alternating series with $E^{(2)} < 0$, $E^{(3)} > 0$, $E^{(4)} < 0$ and so forth means that already a truncation of this series at a low order is very close to the infinite sum. Note that the prefactor associated with a corresponding vacuum amplitude diagram is given by $-(-1)^{h+l}$ where h corresponds to the number of hole lines and l to the number of closed loops. As an example, the second-order diagram in Fig. 13 contains two hole lines and two closed loops and therefore has a prefactor of -1 (note that this rule needs to be combined with the convention for the energy denominators used in the algebraic expressions derived from the diagrams, so the reader is referred to [56, 58, 61] for a more complete and consistent definition of the diagram sign rule).

As has already been discussed for first-order interactions, the unrestricted summations over orbital states in diagrams leads to a number of unphysical contributions which violate the Pauli exclusion principle (so-called EPV diagrams, see above). As an example, the second-order ring diagram with two identical particle-line labels (see left diagram in Fig. 14) describes an interaction of two electrons which are in the same unoccupied state a. As is shown in the figure, the two particle lines may be exchanged on either the upper or the lower interaction vertex without changing the diagram (including the sign). By this operation one then obtains the diagram on the right-hand side in Fig. 14 which possesses the form of the second-order exchange diagram.

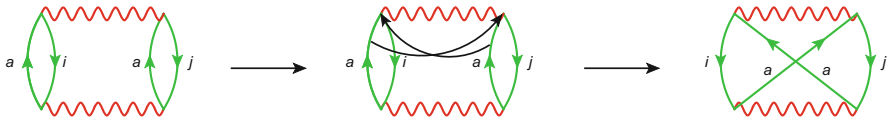


Fig. 14 The exchange of two same labeled particle lines ending at the same interaction line leaves the diagram, including its prefactor, unchanged

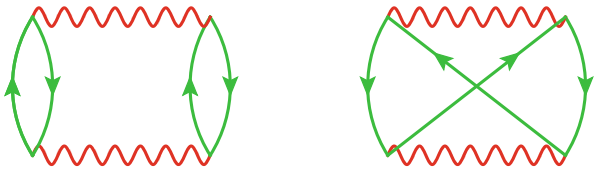


Fig. 15 All second-order vacuum amplitudes (Hartree–Fock basis). The exchange-type diagram on the *right-hand side* cancels the unphysical self-correlation contributions of the second-order ring diagram on the *left-hand side*

The problem is now that in the summation over the ring diagrams in RPA (Fig. 13), many such unphysical contributions occur and are not cancelled as compared to a complete order expansion of the correlation energy. In analogy to the first-order case, where this issue is termed self-Coulomb interaction, the RPA method is stated to include a self-correlation error. Indeed, it has been found that the RPA method usually overestimates electron correlation energies of atoms and molecules significantly because of the uncancelled EPV diagrams [13, 28, 65].

Obviously, since in a complete sum the EPV diagram contributions are eliminated, this must also hold true in each order in the vacuum amplitude expansion. As an example, Fig. 15 displays the two second-order vacuum amplitude diagrams which would occur in a complete perturbation series expansion also containing exchange-type interactions as shown on the bottom of Fig. 11.

As can be seen, the exchange-diagram on the right-hand side in Fig. 15 has the same form as the third diagram in Fig. 14, being associated, however, with a different prefactor being $-(-1)^{2+1} = +1$ (two hole lines, one loop). Therefore, the EPV diagrams arising from the ring-diagram of Fig. 15 is cancelled by corresponding contributions from the second diagram.

While one way to deal with the self-correlation error of RPA would be to correct the RPA correlation energy in low orders by adding corresponding exchange contributions [41], a more effective way is to do this through all orders. As a matter of fact, however, this can be done in many various ways and a number of different exchange-RPA methods have been derived; see [12, 13] for an overview. In this work only two of these methods are presented which can be implemented using efficient algorithms [43, 45] and thus also allow one to calculate the electron correlation energy for extended systems.

The first of these two exchange-RPA methods is the SOSEX method by Kresse et al. [43]. In this method the double excitation amplitudes are calculated just as in the RPA method and are then antisymmetrised. Thus, in first and second order

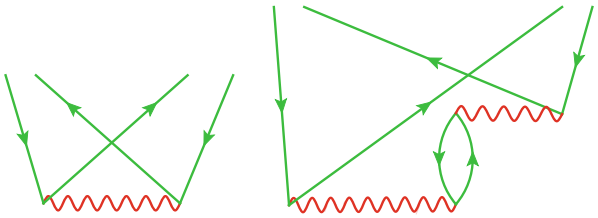


Fig. 16 Antisymmetrised first- and second-order doubles amplitudes described by the SOSEX method

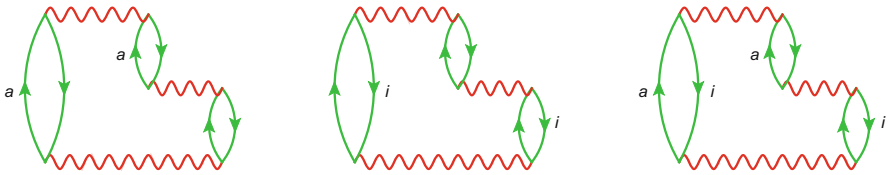


Fig. 17 Labeled third-order EPV diagrams occurring in the RPA vacuum amplitude expansion

(yielding second- and third-order energy contributions), in addition to the two RPA amplitude diagrams shown in Fig. 12, the SOSEX method also contains terms arising from the amplitudes shown in Fig. 16.

One can observe that the antisymmetrised doubles amplitudes can be obtained from the RPA amplitudes just by intersecting the two particle lines of the corresponding RPA diagram. Obviously, by this the SOSEX method contains in second-order, in addition to the ring diagram, the exchange diagram of Fig. 15, and is therefore exact through this order.

The antisymmetrised second-order doubles diagram of the SOSEX method of Fig. 16 cancels all labeled EPV diagram stemming from the third-order ring-diagram, shown in Fig. 17.

In third order, however, there exist further exchange contributions not described by the SOSEX method. These can be obtained by performing the antisymmetrisation of the doubles amplitudes in all orders. A straightforward recipe to realise this, implemented by the RPAX2 method of [45], is to generate the doubles amplitudes as described above for the RPA method, but to add, in each order, a corresponding antisymmetrised amplitude diagram with the two open particle lines exchanged. By this, in first-order the two diagrams of the first row in Fig. 18 are obtained which yield the two second-order vacuum amplitude diagrams of Fig. 15. In second-order the attachment of a (non-antisymmetric) particle–hole interaction diagram (first or third diagram in Fig. 11) yields the two second-order diagrams in the second row of Fig. 18. Finally, these two doubles amplitudes are antisymmetrised, yielding the last two diagrams in Fig. 18.

As can be seen, with this method two additional diagrams to those shown in Fig. 16 are obtained in second order. The corresponding exchange vacuum amplitude diagrams generated are displayed in Fig. 19. Note that the third diagram in the figure can alternatively be obtained by turning the second diagram upside down,

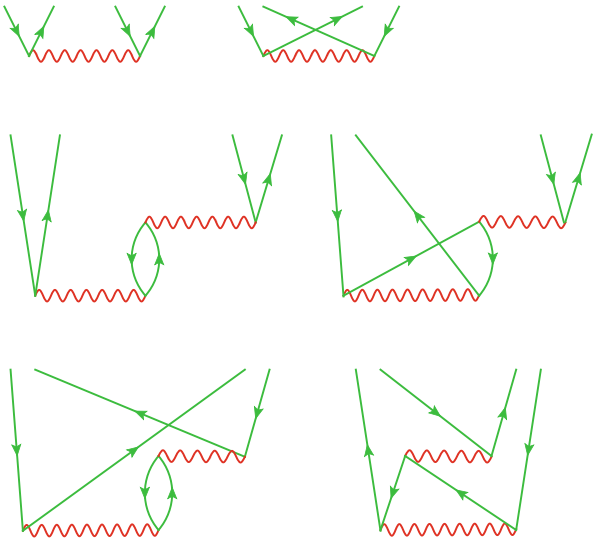


Fig. 18 First- and second-order doubles amplitudes described by the RPAX2 method

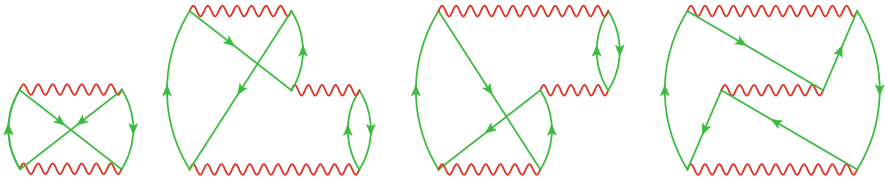


Fig. 19 Exchange vacuum amplitude diagrams of second- and third-order described by the RPAX2 method

which corresponds to taking the complex conjugate of it. Since now both the second and third diagram of Fig. 19 remove the EPV diagrams of Fig. 17, their sum overcorrects the third-order EPV contributions of the ring-diagram. However, the fourth diagram in Fig. 19 again has a positive prefactor (one loop and three hole-lines) and gives the same EPV contributions on magnitude as the ring-diagram. Because of this, in total the RPAX2 method also removes the RPA self-correlation error in third-order.

The SOSEX and RPAX2 methods contain only vacuum amplitudes which are of the particle–hole type, namely which correspond to amplitudes built by the fragments of the first row in Fig. 11, along with corresponding exchange fragments (bottom row in the figure). There also exist, however, exchange-RPA methods which contain further contributions which are of the mixed ph-pp or ph-hh type (third and fourth fragments in Fig. 11). These are presented in Fig. 20.

Again, individually labelled diagrams of Fig. 11 are EPV diagrams which are cancelled if the complete sum of these diagrams is taken. As an example, the second

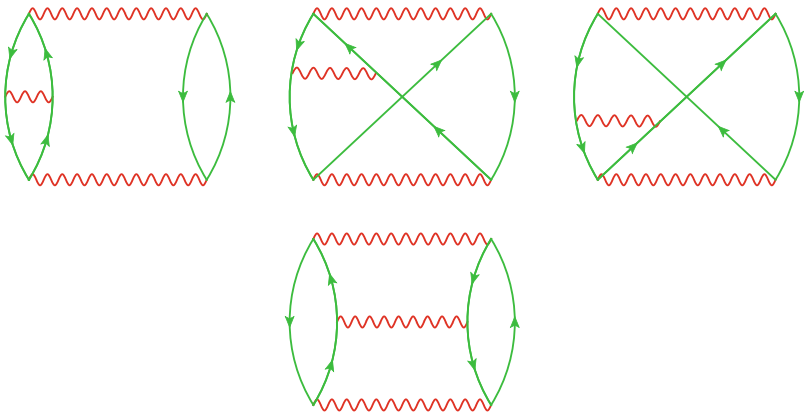


Fig. 20 Third-order mixed particle-hole and particle–particle/hole–hole diagrams not described by SOSEX and RPAX2. These amplitudes are, however, contained in other exchange-RPA variants

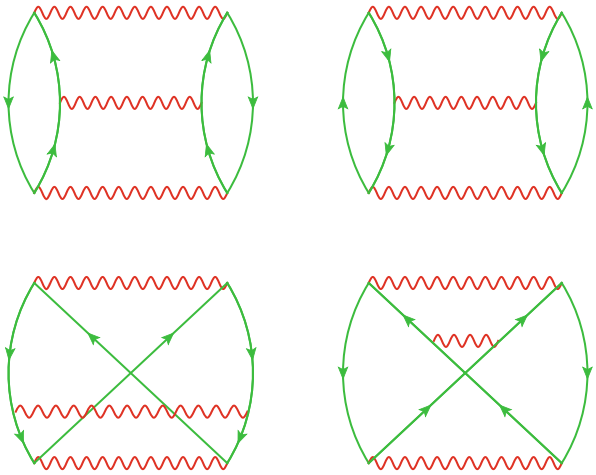


Fig. 21 Third-order particle–particle and hole–hole vacuum amplitudes not described by RPA methods

diagram cancels the EPV contributions of the first diagram because it can be obtained from it by an exchange of the two bottom particle lines and possesses a different prefactor (the first diagram has two closed loops and the second one closed loop).

Any RPA method, however, cannot be exact in third order in the electron–electron interaction. The reason is that particle–particle and hole–hole interactions are not described by the RPA. The corresponding missing third-order vacuum amplitude diagrams are shown in Fig. 21.

3 RPA Methods in a Kohn–Sham Orbital Basis

Commonly RPA methods were formulated based on Hartree–Fock (HF) reference states. This has the advantage that single excitation contributions to the vacuum amplitudes first appear in fourth order in the electron–electron interaction; see Sect. 2. The question is therefore which particular advantage it has to derive RPA methods in a different orbital basis such as the Kohn–Sham single-particle functions. In order to answer this question, first the transformation of the HF RPA amplitude equations to the KS-RPA amplitude equations is considered. The HF RPA amplitude equations, which can be obtained from the coupled-cluster doubles equations omitting all ladder contributions (see [32, 56]) is given by the Riccati-type equation:

$$\mathbf{B} + \varepsilon\mathbf{T} + \mathbf{T}\varepsilon + \mathbf{AT} + \mathbf{TA} + \mathbf{TBT} = \mathbf{0} \tag{1}$$

with $\varepsilon_{ia,jb} = (\varepsilon_a - \varepsilon_i)\delta_{ij}\delta_{ab}$ being a diagonal matrix containing HF orbital energy differences of occupied orbitals labeled i, j and unoccupied orbitals labeled a, b, and $\mathbf{A}$ and $\mathbf{B}$ are matrices defined by the antisymmetrised two-electron integrals $A_{ia,jb} = [ia||bj]$ and $B_{ia,jb} = [ia||jb]$ written in chemist's notation in this work. Note that the indices i and a are combined to a single superindex ia so that the matrices in (1) are of the size $N_{\mathrm{occ}} \times N_{\mathrm{virt}}$ with N_{occ} and N_{virt} being the number of occupied and virtual orbitals, respectively. The matrices $\mathbf{T}$ in (1) are the RPA doubles amplitudes. Note that the use of antisymmetrised integrals in (1) means that exchange vacuum amplitudes presented in Figs. 19 and 20 are accounted for in third-order and the underlying RPA method is sometimes termed also ring coupled-cluster doubles method.

In order to write (1) in terms of a Kohn–Sham orbital basis, a partitioning different to the conventional HF partitioning is chosen for the Hamiltonian operator. As a most natural choice, the Kohn–Sham partitioning is used so that the unperturbed Hamiltonian is just defined by the Kohn–Sham orbital energies $\varepsilon_r^{\mathrm{KS}}$ (with r referring to an arbitrary orbital). This can be formulated in second-quantised form as

$$H_0 = \sum_r \varepsilon_r^{\mathrm{KS}} \{r^\dagger r\} \tag{2}$$

$$\begin{aligned} V = W &+ \sum_{i,a} f_{ia} \left[\{a^\dagger i\} + \{i^\dagger a\}\right] + \sum_r \langle r|\hat{v}_{\mathrm{x}}^{\mathrm{HF}} - v_{\mathrm{xc}}|r\rangle \{r^\dagger r\} \\ &+ \sum_{i \neq j} \langle i|\hat{v}_{\mathrm{x}}^{\mathrm{HF}} - v_{\mathrm{xc}}|j\rangle \{i^\dagger j\} + \sum_{a \neq b} \langle a|\hat{v}_{\mathrm{x}}^{\mathrm{HF}} - v_{\mathrm{xc}}|b\rangle \{a^\dagger b\} \end{aligned} \tag{3}$$

with W being the electron–electron interaction operator and f_{ia} is an occupied-virtual Fock-matrix element. The fluctuation potential V contains differences of the matrix elements of the Hartree–Fock exchange potential $\hat{v}_{\mathrm{x}}^{\mathrm{HF}}$ and the Kohn–Sham exchange correlation potential v_{xc}; see (3).

A first approximation which is used from now on is to approximate the full exchange-correlation (xc) part of the Kohn–Sham potential by

$$v_{\mathrm{xc}}(\mathbf{r}) \approx v_{\mathrm{x}}(\mathbf{r}) \tag{4}$$

that is, the xc-potential is approximated by the local Kohn–Sham exchange potential. Because of this, the following derivations are strictly valid only for the exchange-only case, but we return to the description of correlated single-particle states later on. It is known now that, with a good accuracy, orbitals from a local exact exchange potential can be transformed to HF orbitals using occupied-occupied virtual-virtual unitary transformations and vice versa [66]. This means that the occupied-virtual matrix elements of the Fock matrix in (3) are close to zero so that Brillouins theorem is virtually fulfilled using an exact-exchange Kohn–Sham orbital basis. With this, the partitioning of the Hamiltonian in (2) and (3) is simplified and the RPA amplitude equations of (1) can be written as

$$\mathbf{B} + \boldsymbol{\varepsilon}^{\mathrm{KS}}\mathbf{T} + \mathbf{T}\boldsymbol{\varepsilon}^{\mathrm{KS}} + \widetilde{\mathbf{A}}\mathbf{T} + \mathbf{T}\widetilde{\mathbf{A}} + \mathbf{TBT} = \mathbf{0} \tag{5}$$

with $\boldsymbol{\varepsilon}^{\mathrm{KS}}$ containing the Kohn–Sham orbital energy differences and the matrix $\tilde{A}$ differs from the original matrix $\mathbf{A}$ by additional contributions originating from the differences of the matrix elements of the occupied-occupied and virtual-virtual blocks of the HF and Kohn–Sham exchange matrices, respectively:

$$\widetilde{A}_{ia,jb} = A_{ia,jb} + \Delta_{ia,jb} \text{ with } \Delta_{ia,jb} = \left\langle i\middle|\hat{v}_{\mathrm{x}}^{\mathrm{HF}} - v_{\mathrm{x}}\middle|j\right\rangle\delta_{ab} + \left\langle a\middle|\hat{v}_{\mathrm{x}}^{\mathrm{HF}} - v_{\mathrm{x}}\middle|b\right\rangle\delta_{ij}. \tag{6}$$

With this, the RPA equations in terms of a KS orbital basis has the same form as the RPA equations for HF references states in (1) but with the $\mathbf{A}$ matrix modified according to (6).

While in general the local exchange potential v_{x} has to be obtained by, e.g. the optimised effective potential approach [67–72] in which v_{x} can be calculated from a response-type equation containing the density differences of the HF and KS systems to first-order, in the case of closed-shell two-electron systems it is given exactly by minus one half of the Coulomb potential [36, 38, 73]. Then the Δ matrix in (6) reads $\Delta_{ia,jb} = [ij|ba] - [ib|ja]$ in this particular case of a two-electron system, i.e. with the restriction $i = j = 1$. Identifying that the sum $\mathbf{A} + \Delta$ is then identical to $\mathbf{B}$ (see above) means that the RPA equations can be rewritten as

$$\mathbf{B} + \boldsymbol{\varepsilon}^{\mathrm{KS}}\mathbf{T} + \mathbf{T}\boldsymbol{\varepsilon}^{\mathrm{KS}} + \mathbf{BT} + \mathbf{TB} + \mathbf{TBT} = \mathbf{0}. \tag{7}$$

In the special case of a closed-shell two-electron system, $B_{ia,jb} = \frac{1}{2}[ia|jb] = \frac{1}{2}C_{ia,jb}$, so (7) is identical to the direct RPA equations apart from the prefactor of one half for the two-electron integral matrices taking into account electron exchange effects.

Certainly, the approximations used in order to arrive at (7), especially the restriction to a two-electron system, are questionable in general. Nevertheless, in [45] it has been investigated whether an RPA method based on the amplitude equation (7) also gives a good description of electron correlation effects for general electronic systems. This method was termed RPAX1 in [45]. It was found therein that correlation energies calculated using the RPAX1 method are of about the same magnitude as corresponding coupled-cluster singles doubles (CCSD) correlation energies for a number of small organic molecules. Furthermore, for a set of 16 chemical reactions of organic compounds, the RPAX1 method is almost as accurate as the CCSD method if compared to corresponding accurate CCSD(T) reaction energies [45]. However, as is shown in [45], the RPAX1 method contains unphysical electron-correlation energy terms in third order. This problem is resolved by the RPAX2 method which is derived in [45] and reviewed in Sect. 2.

However, on account of the results of [45], it may nevertheless be concluded that the omission of differences of potential matrix elements (see (6)) combined with the omission of ph-pp/hh contributions (Fig. 20) is a reasonable approximation to be used in KS-RPA methods.

Such effective potential matrix elements also appear, however, in a many-body expansion of the singly excited amplitudes and these in turn contribute to the correlation energy already in second-order if a non-HF reference state is used. Moreover, the one-particle density matrix is wrong already in second-order without the inclusion of single excitations even in an HF orbital basis (thus effecting electron properties). The question is therefore how to account for single excitations in KS-based RPA methods?

To answer this, consider an RPA method in which the orbitals are optimised such as to make the RPA energy stationary with respect to orbital variations. Such a method can be realised, for example, by a unitary transformation of an original set of orbitals [74]. Considering the case that this would be done for the exact energy functional, not for the RPA one, it can be shown that the variationally optimised orbitals obtained in this way are identical to Brueckner orbitals, i.e. orbitals which yield a determinant Φ^{B} which has maximum overlap with the exact wave function Ψ^{exact} [75]:

$$\left\langle \Psi^{\mathrm{exact}} \middle| \Phi^{\mathrm{B}} \right\rangle = \text{maximum}. \tag{8}$$

It can easily be shown [58] that the maximum overlap criterion of (8) is identical to the requirement that single-excitation contributions to the exact wave function vanish, that is:

$$\left\langle \Psi^{\mathrm{exact}} \middle| \left(\Phi^{\mathrm{B}}\right)_i^a \right\rangle = 0, \tag{9}$$

where $(\Phi^{\mathrm{B}})_i^a$ is a singly-excited determinant with occupied orbital i replaced by an unoccupied orbital a. Because of this, in the case of an expansion of the exact wave

function, single excitations need not be considered if the orbitals are optimised such that the energy is minimised, i.e. the orbitals defined by

$$\frac{\partial E}{\partial \theta_{pq}} = 0 \tag{10}$$

(with θ_{pq} denoting an orbital rotation parameter) are identical to the orbitals defined by (8) and (9). Therefore, in the case of the exact wave function, conditions (10) and (9) are redundant.

Noting that the RPA wave function is not the exact one, however, it has to be analysed to which extent, for this particular case, the orbitals obtained by the two criteria (10) and (9) resemble each other. Previous studies for coupled cluster doubles methods [74] as well as coupled-pair functional methods [76] have also shown that, in the case of approximate wave functions, Brueckner orbitals from the projection condition (9) and the variationally optimised orbitals determined by condition (10) are close. Indeed, in [76] it was shown that the singles-amplitude equations determining Brueckner orbitals and the gradient equations for fulfilling the optimisation criterion are identical up to first order in the doubles amplitudes, providing analytical evidence for the similarity of projective and variationally optimised orbitals.

Because of this, it can also be concluded that in the case of approximate wave functions, such as the RPA wave function, single excitations can be omitted altogether if the orbitals are variationally optimised. Such optimised-orbital RPA methods based on Kohn–Sham reference states have previously been developed by several groups [77–79] and it was observed, for example, that the HOMO orbital energies from these methods are close to the exact ionisation potentials for a number of systems [78, 79]. This indicates that the optimised orbitals from an RPA functional are already close to the exact Kohn–Sham orbitals. In order to support this finding further, Fig. 22 displays the exact Kohn–Sham correlation potentials from Umrigar et al. [117] in comparison to the correlation potentials calculated by the optimised-orbital RPA method of [79] for the helium and neon atom. In addition, the diagrams in Fig. 22 also show the correlation potentials calculated by the Brueckner coupled-cluster doubles OEP (BCCD-OEP) method of [80]. As can be seen, both the RPA correlation potentials and the BCCD-OEP correlation potentials are fairly close to the exact KS correlation potentials. This in turn means that Kohn–Sham orbitals, too, resemble Brueckner orbitals or variationally optimised orbitals associated with a given wave function truncation to some extent (see also [75] for a theoretical argument of this point). Furthermore, these numerical results not only justify the use of RPA methods in conjunction with Kohn–Sham reference states, but also indicate why such KS-based RPA methods tend to describe the electronic structure much more accurately than HF-based RPA methods, because KS-based RPA methods to some degree can implicitly describe single excitations as opposed to HF-RPA methods.

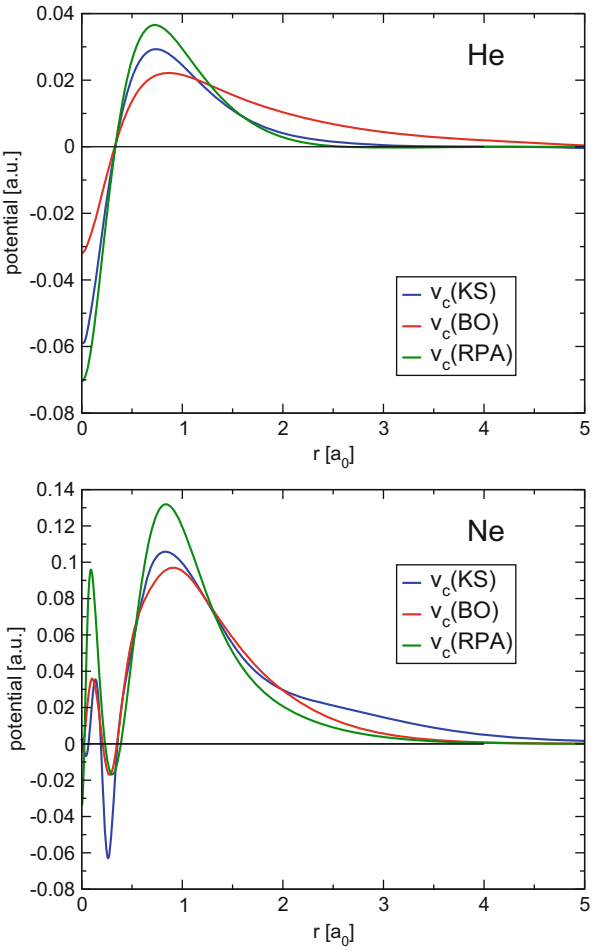


Fig. 22 Kohn–Sham (KS) correlation potentials of the helium (*top*) and neon (*bottom*) atom. Also shown in the diagrams are correlation potentials calculated from the density of the Brueckner determinant (BO) from a BCCD wave function and from the optimised-orbital RPA method

In most practical calculations, however, neither the exact KS orbitals nor variationally optimised orbitals are used in conjunction with the RPA. Instead, many KS-based RPA methods, such as those presented in this work, use orbitals from approximate GGA or hybrid-GGA functionals of a preceding self-consistent KS calculation. The question is thus how close these orbitals from approximate functionals resemble the variationally optimised ones. To analyse this, Fig. 23 shows the determinant overlaps of the optimised-orbital RPA determinant with the HF determinant as well as the determinants from the approximate Perdew–Burke–Ernzerhof (PBE) and PBE0 GGA and hybrid-GGA functionals for the neon atom and a number of small molecules, respectively. As can be seen, compared to the overlaps with the HF determinant, both the PBE and the PBE0 determinant strongly overlap with the optimised-orbital RPA determinant, the latter being slightly stronger than the former. With this, the use of KS orbitals from approximate

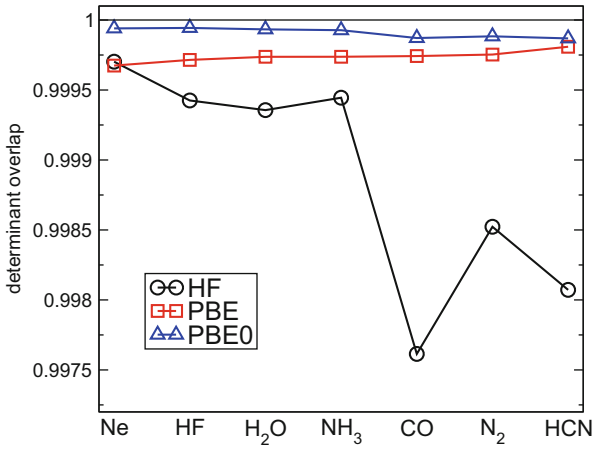


Fig. 23 Determinant overlaps of the optimised-orbital RPA determinant with the HF determinant and the determinants from the PBE and PBE0 xc functionals for some atomic and small molecular systems

functionals in RPA methods is also justified, since these too capture missing single excitations to some extent in contrast to HF-based RPA methods.

4 Performance of Kohn–Sham RPA Methods for Thermochemistry

In this section, the accuracy of Kohn–Sham RPA methods for describing thermochemical properties is investigated. For this, a number of existing databases were used, which are summarised in Table 1.

The databases IP13, EA13, PA8 and πTC13 are part of the Minnesota database for chemistry and solids by Truhlar et al. [94] and the databases G2RC, DARC, BSR36, NBPRC, BH76RC and BH76 are part of the GMTKN30 database by Grimme et al. [95–97]. In most cases, as can be seen in Table 1, the accurate reference values are theoretical results obtained by the CCSD with perturbative triples (CCSD(T)) method [58]. These values were extrapolated to the complete basis set limit using various basis sets; see references given in the last column of the table. In those cases where experimental values are taken as reference, databases IP13 and G2RC, the values are backcorrected by subtracting the zero-point vibrational energies (ZPVE) so that they can be compared directly to calculated results; see e.g. [84]. In the case of the ionisation energy and electron affinity calculations, relaxed cation and anion geometries were used, respectively.

To assess the performance of Kohn–Sham RPA methods for these databases, two exchange-RPA variants along with the standard (direct) RPA method were used, namely the SOSEX [43] and RPAX2 methods [45]; see Sect. 2 for a description of these methods.

In the case of the RPA and the SOSEX methods, the electron–electron interaction energy was calculated using orbitals from a preceding self-consistent

Table 1 Summary of databases for thermochemical properties used in this work

Database	Description	Reference method	References
HEAT	Atomisation energies	est. CCSD(T)/cbs	[81, 82]
IP13	Ionisation energies	exp. and QCISD	[83, 84]
EA13	Electron affinities	G3	[83, 84]
PA8	Proton affinities	est. CCSD(T)/cbs	[85]
πTC13	Thermochemistry of π-systems	est. CCSD(T)/cbs	[86]
G2RC	Reaction energies of selected G2/97 systems	exp.	[87]
DARC	Reaction energies of Diels–Alder reactions	est. CCSD(T)/cbs	[88]
BSR36	Bond separation reactions of saturated hydrocarbons	est. CCSD(T)/cbs	[89]
NBPRC	Oligomerisations and H_2 fragmentations of NH_3/BH_3 systems; H_2 activation reactions with PH_3/BH_3 systems	est. CCSD(T)/cbs	[90, 91]
BH76RC	Reaction energies of the BH76 set	W1 and theor. est.	[92, 93]
BH76	Barrier heights of hydrogen transfer, heavy atom transfer, nucleophilic substitution, unimolecular, and association reactions	W1 and theor. est.	[92, 93]

Kohn–Sham calculation employing the PBE exchange-correlation functional [98]. In contrast to this, the RPAX2 electron–electron interaction energy was calculated using orbitals from PBE exchange-only Kohn–Sham calculations. It was found in [45] that this gives slightly better intermolecular interaction energies compared to RPAX2 used in conjunction with PBE orbitals. It should be noted, however, that, for the thermochemical properties considered in this work, the differences between the results from RPAX2/PBE and RPAX2/PBEx calculations are relatively small. This means that in the following analysis differences caused by the use of different orbitals in the RPA and exchange-RPA calculations can be neglected.

In order to capture effectively the dynamic correlation effects described by the RPA methods, all results, except of those for the IP13, EA13 and PA8 databases, were obtained using the two-point basis-set extrapolation method by Bak et al [99]. For this, the corresponding augmented triple- (aug-cc-pVTZ) and quadruple-zeta (aug-cc-pVQZ) basis sets by Dunning et al. were used. The results for the IP13, EA13 and PA8 databases, however, were calculated using only the aug-cc-pVQZ basis set because the properties of these databases are less basis set dependent. Also, the zeroth- and first-order energies (Fig. 9) have not been extrapolated but were also calculated using solely the aug-cc-pVQZ basis set which should provide energies close to the complete basis set limit in this case.

The RPA and RPAX2 amplitude equations were solved using the density-fitting algorithms described in [45]. In the case of the RPA method, the denominator

decomposition was done using an incomplete Cholesky decomposition with a threshold of 10^{-8} for deleting linearly dependent Cholesky vectors. In the case of both the RPA and the RPAX2 method, the corresponding aug-cc-pV(T,Q)Z-MP2Fit basis sets by Weigend et al. [100] have been used for the two-electron integral decomposition.

With the exception of the calculations of the atomisation energies, core electrons were not correlated in the calculations. Open shell systems treated in this work were calculated with a spin unrestricted reference state to enable full symmetry breaking. All calculations were done using a developers' version of the Molpro quantum chemistry program [101].

4.1 Atomisation Energies, Ionisation Energies and Electron and Proton Affinities

In this section, the performance of RPA, SOSEX and RPAX2 is investigated for the following reactions:

$$\mathrm{M}(\mathbf{R}_e) \rightarrow \sum_i \mathrm{A}_i, \tag{11}$$

$$\mathrm{M}(\mathbf{R}_e) \rightarrow \mathrm{M}\left(\widetilde{\mathbf{R}}_e\right)^+ + e^-, \tag{12}$$

$$\mathrm{M}(\mathbf{R}_e) + e^- \rightarrow \mathrm{M}\left(\widetilde{\mathbf{R}}_e\right)^-, \tag{13}$$

$$\mathrm{M}(\mathbf{R}_e) + \mathrm{H}^+ \rightarrow \mathrm{MH}\left(\widetilde{\mathbf{R}}_e\right)^+, \tag{14}$$

where $\mathrm{M}(\mathbf{R}_\mathrm{e})$ corresponds to the molecule in its equilibrium geometry, $\widetilde{\mathbf{R}}_e$ denotes a relaxed geometry of the given compound and $\sum_i \mathrm{A}_i$ corresponds to the sum of the constituent atoms of molecule M. The reaction in (11) describes the atomisation, the reaction in (12) the ionisation, the reaction in (13) the electron affinity and the reaction in (14) the proton affinity of molecule M.

Clearly, a direct comparison of the calculated reaction energies of the reactions in (11)–(14) to experimental ones is not straightforward because, apart from the electronic ground-state energies accessible through quantum-chemical calculations, other energy terms also need to be accounted for, most notably the ZPVE of the molecules in the given equilibrium structures and, in order to achieve an even higher accuracy, spin-orbit interactions. Since the aim of this work is to analyse specifically the role of electron correlation effects which contribute to the electronic ground-state energies, however, all results in this and all subsequent sections present only the dominating electronic contributions to the given reactions, i.e. vibrational and relativistic effects are excluded. It should therefore be noted that the comparison to experimental results is strictly only possible if at least the

Table 2 HEAT database: atomisation energies in kcal/mol

System	RPA	SOSEX	RPAX2	Ref.[a]
N_2	−221.17	−215.06	−227.98	−228.15
H_2	−108.73	−109.54	−112.03	−109.49
F_2	−29.93	−24.58	−37.58	−38.45
CO	−243.21	−251.86	−260.25	−259.72
O_2	−112.20	−103.74	−116.37	−120.27
C_2H_2	−381.24	−399.44	−407.97	−405.45
CCH	−243.79	−257.97	−264.69	−265.21
CF	−119.98	−127.38	−132.50	−132.67
CH_2	−178.91	−190.99	−191.75	−190.71
CH	−79.98	−82.98	−85.80	−84.07
CH_3	−294.63	−308.19	−311.13	−307.84
CN	−170.31	−161.69	−177.46	−179.38
CO_2	−363.31	−376.54	−389.43	−390.20
H_2O_2	−256.21	−257.28	−270.60	−268.96
H_2O	−224.03	−229.92	−235.56	−233.04
HCN	−297.20	−302.57	−313.86	−313.21
HCO	−262.87	−269.39	−278.88	−279.23
HF	−133.43	−139.82	−143.09	−141.72
HNO	−197.84	−191.88	−206.71	−205.54
HO_2	−165.75	−161.57	−174.16	−175.05
NH_2	−179.03	−179.51	−184.66	−182.44
NH_3	−291.09	−295.12	−302.12	−298.00
NH	−81.95	−80.84	−83.70	−82.97
NO	−146.28	−138.87	−150.82	−152.25
OF	−45.92	−39.19	−50.65	−52.20
OH	−103.76	−105.09	−108.23	−107.17
MAE	10.41	7.84	1.53	

[a]CCSD(T)/cbs, see [82]

ZPVE corrections are added to the presented results, because these can change the electronic contribution to the reaction energies by 10% and even more for a given system; see [102].

Table 2 displays the atomisation energies (see (11)) for the HEAT database of [81, 82] calculated with the RPA, SOSEX and RPAX2 methods. The last column in the table shows the CCSD(T) complete basis set extrapolated (cbs) reference results of [82] for comparison. The description of atomisation energies is actually a challenging test for electron correlation methods because they are calculated from energy differences of systems with a different number of paired electrons, requiring a balanced description of short-range interactions. This of course raises the question whether the CCSD(T) reference method itself can provide a high accuracy for atomisation energies. This has been thoroughly investigated in the past – see, for instance, [102] (Chap 15.7), [103] (Chap. 18.2) and [82] – and it turns out that, for molecules of the size of those in Table 2, the CCSD(T) method yields atomisation

energies which only deviate by about 0.3 kcal/mol on average from the experiment (if ZPVEs and other corrections are accounted for; see above).

The values presented in Table 2 show that the accuracy of the respective RPA methods obviously depends on their capturing of exchange interactions. Most notably, the mean absolute error (MAE) for the respective RPA method, displayed in the last line in the table, decreases from 10.4 over 7.8 to only 1.5 kcal/mol with no (RPA), screened second-order (SOSEX) and all-order exchange interactions (RPAX2). A closer inspection of the sources of the relatively large error of the RPA method shows that it has especially difficulties in describing the atomisation energies for molecules containing double or triple bonds e.g. CO_2 and C_2H_2, for which the RPA method strongly underestimates the atomisation energy by over 20 kcal/mol. Compared to this, the SOSEX method gives a much more balanced description for the atomisation energies of the molecules in the table; however, it does not significantly reduce the mean average error of the RPA method since, for some molecules, the SOSEX atomisation energies deviate more strongly from the reference energies than the RPA ones; see Table 2. In contrast to this, the RPAX2 atomisation energies are almost always closer to the CCSD(T) reference values than the RPA ones, the exception being only the H_2 molecule. Overall, the RPAX2 method gives a strong improvement over both the RPA the and SOSEX methods for the considered property. While not directly comparable to results from the literature because of differing databases, it is nevertheless remarkable that the small error of 1.5 kcal/mol for the atomisation energies of the HEAT database obtained by the RPAX2 method is significantly smaller than typical errors for a range of established [103] (Chap. 18.2) and modern [104] (Chap. 8.6.1) [97, 105, 106] (hybrid/meta)-GGA based density functionals for which errors of the order of typically 3–6 kcal/mol are obtained.

Tables 3, 4 and 5 show the results for ionisation energies, electron affinities and proton affinities, respectively (see (12)–(14)). Conventionally, the former two properties are sometimes also calculated using Koopman's theorem in the framework of HF and Kohn–Sham DFT methods. By this, however, it is only possible to capture the correlation contributions while relaxation contributions, both electronic and structural, cannot be described with this approach. Also, in the case of DFT methods, Koopman's theorem cannot be extended to the description of electron affinities as in case of HF because, in Kohn–Sham DFT, where the electrons are bound by a local effective potential, all electrons 'feel' the effective field of $N-1$ electrons so that the unoccupied orbital levels in an $N+1$ system are not proper electron affinities [107]. Furthermore, with approximate DFT methods the extra electron in the radical anion is often loosely bound so that results for electron affinities can be subject to large overdelocalisation errors.

In this work, however, all results presented in Tables 3 and 4 correspond to adiabatic ionisation potentials and electron affinities, respectively. The results in Table 3 are therefore not directly comparable to the ionisation energies of [79] which were obtained from the HOMO energies using the orbital-optimised RPA method. Nevertheless, the RPA ionisation energies shown in Table 3 roughly deviate by the same magnitude from the experimental values compared to the

Table 3 IP13 database: ionisation energies in kcal/mol

System	RPA	SOSEX	RPAX2	Ref.[a]
C	265.77	258.50	257.76	259.74
S	248.63	233.97	236.24	238.34
SH	247.58	237.43	239.13	238.36
Cl	304.36	296.31	297.51	299.31
Cl_2	266.74	268.57	266.60	265.30
OH	309.09	296.83	299.38	298.90
O	325.53	308.39	311.08	313.67
O_2	283.70	279.79	276.83	278.90
P	239.94	241.60	241.50	242.80
PH	237.51	234.42	233.14	234.10
PH_2	231.81	225.99	224.27	226.30
S_2	220.77	218.24	215.61	216.00
Si	194.09	186.17	189.62	188.05
MAE	6.27	2.08	1.49	

[a]exp. and QCISD, see [83, 84]

Table 4 EA13 database: electron affinities in kcal/mol

System	RPA	SOSEX	RPAX2	Ref.[a]
C	32.50	24.72	26.23	29.19
S	53.52	45.37	46.92	47.91
SH	57.41	52.34	53.37	53.84
Cl	84.08	82.96	83.05	84.24
Cl_2	65.96	53.35	53.45	55.60
OH	46.40	35.30	41.03	42.30
O	40.93	26.15	31.40	33.77
O_2	17.12	6.54	8.29	10.80
P	32.20	16.30	15.65	16.92
PH	30.29	20.94	22.93	23.20
PH_2	34.46	27.47	28.98	29.40
S_2	44.99	38.24	37.99	38.50
Si	31.54	29.86	30.45	32.33
MAE	5.79	2.96	1.40	

[a]G3, see [83, 84]

orbital-optimised RPA results of [79], namely by 5–6 kcal/mol. This is about the same order of magnitude as is obtained with standard density-functional methods if the ΔSCF method is used [104, 105]. Compared to this, both the SOSEX and RPAX2 functionals yield a substantial improvement over the RPA results, for both the ionisation energies and the electron and proton affinities; see Tables 3, 4, and 5. While the RPAX2 values are somewhat more accurate than the SOSEX values, both the SOSEX and RPAX2 methods describe these properties clearly better than a range of modern density-functionals tested for these particular databases; see [86]. As an example, the M05 functional by Zhao and Truhlar has MAEs of 2.9, 3.0 and 2.2 kcal/mol for the IP13, EA13 and PA8 databases, respectively [86].

Table 5 PA8 database: proton affinities in kcal/mol

System	RPA	SOSEX	RPAX2	Ref.[a]
NH_3	214.08	212.87	212.50	211.90
H_2O	173.50	172.77	172.28	171.80
C_2H_2	161.70	160.46	159.09	156.60
SiH_4	160.06	156.42	157.10	156.50
PH_3	196.21	194.22	193.85	193.10
H_2S	178.30	174.43	174.91	173.70
HCl	142.72	138.22	139.01	137.10
H_2	107.63	105.98	106.32	105.90
MAE	3.45	1.12	1.06	

[a]CCSD(T)/cbs, see [85]

Table 6 πTC13 database: isomeric energy differences between cumulenes and polyenes (lines 1–3), proton affinities of five conjugated polyenes (lines 4–8) and proton affinities of the five conjugated Schiff bases (lines 9–13)

System	RPA	SOSEX	RPAX2	Ref.[a]
Polyenes – cumulenes				
C_3H_4	−0.34	−1.47	−1.04	−1.40
C_5H_4	−6.62	−8.85	−8.22	−8.80
C_7H_4	−10.96	−13.33	−12.86	−14.30
Conjugated polyenes				
C_2H_4 - $C_2H_5^{\oplus}$	173.24	170.18	168.71	167.81
C_4H_6 - $C_4H_7^{\oplus}$	201.69	197.02	196.10	193.45
C_6H_8 - $C_6H_9^{\oplus}$	219.60	214.18	213.54	209.68
C_8H_{10} - $C_8H_{11}^{\oplus}$	230.96	225.15	224.58	219.67
$C_{10}H_{12}$ - $C_{10}H_{13}^{\oplus}$	239.01	232.99	232.43	225.95
Conjugated Schiff bases				
CH_2NH - $CH_2NH_2^{\oplus}$	219.25	217.43	215.68	214.46
C_3H_4NH - $C_3H_4NH_2^{\oplus}$	232.55	230.24	228.37	226.15
C_5H_6NH - $C_5H_6NH_2^{\oplus}$	241.46	238.90	236.93	233.44
C_7H_8NH - $C_7H_8NH_2^{\oplus}$	247.73	244.96	242.94	238.16
$C_9H_{10}NH$ - $C_9H_{10}NH_2^{\oplus}$	252.45	249.57	247.46	240.97
MAE	7.29	4.00	3.03	

Energies are in kcal/mol
[a]CCSD(T)/cbs, see [86]

Proton affinities were also studied for a number of π-systems from the πTC13 database [86]; see Table 6. In addition, the table contains the isomeric energy differences between three polyene/cumulene systems. One can see that, compared to the proton affinities of the eight small molecules in Table 5, the errors of the three RPA methods considered are now somewhat larger and clearly increase with the size of the polyene or Schiff bases of the database. The same is also observed for a range of density-functionals which have been tested for the πTC13 database

[86]. As an example, the M05 functional, which performs best among the 15 functionals selected in [86], has an average unsigned error of 5.1 kcal/mol for the π-systems of Table 6. Thus, again, the performance of the RPA is comparable to the performance of standard DFT methods, while both the SOSEX and the RPAX2 methods yield proton affinities which are significantly more accurate than with current density functionals.

4.2 Reaction Energies

The accuracy of RPA, SOSEX and RPAX2 for reaction energies is assessed by using a variety of different chemical reactions contained in five different databases. The G2RC database is a selection of 25 chemical reactions of small closed-shell molecules containing main-group elements from the first to third row of the periodic table [87]. The DARC database contains 15 Diels–Alder reactions; for reactions of the type see [88]. The BSR36 database contains 36 bond separation reactions of saturated hydrocarbons, i.e. chemical reactions of the general form

$$(m+1)\mathrm{C_2H_6} \rightarrow \mathrm{CH_3(CH_2)_{\mathit{m}}CH_3} + m\mathrm{CH_4}.$$

This database was developed by Steinmann et al. in order to study the effect of intramolecular correlation effects in DFT + D methods and thus contains relatively large reactants (with m up to 17) [89]. The NBPRC database is a set of six oligomerisation and dihydrogen fragmentation reactions of ammonia/borane systems as well as six H_2 activation reactions of corresponding ammonia/borane systems [90, 91]. The general form of the reactions of this database are

$$\begin{gathered} m\mathrm{NH}_n + m\mathrm{BH}_n \rightarrow m\mathrm{NH}_n\mathrm{BH}_n \\ m\mathrm{NH}_n\mathrm{BH}_n \rightarrow \mathrm{N}_n\mathrm{B}_n\mathrm{H}_{n-2} + \mathrm{H}_2 \\ \mathrm{PR}^1_3 + \mathrm{BR}^2_3 \rightarrow \mathrm{R}^1_3\mathrm{P} - \mathrm{BR}^2_3 \\ \mathrm{R}^1_3\mathrm{P} - \mathrm{BR}^2_3 + \mathrm{H}_2 \rightarrow \left[\mathrm{PR}^1_3\mathrm{H}\right]^+\left[\mathrm{BR}^2_3\mathrm{H}\right]^- \end{gathered}$$

with R^1 and R^2 denoting different substituents; see [96]. These reactions have been considered in [96, 108] because it turns out that standard DFT methods such as the B3LYP method [109] have difficulties in describing the formation of boron–nitrogen bonds; see [110].

Finally, the reactions of the BH76RC database are considered, which include four heavy atom transfer and 17 hydrogen transfer reactions of the type

$$\mathrm{R^1A} + \mathrm{R^2} \rightarrow \mathrm{R^1} + \mathrm{R^2A}$$

with A denoting either hydrogen or a heavy atom (in this database only elements from the second row are considered), two nucleophilic substitution reactions, i.e. reactions of the type

$$A^- + BR \rightarrow AR + B^-$$

and finally five unimolecular and association reactions of the type

$$R^1 + R^2 \rightarrow R^1 - R^2$$

and

$$R^1 \rightarrow R^2.$$

The reactions of the BH76RC database were developed by Truhlar et al. in order to study the forward and reverse reaction barriers of the chemical reaction considered; see [92, 93].

In Sect. 4.1 it has already been mentioned that, in order to be able to compare the reaction energies from quantum chemical calculations to experimental ones, it is necessary to decompose the latter into distinct contributions. These are, at 0 K, the nonrelativistic electronic reaction enthalpies at the equilibrium, a vibrational contribution taking into account the zero-point vibrations (ZPVE) of the molecules, and a relativistic correction term. As in the case of atomisation energies (see Sect. 4.1), the molecular ZPVEs can be significant and thus cannot be neglected if a comparison with experimental reaction energies is made. It is fortuitous, however, that, for chemical reactions where the reactants are in their ground-state equilibrium, it usually suffices to describe the molecular ZPVEs using the harmonic approximation. As compared to this, the relativistic corrections are usually very small if the reactants contain only elements from the first to third rows; see [102].

In the current work, the only database for which backcorrected (i.e. ZPVE and spin-orbit excluding) experimental results are used as reference results is the G2RC database; see Table 7 and [87]. The DARC, BSR36 and NBPRC databases (see Tables 8, 9 and 10) all use estimated complete basis set CCSD(T) results as reference values instead. It can be assumed that these are accurate to within about 0.4 kcal/mol if compared to experimental reaction enthalpies [102]. In the case of the BH76RC database, the W1 method by Martin and Oliveira [111] is used as the reference method. Basically, this method is designed to extrapolate to the complete basis set limit of a CCSD(T) calculation and thus can be expected to yield accuracies similar to raw CCSD(T)/cbs calculations for reaction energies, i.e. about 0.5 kcal/mol [111].

The RPA, SOSEX and RPAX2 reaction energies for the individual databases are compiled in Tables 7, 8, 9, 10 and 11. Considering the G2RC reaction (Table 7), it can be observed that there are few reactions for which all RPA methods fail to reproduce accurately the experimental reference results, namely the water abstraction reaction of $CH_3(SO)CH_3$, the silicon-carbon swapping reaction $SiCl_4 + CF_4$ and the hydrogenation reaction of SO_2. The SOSEX method in particular yields reaction energies which strongly deviate from the experimental values for these three cases, e.g. for the SO_2 hydrogenation the error is larger than 18 kcal/mol and explains that overall the SOSEX method performs worse than RPA and RPAX2.

Table 7 G2RC database: reaction energies of selected G2/97 systems in kcal/mol

System	RPA	SOSEX	RPAX2	Ref.[a]
$CH_3(SO)CH_3 \rightarrow C_2H_4S + H_2O$	−6.19	−8.67	−6.48	−0.96
$HCOOH \rightarrow CO_2 + H_2$	−0.83	−3.74	−1.78	−2.03
$CH_3CHO \rightarrow CO + CH_4$	−5.26	−3.37	−4.24	−2.55
$Si_2H_6 + H_2 \rightarrow 2SiH_4$	−4.70	−4.62	−4.23	−3.88
$CO + H_2O \rightarrow CO_2 + H_2$	−5.09	−4.06	−5.48	−7.01
$CH_3COCl + H_2O \rightarrow CH_3COOH + HCl$	−9.78	−11.88	−10.61	−9.96
$CH_3CN + H_2O \rightarrow CH_3(CO)NH_2$	−17.30	−21.16	−19.09	−20.13
$C_2H_2 + HF \rightarrow CH_2CHF$	−26.45	−27.69	−26.53	−25.37
$F_2 + Cl_2 \rightarrow 2FCl$	−25.70	−28.61	−27.07	−26.42
$BCl_3 + AlF_3 \rightarrow BF_3 + AlCl_3$	−27.67	−32.99	−30.39	−26.85
$SiCl_4 + CF_4 \rightarrow SiF_4 + CCl_4$	−23.69	−15.81	−20.04	−26.33
$H_2CO + H_2 \rightarrow CH_3OH$	−27.00	−32.25	−29.27	−29.23
$C_2H_2 + C_2H_4 \rightarrow$ cyclobutene	−30.97	−35.62	−31.93	−32.73
$SiO + 3H_2 \rightarrow SiH_4 + H_2O$	−33.68	−43.86	−37.25	−33.91
$N_2 + 3H_2 \rightarrow 2NH_3$	−33.54	−44.90	−38.95	−38.87
$CH_4 + 2Cl_2 \rightarrow CH_2Cl_2 + 2HCl$	−45.20	−52.70	−48.67	−47.08
$C_2H_2 + H_2 \rightarrow C_2H_4$	−48.10	−51.64	−49.32	−48.42
$SO_2 + 3H_2 \rightarrow H_2S + 2H_2O$	−62.98	−78.65	−68.31	−60.47
$CO + 3H_2 \rightarrow H_2O + CH_4$	−58.89	−68.70	−63.80	−64.78
$OF_2 + H_2 \rightarrow F_2 + H_2O$	−66.09	−74.34	−69.92	−68.72
$N_2O + H_2 \rightarrow N_2 + H_2O$	−79.90	−86.73	−82.20	−80.74
$C_2H_4 + CH_2 \rightarrow$ cyclopropane	−105.99	−111.19	−111.17	−109.06
$H_2 + F_2 \rightarrow 2HF$	−127.36	−143.99	−135.61	−134.29
$3C_2H_2 \rightarrow C_6H_6$	−148.40	−166.37	−155.07	−151.57
$Li_2 + F_2 \rightarrow 2LiF$	−201.98	−216.13	−207.10	−212.70
MAE	2.71	5.30	2.14	

[a]Experimental, see [87]

A comparison of the MAEs of RPA and RPAX2 for the G2RC database (see last line in Table 7) to results from other quantum chemistry methods shows that they lie in the range of the errors obtained by the MP2 and SCS-MP2 methods which yield 3.6 and 2.2 MAEs, respectively (see [97] and Fig. 24). In contrast to this, DFT methods (raw and dispersion corrected) typically have larger errors for this benchmark set [97]. For example, the reaction energies from the M06 functional [112] (with or without dispersion correction) deviate by 4.3 kcal/mol from the experimental reference results on average [97].

For the Diels–Alder reactions displayed in Table 8, no such strong deviations as for the G2RC set of the reactions from the reference results are observed with the (exchange-)RPA methods. However, in the case of these chemical reactions, it can also be seen that the SOSEX reaction energies in most cases do not improve the RPA reaction energies, again indicating that the exchange interactions described by the SOSEX method obviously give an unbalanced description of the correlation

Table 8 DARC database: Diels–Alder reactions in kcal/mol

System	RPA	SOSEX	RPAX2	Ref.[a]
$C_2H_4 + CH_2CHCHCH_2 \rightarrow [C_6H_{10}]^b$	−42.33	−49.08	−44.38	−43.80
$C_2H_2 + CH_2CHCHCH_2 \rightarrow [C_6H_8]^b$	−57.96	−64.91	−60.04	−59.30
$C_2H_4 + \text{cyc-}C_5H_6 \rightarrow [C_7H_{10}]^b$	−28.28	−31.88	−28.69	−30.00
$C_2H_2 + \text{cyc-}C_5H_6 \rightarrow [C_7H_8]^b$	−32.69	−35.29	−32.60	−33.10
$C_2H_4 + \text{cyc-}C_6H_8 \rightarrow [C_7H_{12}]^b$	−35.47	−40.36	−36.54	−36.50
$C_2H_2 + \text{cyc-}C_6H_8 \rightarrow [C_7H_{10}]^b$	−46.99	−52.15	−48.25	−48.20
$C_4H_4O + C_4H_2O_3 \rightarrow [C_8H_6O_4]^b$	−13.15	−14.26	−12.43	−14.40
$C_4H_4O + C_4H_2O_3 \rightarrow [C_8H_6O_4]^{b,c}$	−14.86	−16.37	−14.35	−16.20
$C_4H_4O + C_4H_3O_2N \rightarrow [C_8H_7O_3N]^b$	−15.75	−17.30	−15.14	−17.20
$C_4H_4O + C_4H_3O_2N \rightarrow [C_8H_7O_3N]^{b,c}$	−17.59	−19.51	−17.21	−19.20
$\text{cyc-}C_5H_6 + C_4H_2O_3 \rightarrow [C_9H_8O_3]^b$	−30.60	−33.85	−30.78	−31.60
$\text{cyc-}C_5H_6 + C_4H_2O_3 \rightarrow [C_9H_8O_3]^{b,c}$	−30.97	−34.25	−31.20	−32.10
$\text{cyc-}C_5H_6 + C_4H_3O_2N \rightarrow [C_9H_9O_2N]^b$	−32.79	−36.51	−33.12	−34.10
$\text{cyc-}C_5H_6 + C_4H_3O_2N \rightarrow [C_9H_9O_2N]^{b,c}$	−33.03	−36.78	−33.41	−34.40
MAE	1.26	2.33	1.06	

[a]CCSD(T)/cbs, see [88]
[b]Denotes a molecular transition state
[c]Denotes an alternate conformer to the reaction product of the respective reaction of the previous line

effects within the educts and products. This is again strongly corrected by the RPAX2 method, which yields a comparable average error of 1.1 kcal/mol to the CCSD(T) reference reaction energies as RPA; see Table 8. This is clearly better than with, for example, the MP2 method, yielding an error of 4.0 kcal/mol or the B3LYP method with an error of 15 kcal/mol, the latter thus dramatically failing in describing the considered Diels-Alder reactions [97]. More modern density-functionals such as the M06 functional or the B2PLYP double hybrid functional [113] are also much less accurate than the RPA methods considered here: they yield MAEs of 3.7 and 5.0 kcal/mol, respectively, for the DARC database if combined with the D3 dispersion correction by Grimme [91, 97].

As opposed to the G2RC and DARC reactions, for the reactions of the BSR36, NBPRC and NH76RC databases the SOSEX method tends to describe the respective reactions better than the RPA method. This is true in particular for the bond separation reactions of Table 9 and the reactions involving B/N/P-compounds of Table 10. Again, the RPAX2 reaction energies have significantly smaller errors for these sets than the SOSEX reaction energies, as can be observed in Tables 9, 10, and 11.

In comparison to a large number of benchmark results for standard GGA, hybrid-and meta-GGA functionals as well as double-hybrid functionals (see [96, 97]), the performance of the RPAX2 method for the selected databases is remarkable. In order to highlight this, Fig. 24 shows the MAEs of the RPA methods for the different databases in comparison to the errors of a selection of other wave-function and DFT methods, namely second-order Møller–Plesset

Table 9 BSR36 database: bond separation reactions of saturated hydrocarbons in kcal/mol

System	RPA	SOSEX	RPAX2	Ref.[a]
$5C_2H_6 \rightarrow [C_6H_{14}]^1 + 4CH_4$	8.50	9.67	9.84	9.81
$5C_2H_6 \rightarrow [C_6H_{14}]^2 + 4CH_4$	8.53	9.70	9.66	9.66
$5C_2H_6 \rightarrow [C_6H_{14}]^3 + 4CH_4$	9.97	11.23	11.37	11.37
$5C_2H_6 \rightarrow [C_6H_{14}]^4 + 4CH_4$	7.87	9.01	9.09	9.04
$5C_2H_6 \rightarrow [C_6H_{14}]^5 + 4CH_4$	7.81	8.87	8.72	8.71
$6C_2H_6 \rightarrow [C_6H_{14}]^6 + 5CH_4$	9.80	11.11	10.95	10.91
$6C_2H_6 \rightarrow [C_6H_{14}]^7 + 5CH_4$	11.43	13.00	13.09	13.05
$6C_2H_6 \rightarrow [C_6H_{14}]^8 + 5CH_4$	10.54	11.96	11.92	11.89
$6C_2H_6 \rightarrow [C_6H_{14}]^9 + 5CH_4$	11.91	13.40	13.55	13.53
$6C_2H_6 \rightarrow [C_6H_{14}]^{10} + 5CH_4$	9.87	11.22	11.52	11.43
$6C_2H_6 \rightarrow [C_6H_{14}]^{11} + 5CH_4$	11.14	12.57	12.99	12.97
$6C_2H_6 \rightarrow [C_6H_{14}]^{12} + 5CH_4$	11.04	12.48	12.92	12.77
$6C_2H_6 \rightarrow [C_6H_{14}]^{13} + 5CH_4$	9.84	11.22	11.30	11.23
$6C_2H_6 \rightarrow [C_6H_{14}]^{14} + 5CH_4$	8.73	10.01	10.30	10.16
$7C_2H_6 \rightarrow [C_6H_{14}]^{15} + 6CH_4$	12.70	14.17	15.06	15.05
$5C_2H_6 \rightarrow [C_5H_{10}] + 5CH_4$	2.07	2.68	2.13	2.38
$6C_2H_6 \rightarrow [C_5H_{13}]^1 + 6CH_4$	9.77	11.10	10.80	10.67
$6C_2H_6 \rightarrow [C_5H_{13}]^2 + 6CH_4$	5.67	6.70	6.12	6.35
$7C_2H_6 \rightarrow [C_7H_{14}]^1 + 7CH_4$	13.54	15.36	14.99	14.88
$7C_2H_6 \rightarrow [C_7H_{14}]^2 + 7CH_4$	9.47	10.82	10.47	10.65
$7C_2H_6 \rightarrow [C_7H_{14}]^3 + 7CH_4$	9.07	10.47	9.92	10.11
$7C_2H_6 \rightarrow [C_7H_{14}]^4 + 7CH_4$	8.06	9.31	9.05	9.13
$7C_2H_6 \rightarrow [C_7H_{14}]^5 + 7CH_4$	9.36	10.85	10.30	10.48
$7C_2H_6 \rightarrow [C_7H_{14}]^6 + 7CH_4$	8.79	10.13	9.66	9.82
$11C_2H_6 \rightarrow [C_{10}H_{18}]^1 + 12CH_4$	17.36	19.70	19.42	19.30
$11C_2H_6 \rightarrow [C_{10}H_{18}]^2 + 12CH_4$	8.04	9.52	10.13	9.67
$11C_2H_6 \rightarrow [C_{10}H_{18}]^3 + 12CH_4$	13.56	15.55	15.42	15.19
$11C_2H_6 \rightarrow [C_{10}H_{18}]^4 + 12CH_4$	23.86	26.90	26.30	26.05
$11C_2H_6 \rightarrow [C_{10}H_{18}]^5 + 12CH_4$	21.34	24.02	23.81	23.45
$13C_2H_6 \rightarrow [C_{12}H_{22}] + 14CH_4$	25.31	28.51	28.44	27.96
$12C_2H_6 \rightarrow [C_{12}H_{24}]^1 + 12CH_4$	22.96	25.96	25.74	25.41
$12C_2H_6 \rightarrow [C_{10}H_{16}] + 14CH_4$	25.62	28.48	27.46	27.56
$14C_2H_6 \rightarrow [C_{12}H_{20}]^1 + 16CH_4$	36.27	40.52	39.43	39.56
$14C_2H_6 \rightarrow [C_{12}H_{20}]^2 + 16CH_4$	29.43	32.75	32.34	32.42
$16C_2H_6 \rightarrow [C_{14}H_{24}] + 18CH_4$	46.85	52.39	51.25	51.44
$18C_2H_6 \rightarrow [C_{14}H_{20}] + 22CH_4$	43.81	48.64	46.90	47.06
MAE	1.70	0.39	0.14	

[a]CCSD(T)/cbs, see [89]

(MP2), spin-component-scaled MP2 (SCS-MP2) [114], the Becke–Lee–Yang-Parr three parameter hybrid functional (B3LYP), the Minnesota 2006 functional without and with Grimmes dispersion correction (M06 and M06-D3) and the double hybrid

Table 10 NBPRC database: oligomerisations and H_2 fragmentations of NH_3/BH_3 systems/H_2 activation reactions with PH3/BH3 systems

System	RPA	SOSEX	RPAX2	Ref.[a]
$NH_3 + BH_3 \rightarrow NH_3BH_3$	−28.96	−30.32	−30.61	−31.84
$NH_3 + BH_3 \rightarrow NH_2BH_2 + H_2$	−1.20	0.49	−0.32	−0.06
$NH_2BH_2 \rightarrow NHBH + H_2$	36.56	40.62	37.96	37.27
$2NH_2BH_2 \rightarrow$ cyc-$(BH_2NH_2)_2$	−17.72	−18.81	−18.07	−18.94
$3NH_2BH_2 \rightarrow B_3N_3H_6 + 3H_2$	−45.84	−48.33	−47.42	−48.27
$3NH_2BH_2 \rightarrow B_3N_3H_{12}$	−42.27	−44.69	−43.58	−45.06
$BH_3 + PH_3 \rightarrow BH_3PH_3$	−21.01	−23.11	−23.26	−24.40
$BH_3PH_3 + H_2 \rightarrow [PH_4BH_4]$[b]	40.73	42.19	41.39	40.40
$BF_3 + P(CH_3)_3 \rightarrow BF_3P(CH_3)_3$	−13.55	−13.87	−14.44	−14.60
$BF3P(CH_3)_3 + H_2 \rightarrow [P(CH_3)_3H\text{-}BF_3H]$[b]	19.99	19.15	19.19	18.80
$P(CH_3)_3 \rightarrow BCl_3P(CH_3)_3 + BCl_3$	−28.65	−30.98	−30.78	−30.80
$BCl_3P(CH_3)_3 + H_2 \rightarrow [P(CH_3)_3HBCl_3H]$[b]	18.59	17.17	17.74	17.60
MAE	1.69	0.89	0.68	

Energies in kcal/mol
[a]CCSD(T)/cbs, see [90, 91]
[b]Denotes a molecular transition state

functional B2PLYP-D3 by Grimme [113], including the D3 long-range dispersion correction [91]. An inspection of the diagrams in Fig. 24 shows that, with few exceptions, the RPA methods on average describe the different chemical reactions in the databases better than corresponding second-order perturbation theory methods and the selection of DFT methods. This is particularly true if the (exchange-)RPA errors are compared to those from the standard B3LYP hybrid functional, which strongly fails to describe the Diels–Alder reactions, the bond-separation reactions of the BSR36 database and the reactions involving N/B/P-compounds in the NBPRC database with reasonable precision; see Fig. 24. One can observe that the more modern M06 and M06-D3 functionals, which contain a large number of empirical parameters, in most cases outperform the B3LYP method for the databases considered, but even compared to these methods the RPA and in particular the RPAX2 method yield smaller errors on average.

The MP2, SCS-MP2 and B2PLYP-D3 methods all contain explicit second-order perturbation theory corrections to the correlation energy, i.e. involve the terms displayed in Fig. 15. While the MP2 method accounts for both terms in second-order completely (but misses all higher order terms compared to RPA methods), in the SCS-MP2 method these two terms are scaled with empirical prefactors with the idea of improving the description of parallel and antiparallel spin-component contributions to the correlation energy [114] (note that the antiparallel and parallel spin-component parameters p_S and p_T can be easily transformed into corresponding prefactors for the direct and exchange term of Fig. 15, yielding a factor of $(Ps + P_T)/2$ for the direct term and p_T for the exchange term). The drawback of this correction approach is, however, that the SCS-MP2 method is not self-correlation free in second-order because the direct and exchange contributions have different

Table 11 BH76RC database: reaction energies in kcal/mol

System	RPA	SOSEX	RPAX2	Ref.[a]
$H + N_2O \rightarrow OH + N_2$	−68.56	−71.65	−67.27	−65.08
$H + CH_3F \rightarrow HF + CH_3$	−26.05	−28.45	−27.49	−26.64
$H + F_2 \rightarrow HF + F$	−102.90	−114.05	−104.98	−103.91
$CH_3 + ClF \rightarrow CH_3F + Cl$	−55.25	−55.73	−53.21	−52.74
$F^- + CH_3Cl \rightarrow Cl^- + CH_3F$	−31.31	−36.08	−32.92	−32.65
$[FCH_3Cl]^b \rightarrow [FCH_3Cl]^{b,c}$	−26.17	−31.21	−27.98	−26.73
$OH^- + CH_3F \rightarrow CH_3OH + F^-$	−18.86	−20.79	−19.25	−20.11
$[HOCH_3F]^b \rightarrow [HOCH_3F]^{b,c}$	−34.76	−37.92	−35.81	−36.24
$H + N_2 \rightarrow HN_2$	3.03	4.25	3.99	3.97
$H + CO \rightarrow HCO$	−19.46	−17.25	−18.37	−19.51
$H + C_2H_4 \rightarrow C_2H_5$	−38.87	−43.02	−40.68	−40.03
$CH_3 + C_2H_4 \rightarrow C_3H_7$	−24.37	−28.14	−26.03	−26.12
$HNC \rightarrow HCN$	−14.33	−12.48	−13.92	−15.05
$H + HCl \rightarrow H_2 + Cl$	−8.29	−1.82	−3.03	−3.00
$OH + H_2 \rightarrow H_2O + H$	−11.52	−15.00	−15.04	−16.10
$CH_3 + H_2 \rightarrow CH_4 + H$	−1.61	−2.60	−2.24	−3.20
$OH + CH_4 \rightarrow H_2O + CH_3$	−9.91	−12.40	−12.79	−12.90
$OH + NH_3 \rightarrow H_2O + NH_2$	−8.69	−9.41	−10.08	−9.50
$HCl + CH_3 \rightarrow Cl + CH_4$	−9.90	−4.42	−5.27	−6.20
$OH + C_2H_6 \rightarrow H_2O + C_2H_5$	−12.84	−15.83	−16.33	−16.50
$F + H_2 \rightarrow HF + H$	−24.58	−29.93	−30.72	−31.60
$O + CH_4 \rightarrow OH + CH_3$	6.67	7.20	6.16	5.60
$H + PH_3 \rightarrow H_2 + PH_2$	−23.45	−22.04	−22.54	−20.10
$H + OH \rightarrow H_2 + O$	−5.06	−4.60	−3.92	−2.40
$H + H_2S \rightarrow H_2 + HS$	−16.91	−13.60	−14.38	−13.80
$O + HCl \rightarrow OH + Cl$	−3.23	2.78	0.89	−0.60
$NH_2 + CH_3 \rightarrow NH + CH_4$	−13.27	−13.66	−13.47	−14.40
$NH_2 + C_2H_5 \rightarrow NH + C_2H_6$	−10.35	−10.23	−9.93	−10.80
$C_2H_6 + NH_2 \rightarrow C_2H_5 + NH_3$	−4.15	−6.42	−6.25	−7.00
$NH_2 + CH_4 \rightarrow NH_3 + CH_3$	−1.23	−2.99	−2.71	−3.30
MAE	2.19	2.03	0.83	

[a]W1 and theor. est., see [92, 93]
[b]Denotes a molecular transition state
[c]Denotes an alternate conformer to the respective educt

prefactors in the original parametrisation by Grimme. As a consequence, the EPV contributions of the first diagram in Fig. 15 are not properly corrected by the corresponding cancelling contributions of the second diagram in the figure; see Sect. 2 for a discussion of the importance of the self-correlation correction. The B2PLYP-D3 method does not lack this inconsistency, but involves the terms in Fig. 15 multiplied by an empirical prefactor and mixes them with a local GGA-type functional as in standard hybrid functionals (this method is therefore also referred to

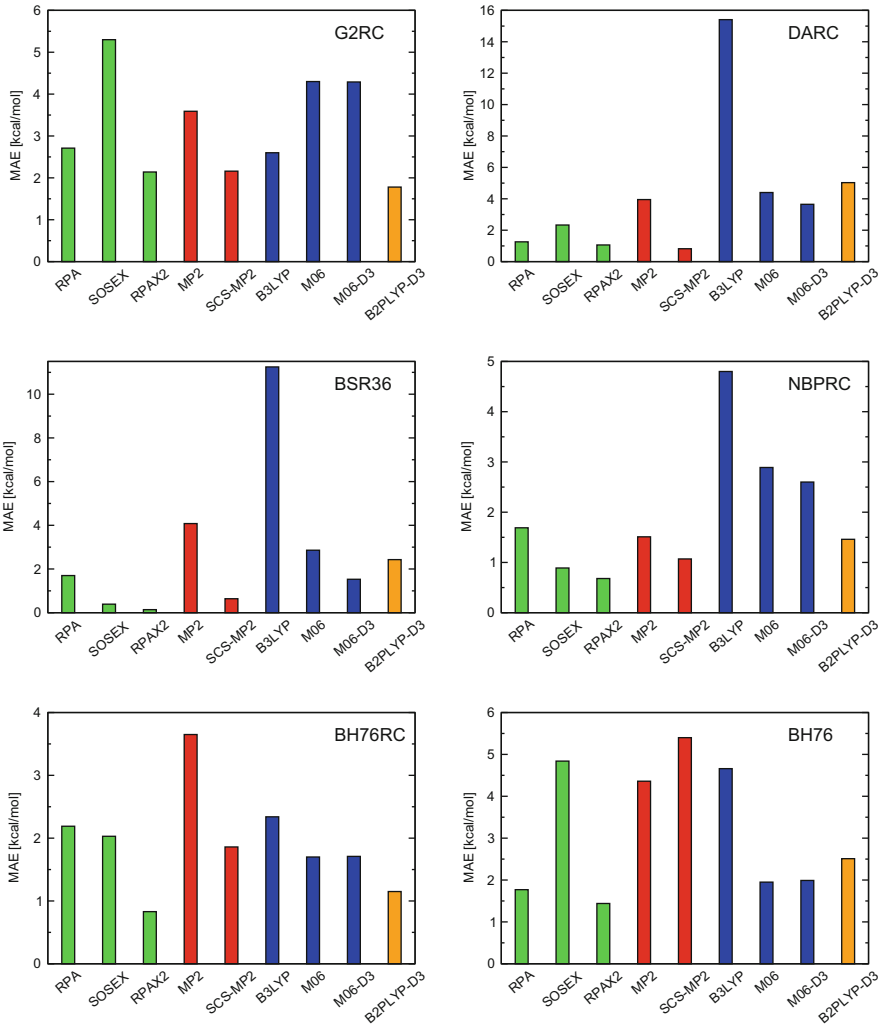


Fig. 24 Mean absolute errors of the different RPA methods compared to the corresponding errors of other wave-function and density-functional theory methods for the different databases studying reaction energies; see text. Also included in the figure is a diagram showing the MAEs for the barrier heights of the BH76 set; see Sect. 4.3

as double hybrid functional). While in this way the SCS-MP2 method can be seen as a correction approach to the MP2 method, since it is based on Hartree–Fock, the B2PLYP method is more akin to a DFT method which tries to capture long-range correlation effects missing in standard (hybrid-)GGA functionals.

A comparison of the performance of these second-order perturbation theory-based methods to the RPA methods for the chemical reactions considered shows that, with few exceptions, the latter are clearly performing better. This means that

obviously the inclusion of higher than second-order correlation energy contributions plays a significant role in the accurate prediction of chemical reaction energies. It also shows that methods involving empirical parameters such as the SCS-MP2 method and the B2PLYP method cannot be considered to be as reliable as the parameter-free RPA methods for the thermodynamic properties considered here. So the advantage, if any, of these empirical perturbative methods lies in the fact that, at least if compared to SOSEX or RPAX2, they are computationally slightly more efficient, because no amplitude equations need to be solved for calculating the second-order correlation energy terms. So one might assume that these methods become more attractive than RPA methods for larger systems. Note, however, that, for very large chemical systems, another point needs to be taken care of if perturbative methods are applied, namely small or even vanishing band gaps of these systems leading to strong non-dynamical correlation effects, which can even lead to a breakdown of perturbation theory methods.

In summary, it can be concluded that the RPA, SOSEX and RPAX2 methods describe chemical reaction energies quite accurately if compared to other commonly used wave-function and DFT methods. This is especially true for the RPAX2 method which, not unexpectedly, performs best among the three RPA methods and yields errors which typically lie in the range of 1–2 kcal/mol for all chemical reactions considered here; see Fig. 24.

4.3 Barrier Heights

In this section, the performance of the RPA, SOSEX and RPAX2 methods for describing barrier heights of chemical reactions is analysed, i.e. compared to the analysis of Sect. 4.2 the energy differences between the educts/products and a corresponding transition state are examined. Since the Hartree–Fock method does not reliably predict barriers for chemical reactions and yields average and maximum errors of typically 10 and 30 kcal/mol [103], respectively, an adequate description of electron correlation effects is of great importance for describing barrier heights for chemical reactions. Standard GGA-type DFT methods are able to describe these correlation effects to some extent, but often underestimate barrier heights on magnitude [92, 93, 103]. The only satisfactory description of this property seems to require at least hybrid-type functionals [93, 103] which correct parts of the self-interaction errors of the Coulomb energy in first order; see Fig. 9. However, the performances of different hybrid functionals for reaction barriers are very varied and it is difficult to choose between the existing functionals. For example, the most popular B3LYP hybrid-functional significantly underestimates barrier heights [93] and it appears to be necessary to derive functionals which are explicitly parametrised against this property to achieve a satisfying accuracy. The first such functional developed was the MPW1K functional [115], which performs much better than ‘older’ hybrid-functionals for thermokinetic properties.

Table 12 BH76 database: barrier heights of hydrogen transfer, heavy atom transfer, nucleophilic substitution, unimolecular, and association reactions

System	RPA		SOSEX		RPAX2		Ref.[a]	
$H + N_2O \rightarrow [N_2OH]^b \leftarrow OH + N_2$	15.61	84.16	21.73	93.38	18.75	86.02	18.14	83.22
$H + HF \rightarrow [HFH]^b$	38.78		45.67		42.88		42.18	
$H + HCl \rightarrow [HClH]^b$	17.77		20.98		19.13		18.00	
$H + CH_3F \rightarrow [HFCH_3]^b \rightarrow HF + CH_3$	27.40	53.45	35.09	63.54	31.00	58.49	30.38	57.02
$H + F_2 \rightarrow [HF_2]^b \leftarrow HF + F$	0.12	103.02	8.86	122.90	5.09	110.08	2.27	106.18
$CH_3 + ClF \rightarrow [CH_3FCl]^b \leftarrow CH_3F + Cl$	2.37	57.62	13.51	69.24	8.59	61.79	7.43	60.17
$F^- + CH_3F \rightarrow [FCH_3F]^b \leftarrow [FCH_3F]$	−1.95	10.98	3.60	16.19	−0.07	12.72	−0.34	13.28
$Cl^- + CH_3Cl \rightarrow [ClCH_3Cl]^b \leftarrow [ClCH_3Cl]$	1.11	11.15	5.12	15.06	1.85	12.35	3.10	13.61
$F^- + CH_3Cl \rightarrow [FCH_3Cl]^b \leftarrow Cl^- + CH_3F$	−13.12	18.19	−10.11	25.97	−12.45	20.46	−12.54	20.11
$[FCH_3Cl]^1 \rightarrow [FCH_3Cl]^b \leftarrow [FCH_3Cl]^2$	1.32	27.49	3.79	34.99	2.08	30.06	2.89	29.62
$OH^- + CH_3F \rightarrow [HOCH_3F]^b \leftarrow CH_3OH + F^-$	−3.36	15.50	1.63	22.42	−1.19	18.07	−2.78	17.33
$[HOCH_3F]^2 \rightarrow [HOCH_3F]^b \leftarrow [HOCH_3F]^1$	9.42	44.17	14.36	52.28	11.20	47.02	10.96	47.20
$H + N_2 \rightarrow [HN_2]^b \leftarrow HN_2$	14.00	10.97	17.41	13.16	15.58	11.59	14.69	10.72
$H + CO \rightarrow [HCO]^b \leftarrow HCO$	4.27	23.73	6.27	23.52	4.39	22.76	3.17	22.68
$H + C_2H_4 \rightarrow [C_2H_5]^b \rightarrow C_2H_5$	3.12	41.99	3.41	46.42	2.61	43.29	1.72	41.75
$CH_3 + C_2H_4 \rightarrow [C_3H_7]^b \leftarrow C_3H_7$	7.86	32.23	9.05	37.19	7.67	33.70	6.85	32.97
$HCN \rightarrow [HCN]^b \leftarrow HNC$	46.28	31.95	46.23	33.76	46.74	32.82	48.16	33.11
$H + HCl \rightarrow [H_2Cl]^b \leftarrow H_2 + Cl$	3.38	11.67	8.36	10.18	6.30	9.33	5.70	8.70
$OH + H_2 \rightarrow [H_3O]^b \leftarrow H_2O + H$	6.25	17.77	10.79	25.80	7.46	22.49	5.10	21.20
$CH_3 + H_2 \rightarrow [CH_5]^b \leftarrow CH_4 + H$	12.73	14.34	14.42	17.02	13.18	15.43	12.10	15.30
$OH + CH_4 \rightarrow [CH_5\ O]^b \leftarrow H_2O + CH_3$	6.80	16.71	12.31	24.71	8.15	20.94	6.70	19.60
$OH + NH_3 \rightarrow [NH_4O]^b \leftarrow H_2O + NH_2$	0.72	9.40	14.23	23.64	6.49	16.57	3.20	12.70
$HCl + CH_3 \rightarrow [CH_4Cl]^b \leftarrow Cl + CH_4$	1.16	11.05	4.42	8.85	2.43	7.70	1.70	7.90
$OH + C_2H_6 \rightarrow [C_2H_7O]^b \leftarrow H_2O + C_2H_5$	4.59	17.43	10.14	25.97	5.68	22.01	3.40	19.90
$F + H_2 \rightarrow [FH_2]^b \leftarrow HF + H$	2.34	26.91	10.42	40.35	5.32	36.04	1.80	33.40

$O + CH_4 \rightarrow [CH_4O]^b \leftarrow OH + CH_3$	14.26	7.59	20.42	13.21	16.23	10.06	13.70	8.10
$H + PH_3 \rightarrow [PH_4]^b \leftarrow H_2 + PH_2$	3.25	26.70	4.92	26.96	3.44	25.98	3.10	23.20
$H + OH \rightarrow [OH_2]^b \leftarrow H_2 + O$	8.58	13.63	14.30	18.90	11.54	15.46	10.70	13.10
$H + H_2S \rightarrow [SH_3]^b \leftarrow H_2 + HS$	3.81	20.71	5.76	19.36	4.27	18.64	3.50	17.30
$O + HCl \rightarrow [OHCl]^b \leftarrow OH + Cl$	6.56	9.79	28.18	25.40	14.20	13.31	9.80	10.40
$NH_2 + CH_3 \rightarrow [NH_2CH_3]^b \leftarrow NH + CH_4$	8.74	22.02	12.95	26.61	10.08	23.54	8.00	22.40
$NH_2 + C_2H_5 \rightarrow [NH_2C_2H_5]^b \leftarrow NH + C_2H_6$	9.60	19.95	13.67	23.90	10.93	20.86	7.50	18.30
$C_2H_6 + NH_2 \rightarrow [C_2H_6NH_2]^b \leftarrow C_2H_5 + NH_3$	12.33	16.48	15.85	22.27	12.84	19.09	10.40	17.40
$NH_2 + CH_4 \rightarrow [NH_2CH_4]^b$	14.46		18.43		15.47		14.50	
$NH_3 + CH_3 \rightarrow [NH_3CH_3]^b$	15.69		21.42		18.18		17.80	
$C_5H_8 \rightarrow [HC_5H_7]^b$	39.65		38.00		38.91		38.40	
MAE		1.77		4.84		1.44		

Energies are in kcal/mol

[a]W1 and theor. est., see [92, 93]

[b]Denotes a molecular transition state

Here we examine whether the parameter-free RPA methods are able to describe reaction barriers with a good accuracy. For this, the reactions of the BH76 database [92, 93] are considered, which are compiled in Table 12. This table contains both the activation energies for the forward and reverse reactions, so each column in the table has two subcolumns (except for the reactions where the educts and products are identical). The last line in the table again shows the MAEs for the three RPA methods compared to the reference values of the last column. As with some chemical reaction databases (see Sect. 4.2), one can see that the description of electron exchange effects through the SOSEX method does not lead to an improvement over the RPA barrier heights. On the contrary, the average error for the SOSEX barrier heights is even more than twice as large as the RPA ones. This is because of a strong overestimation of the barrier heights, most notably found for the hydrogen transfer reaction in the $O \cdots H \cdots Cl$ system with deviations of 18 and 15 kcal/mol from the reference barrier heights, respectively. It is difficult to explain what the origin of this poor performance of the SOSEX method for this property might be, yet, considering the reasonable description of the corresponding underlying reaction energies of the BH76RC set (see Table 11), one can narrow the failure of the SOSEX method to a poor description of the transition states, which are obviously too weakly bound with SOSEX. Compared to this, the RPAX2 method yields barrier heights much closer to the reference values with a mean absolute deviation of 1.4 kcal/mol, i.e. even slightly smaller than for RPA; see Table 12.

A comparison of the MAEs of the RPA methods to those of other wave-function and DFT methods (see Fig. 24) shows that, just as with the SOSEX method, the MP2, SCS-MP2 and B3LYP methods are also unable to predict barrier heights with reasonable accuracy. In contrast to this, more modern density-functionals such as the M06 and B2PLYP-D3 functional have smaller average errors of 2–2.5 kcal/mol, respectively. The best performance for barrier heights of the different methods considered is again achieved by the RPAX2 method.

5 Summary

Three KS-RPA correlation methods were considered in this work for studying a number of thermochemical properties. It was found that, compared to a range of commonly used second-order perturbation theory-based wave-function and DFT methods, these RPA methods often yield a higher accuracy for chemical reaction energies. In the case of the direct RPA method, excluding any high-order exchange interactions, this is a remarkable result when considering the fact that this method was originally developed for describing electron–electron interactions in the homogeneous electron gas in the high density limit. While this system is certainly not a good model system of a molecule where the electrons are confined in an external potential stemming from the nuclei, the question arises as to why the RPA can also be successfully applied to molecular systems? It was argued in this work that one reason for the good performance of KS-RPA methods is the fact that the Kohn–

Sham determinant is close to the Brueckner determinant and, thus, to some extent, includes single-excitation contributions which are missing in Hartree–Fock-based RPA methods. Kohn–Sham RPA methods can thus be regarded as approximate Brueckner coupled-cluster doubles methods.

A main advantage of RPA methods over common DFT methods and empirical wave-function methods such as the SCS-MP2 method [114] is the fact that they are free from any empirical parameters and that they can systematically be improved. The latter was demonstrated in this work by using, in addition to the standard direct RPA method, two exchange RPA methods, which account for electron exchange effects in second and higher orders, namely the SOSEX [43] and the RPAX2" method [45]. It has been observed that, with these methods, the description of thermodynamic properties is often more accurate than with the direct RPA method, which possesses a self-correlation error in second and higher orders. Furthermore, a comparison of the performance of SOSEX and RPAX2 themselves shows a clear advantage of RPAX2 over the SOSEX method. This again can be attributed to the fact that the RPAX2 method is more complete in each order of perturbation theory than the SOSEX method; see Sect. 2. In fact, it has been observed that the RPAX2 method yields only very small errors for atomisation energies, ionisation energies and chemical reaction energies, the latter being even more accurate than those from the CCSD method if compared to CCSD(T) reference reaction energies [45]. This indicates that electron correlation effects in third and higher orders, which are missing in RPAX2, are either small or tend to cancel each other in a complete summation of perturbation theory terms contributing to the electron correlation energy.

In summary, the combination of RPA methods with Kohn–Sham reference states leads to attractive alternative methods for studying thermochemical properties of molecular systems.

References

1. Powell CJ, Swan JB (1959) Phys Rev 115:869
2. Pines D, Bohm D (1952) Phys Rev 85:338
3. Mattuck RD (1992) A guide to Feynman diagrams in the many-body problem. Dover, Mineola
4. Bohm D, Pines D (1951) Phys Rev 82:625
5. Bohm D, Pines D (1953) Phys Rev 92:609
6. Pines D (1953) Phys Rev 92:636
7. Gell-Mann M, Brueckner KA (1957) Phys Rev 106:364
8. Sawada K, Brueckner KA, Fukuda N (1957) Phys Rev 108:507
9. Hubbard J (1957) Proc R Soc Lond A240:539
10. Hubbard J (1958) Proc Roy Soc London A243:336
11. Furche F (2001) Phys Rev B 64:195120
12. Angyan J, Liu RF, Toulouse J, Jansen G (2011) J Chem Theory Comput 7:3116
13. Heßelmann A, Görling A (2011) Mol Phys 109:2473
14. Eshuis H, Bates JE, Furche F (2012) Theor Chem Acc 131:1084
15. Ren X, Rinke P, Joas C, Scheffler M (2012) J Mater Sci 47:7447

16. Pines D, Nozieres P (1966) The theory of quantum liquids: normal Fermi liquids, vol I. W.A. Benjamin, New York/Amsterdam
17. March NH, Young WH, Sampanthar S (1995) The many-body problem in quantum mechanics. Dover, New York
18. MacLachlan AD, Ball MA (1964) Rev Mod Phys 36:844
19. Ball MA, McLachlan AD (1964) Mol Phys 7:501
20. Ehrenreich H, Cohen MH (1959) Phys Rev 115:786
21. Dunning TH, McKoy V (1967) J Chem Phys 47:1735
22. Linderberg J, Jørgensen P, Oddershede J, Ratner M (1972) J Chem Phys 56:6213
23. Jørgensen P, Oddershede J, Ratner M (1974) J Chem Phys 61:710
24. Oddershede J (1978) Adv Quantum Chem 11:275
25. Szabo A, Ostlund NS (1977) J Chem Phys 67:4351
26. Szabo A, Ostlund NS (1977) Int J Quantum Chem S11:389
27. McWeeny R (1996) Methods of molecular quantum mechanics. Academic, London
28. Heßelmann A (2011) J Chem Phys 134:204107
29. Callen HB, Welton TA (1951) Phys Rev 83:34
30. Langreth DC, Perdew JP (1975) Solid State Commun 17:1425
31. Langreth DC, Perdew JP (1977) Phys Rev B 15:2884
32. Scuseria GE, Henderson TM, Sorensen DC (2008) J Chem Phys 129:231101
33. Jansen G, Liu RF, Angyan JG (2010) J Chem Phys 133:154106
34. Eshuis H, Furche F (2011) J Phys Chem Lett 2:983
35. Lu D, Li Y, Rocca D, Galli G (2009) Phys Rev Lett 102:206411
36. Fuchs M, Niquet YM, Gonze X, Burke K (2005) J Chem Phys 122:094116
37. Henderson TM, Scuseria GE (2010) Mol Phys 108:2511
38. Heßelmann A, Görling A (2011) Phys Rev Lett 106:093001
39. Mori-Sanchez P, Cohen AJ, Yang W (2012) Phys Rev A 85:042507
40. Heßelmann A, Görling A (2013) J Chem Theory Comput 9:4382
41. Jiang H, Engel E (2007) J Chem Phys 127:184108
42. Toulouse J, Gerber IC, Jansen G, Savin A, Angyan JG (2009) Phys Rev Lett 102:096404
43. Grüneis A, Marsman M, Harl J, Schimka L, Kresse G (2009) J Chem Phys 131:154115
44. Heßelmann A, Görling A (2010) Mol Phys 108:359
45. Heßelmann A (2012) Phys Rev A 85:012517
46. Gould T (2012) J Chem Phys 137:111101
47. Eshuis H, Yarkony J, Furche F (2010) J Chem Phys 132:234114
48. Nguyen HV, de Gironcoli S (2009) Phys Rev B 79:205114
49. Furche F, Van Voorhis T (2005) J Chem Phys 122:164106
50. Fabiano E, Sala FD (2012) Theor Chem Acc 131:1278
51. Eshuis H, Furche F (2012) J Chem Phys 136:084105
52. Toulouse J, Colonna F, Savin A (2004) Phys Rev A 70:062505
53. Janesko BG, Henderson TM, Scuseria GE (2009) J Phys Chem 130:081105
54. Toulouse J, Zhu W, Angyan JG, Savin A (2010) Phys Rev A 82:032502
55. Janesko BG, Henderson TM, Scuseria GE (2009) J Phys Chem 131:034110; Erratum: J Phys Chem 133 (2010) 179901
56. Harris FE, Monkhorst HJ, Freeman DL (1992) Algebraic and diagrammatic methods in many-fermion theory. Oxford University Press, New York
57. Lindgren I (1986) Atomic many–body theory. Springer, Berlin
58. Shavitt I, Bartlett RJ (2009) Many-body methods in chemistry and physics: MBPT and coupled–cluster theory. Cambridge University Press, Cambridge
59. Kucharski SA, Bartlett RJ (1986) Adv Quantum Chem 18:281
60. Feynman RP (1949) Phys Rev 76:769
61. Szabo A, Ostlund NS (1996) Modern quantum chemistry. Dover, Mineola
62. Negele JW, Orland H (1998) Quantum many-particle systems. Perseus Publishing, Cambridge
63. Hubac I (1980) Int J Quantum Chem 17:195
64. Thouless DJ (1960) Nucl Phys 21:225

65. Hellgreen M, von Barth U (2010) J Chem Phys 132:044101
66. Ipatov A, Heßelmann A, Görling A (2010) Int J Quantum Chem 110:2202
67. Sharp RT, Horton GK (1953) Phys Rev 90:317
68. Talman JD, Shadwick WF (1976) Phys Rev A 14:36
69. Grabo T, Kreibich T, Kurth S, Gross EKU (1998) Strong coulomb correlations in electronic structure: beyond the local density approximation. Gordon & Breach, Tokyo, p 203, chap. 4
70. Engel E, Dreizler RM (1999) J Comput Chem 20:31
71. Görling A (1999) Phys Rev Lett 83:5459
72. Ivanov S, Hirata S, Bartlett RJ (1999) Phys Rev Lett 83:5455
73. Gimon T, Ipatov A, Heßelmann A, Görling A (2009) J Chem Theory Comput 5:781
74. Sherrill CD, Krylov AI, Byrd EFC, Head-Gordon M (1998) J Chem Phys 109:4171
75. Lindgren I, Salomonson S (2002) Int J Quantum Chem 90:294
76. Kollmar C, Heßelmann A (2010) Theor Chem Acc 127:311
77. Hellgren M, Rohr DR, Gross EKU (2012) J Chem Phys 136:034106
78. Verma P, Bartlett RJ (2012) J Chem Phys 136:044105
79. Bleiziffer P, Heßelmann A, Görling A (2013) J Chem Phys 139:084113
80. Heßelmann A (2005) J Chem Phys 122:244108
81. Tajti A, Szalay PG, Csaszar AG, Kallay M, Gauss J (2004) J Chem Phys 121:11599
82. Harding ME, Vazquez J, Ruscic B, Wilson AK, Gauss J, Stanton JF (2008) J Chem Phys 128:114111
83. Curtiss LA, Raghavachari K, Redfern PC, Rassolov V, Pople JA (1998) J Chem Phys 109:7764
84. Lynch BJ, Zhao Y, Truhlar DG (2003) J Phys Chem A 107:1384
85. Zhao Y, Truhlar DG (2006) J Phys Chem A 110:10478
86. Zhao Y, Truhlar DG (2006) J Chem Phys 125:194101
87. Curtiss LA, Raghavachari K, Redfern PC, Pople JA (1997) J Chem Phys 106:1063
88. Johnson ER, Mori-Sanchez P, Cohen AJ, Yang W (2008) J Chem Phys 129:204112
89. Steinmann SN, Csonka G, Carminboeuf C (2009) J Chem Theory Comput 5:2950
90. Grimme S, Djukic JP (2010) Inorg Chem 49:2911
91. Grimme S, Antony J, Ehrlich S, Krieg H (2010) J Chem Phys 132:154104
92. Zhao Y, Lynch BJ, Truhlar DG (2004) J Phys Chem A 108:2715
93. Zhao Y, Gonzalez-Garcia N, Truhlar DG (2005) J Phys Chem A 109:2012
94. Peverati R, Truhlar DG (2014) The quest for a universal density functional: the accuracy of density functionals across a broad spectrum of databases in chemistry and physics. Phil Trans R Soc A 372:20120476/1-51 (part of a theme issue on "DFT for Physics, Chemistry and Biology," guest edited by T. van Mourik, M. Buehl, and M.-P. Gaigeot)
95. Goerigk L, Grimme S (2011) Phys Chem Chem Phys 13:6670
96. Goerigk L, Grimme S (2011) J Chem Theor Comput 7:291
97. Grimme S (2011) GMTKN30: a database for general main group thermochemistry, kinetics, and non-covalent interactions. http://toc.unimuenster.de/GMTKN/GMTKN30/GMTKN30main.html
98. Perdew JP, Burke K, Ernzerhof M (1996) Phys Rev Lett 77:3865
99. Bak KL, Jørgensen P, Olsen J, Helgaker T, Klopper W (2000) J Chem Phys 112:9229
100. Weigend F, Köhn A, Hättig C (2002) J Chem Phys 116:3175
101. Werner HJ, Knowles PJ, Lindh R, Manby FR, Schüz M et al. (2010) Molpro, version 2010.2, a package of ab initio programs. http://www.molpro.net
102. Helgaker T, Jørgensen P, Olsen J (2012) Molecular electronic–structure theory. Wiley, Chichester
103. Levine IN (2009) Quantum chemistry, 6th edn. Pearson Prentice Hall, New Jersey
104. Cramer CJ (2008) Essentials of computational chemistry, 2nd edn. Wiley, New York
105. Tao J, Perdew JP, Staroverov VN, Scuseria GE (2003) Phys Rev Lett 91:146401
106. Bakowies D (2013) J Phys Chem A 117:228
107. Gritsenko O, Baerends EJ (2004) J Chem Phys 121:655
108. Goerigk L, Grimme S (2010) J Chem Theory Comput 6:107
109. Becke AD (1993) J Chem Phys 98:1372
110. Gilbert TM (2004) J Phys Chem A 108:2550

111. Martin JML, Oliveira GD (1999) J Chem Phys 111:1843
112. Zhao Y, Truhlar DG (2006) Theor Chem Acc 120:215
113. Grimme S (2006) J Chem Phys 124:034108
114. Grimme S (2003) J Chem Phys 118:9095
115. Lynch BJ, Fast PL, Harris M, Truhlar DG (2000) J Phys Chem A104:4811
116. Rocca D (2014) J Chem Phys 140:18A501
117. Umrigar CJ, Gonze X (1994) Phys Rev A 50:3827

Top Curr Chem (2015) 365: 145–174
DOI: 10.1007/128_2014_619

Published online: 21 April 2015

Non-analytic Spin-Density Functionals

Martín A. Mosquera and Adam Wasserman

Abstract We examine the integer discontinuity (or derivative discontinuity) of the exact energy functionals of Kohn–Sham density-functional theory for the spin-polarized case. The integer discontinuity and its cause, the piecewise linearity of the energy in the grand canonical ensemble, are required to improve the predictive power of density-functional approximations to the exchange-correlation energy. We show how any spin-polarized DFA can be adapted to display the proper integer discontinuity. The formalism we present here can be used to improve functionals further within spin density-functional theory and fragment-based formulations of DFT.

Keywords Derivative discontinuity · Electronic spin · Ensemble · Magnetic field · Molecular dissociation

Contents

M.A. Mosquera and A. Wasserman (✉)
Department of Chemistry, Purdue University, West Lafayette, IN 47907, USA
e-mail: awasser@purdue.edu

1 Introduction

Density-functional theory (DFT) is employed by scientists to calculate properties of molecules and materials. The ground-state energy of a system of electrons is used to predict ground-state geometries, spectra, and forces involved in the dynamics of molecules. DFT establishes that the electronic density is sufficient in principle to calculate all ground-state properties. Nevertheless, approximations to the energy as a functional of the density are necessary in practice to carry out the calculations. Kohn–Sham theory [1] is the preferred paradigm to search for approximations to the energy, which is decomposed into several functionals: (1) the kinetic energy of a system of non-interacting electrons, (2) an energy functional accounting for a semi-classical repulsion (Hartree repulsion energy), (3) the external energy of the electrons, and (4) an unknown functional, the exchange-correlation (XC) energy. This latter functional, if properly approximated, leads to efficient computational ab initio methods for the calculation of ground-state molecular observables.

Density-functional approximations (DFAs) to the XC energy work well for atoms and molecules around their equilibrium geometry [2, 3], but challenges remain [4–6]. Molecular properties such as the energies involved in bond breaking (or formation) are not accurate enough to describe chemical reactions [2]. XC energy functionals are also needed to describe molecules involving transition metal atoms, where the coupling of magnetic moments is non-negligible [7]. Although the exact energy functional is unknown, its properties are accessible in some cases. For example, when two atoms are far apart but exchange electrons adiabatically [8], the change of energy in the charge-transfer process displays a piecewise linear dependency on the average number of electrons transferred, and the derivative of the energy is discontinuous at the point where the atoms are in their neutral state. Most known DFAs fail this simple discontinuity test [9].

The integer discontinuity is found by using the statistics of the grand canonical ensemble. In this picture, a molecule or atom exchanges electrons adiabatically with a reservoir. This exchange takes place if the work function of the metal is tuned to the change of energy when the quasi-isolated molecule loses or gains an electron [10]. In other words, the energy must be conserved in the exchange of electrons. The lack of integer discontinuity limits the accuracy of DFA calculations for chemical systems in general [9]. One can devise 1d systems that can be solved with standard numerical techniques to illustrate the integer discontinuity. Tempel

et al. [11] considered a system of two 1d electrons interacting via a soft-coulomb potential that are subject to a double-well potential, also described by soft-coulomb electron-nucleus interaction potentials (this model follows a similar one studied earlier by Perdew [10]). By varying the distance, they noticed a step and a peak in the potential. The height of the step is the difference between the ionization energies of the atoms. The step is caused by the ionization theorem, stating that the highest occupied orbital energy is the negative of the lowest ionization energy. The peak is produced by the tendency of the atoms to neutralize their charges when they are separated by a large distance. Because standard XC functionals miss this peak, the atoms render fractional charges instead.

The need for better spin-resolved electronic structure methods for solids and molecules is justified by applications in engineering, where tools to design and model ferromagnetic materials are needed. Devices based on magnetic materials are useful for health, information, communication, and computer technologies. Molecules and solids can be found in states of high spin-multiplicity (especially during chemical reactions), not just doublet and singlet states (for example, ferromagnetism in molecular chemistry [12, 13]). For development of quantum computers, molecules with high spin moments also need to be modeled [14].

SDFT is the de facto method for molecular calculations, yet most studies on the derivative discontinuity focus on the theory of functionals of the total electronic density. On the other hand, for dissociation of molecules, spin-polarized DFAs fail to yield products in their neutral state; in addition, they predict the products to be in a mixture of spin states, giving rise to non-natural spin moments on each fragment. The motivation for the piecewise linearity of the energy also derives from lattices when considering their unit-cell with its molecules or atoms at large separations. The electrons in these systems can hop from one subsystem to another. Projection of the wavefunction onto a state where the electrons are localized gives the probability amplitude that the system is found on such a state. In the case of a gas, we could assume that the collisions between the molecules change their spin-momentum in a random fashion, such that the grand-canonical ensemble prescription applies to each atom or molecule.

In this work we consider the ensemble formulation of SDFT. We show how the exact energy functional with integer discontinuity can be constructed. The background formalism to define the spin-polarized density functionals of this work is presented. A formula for the XC energy functional in the ensemble is shown, which follows from work developed by us in [15] which introduces the integer discontinuity to the exact XC energy functional and permits extension of any DFA to feature the integer discontinuity. We also extend the Levy constrained search to the zero-temperature ensemble; our definition does not use density matrices but wave-functions over the Fock space. Finally, we discuss how partition spin-density-functional theory [16] can be used to satisfy the constraints of molecular dissociation.

2 Definitions

The state of a system of non-relativistic electrons is determined by a normalizable wave-function that decays rapidly in the asymptotics. If the number of electrons is a positive natural number J, then the wave-function depends on $4J$ Cartesian coordinates: three spatial and one spin coordinate per electron. The wavefunction is a member of the Hilbert space of antisymmetric wave-functions that are square integrable, and vanishes in the asymptotic region. For convenience we employ spaces of kets throughout this work; we denote the Hilbert space of kets corresponding to J electrons as $\mathcal{H}_J$, and the symbol $|\psi\rangle$ represents a member of $\mathcal{H}_J$.

Most of this work concerns states with a number of particles in the positive real line. Hence, we require the Fock space ($\boldsymbol{\mathcal{F}}$), formally defined as

$$\boldsymbol{\mathcal{F}} = \bigoplus_{j=0}^{\infty} \mathcal{H}_j,$$

where $\mathcal{H}_0$, for convenience, is the set of complex numbers, assigned to states with no electrons. The operators corresponding to the physical observables of the system are expressed in terms of creation $\left(\hat{\psi}^{\dagger}_{\sigma}(\mathbf{r})\right)$ and annihilation $\left(\hat{\psi}_{\sigma}(\mathbf{r})\right)$ operators; they satisfy the anticommutation relation $\left\{\hat{\psi}_{\sigma}(\mathbf{r}), \hat{\psi}^{\dagger}_{\sigma'}\left(\mathbf{r}'\right)\right\} = \delta_{\sigma\sigma'}\delta\left(\mathbf{r}' - \mathbf{r}\right)$, where σ denotes the spin-coordinate, $\uparrow$ or $\downarrow$ for electrons.

To complete the definition of our system of electrons, let us introduce the Hamiltonian:

$$\hat{H}_v = \hat{H}_0 + \int \mathrm{d}^3\mathrm{r}\, \hat{n}(\mathbf{r})\, v(\mathbf{r}). \tag{1}$$

$\hat{H}_0$ is the internal energy Hamiltonian. In the absence of spin-orbit and spin-spin couplings, $\hat{H}_0 = \hat{T} + \hat{W}$, where $\hat{T}$ and $\hat{W}$ are the kinetic energy and electron–electron repulsion operators, respectively. The spin-density operator is expressed as $\hat{n}_{\sigma}(\mathbf{r}) = \hat{\psi}^{\dagger}_{\sigma}(\mathbf{r})\hat{\psi}_{\sigma}(\mathbf{r})$, and the total density operator is $\hat{n}(\mathbf{r}) = \hat{n}_{\uparrow}(\mathbf{r}) + \hat{n}_{\downarrow}(\mathbf{r})$. The 1-body external potential v represents the interaction between a single electron and the nuclei (whose positions are fixed). The Hamiltonian operator is thus determined by the 1-body external potential.

Let M be a natural number (including 0) that runs over the number of particles, and $\mathbf{k}$ be a vector running over the quantum numbers of the system. Suppose $\{|\psi_{M,\mathbf{k}}\rangle\}$ is a basis set for the Hilbert space $\mathcal{H}_M$, for $M = 1, 2, \ldots$. Define the Liouville space $\mathcal{L}$ as the set of density matrices of the form

$$\hat{\Gamma} = \bigoplus_{M,\mathbf{k}} \omega_{M,\mathbf{k}} \left|\psi_{M,\mathbf{k}}\right\rangle \left\langle\psi_{M,\mathbf{k}}\right|,$$

$\omega_{M,\mathbf{k}} \geq 0$ and $\sum_{M,\mathbf{k}} \omega_{M,\mathbf{k}} = 1$. For a given basis set, a density matrix is completely determined by the weights $\{\omega_{M,\mathbf{k}}\}$. Let $\hat{N}_{\sigma} = \int \mathrm{d}^3\mathbf{r}\, \hat{n}_{\sigma}(\mathbf{r})$ be the operator

corresponding to the measurement of the number particles with spin σ. The operator for the total number of particles is

$$\hat{N} = \hat{N}_{\uparrow} + \hat{N}_{\downarrow}.$$

If two kets $|\psi_{M'}\rangle$ and $|\psi_M\rangle$ ($M, M' \in \mathbb{N}$ and $M \neq M'$) are eigenkets of the particle number operator ($\hat{N}|\psi_j\rangle = j|\psi_j\rangle, j = M, M'$), then they are orthogonal, $\langle\psi_{M'}|\psi_M\rangle = 0$. A ket $|\psi\rangle$ in the Fock space is expressed as a direct sum of kets belonging to different Hilbert spaces:

$$|\psi\rangle = c_0|0\rangle \oplus c_1|\psi_1\rangle \oplus c_2|\psi_2\rangle \oplus \cdots,$$

where the coefficients $\{c_i\}$ are complex.

3 Variational Principles

For a given number of electrons $J \in \mathbb{N}$, the ground-state energy and ket are obtained by infimizing the average energy of the system $\langle\psi|\hat{H}_v|\psi\rangle$ over the Hilbert space $\mathcal{H}_J$. Let e be the functional defined by

$$e[|\psi\rangle; v] = \langle\psi|\hat{H}_v|\psi\rangle,$$

$|\psi\rangle \in \mathcal{F}$. For systems with any number of particles $N \in \mathbb{R}$, the ground-state energy, E_0, reads

$$E_0(N; [v]) = \inf\{e[|\psi\rangle; v] \,|\, \langle\psi|\hat{N}|\psi\rangle = N,\ |\psi\rangle \in \mathcal{F}\}. \tag{2}$$

The Hamiltonian $\hat{H}_v$ is particle-conserving, $[\hat{N}, \hat{H}_v] = 0$, a property allowing us to write the functional $e[|\psi\rangle; v]$ as follows:

$$e[|\psi\rangle; v] = \sum_{M=1}^{\infty} \omega_M e[|\psi_M\rangle; v], \tag{3}$$

where $\omega_M = |c_M|^2$. Hence, the infimum in (2) is written as

$$E_0(N; [v]) = \inf\left\{\sum_{M\in\mathbb{N}} \omega_M E_0(M; [v]) \,\Big|\, \sum_{M=1}^{\infty} M\omega_M = N, \sum_{M=1}^{\infty} \omega_M = 1, 0 \leq \omega_M \leq 1\right\}, \tag{4}$$

where

$$E_0(M;[v]) = \inf\left\{e\left[\left|\psi_M\right\rangle;v\right] \middle| \left|\psi_M\right\rangle \in \mathcal{H}_M\right\}.$$

The above is the energy of a system with discrete number of electrons, i.e., the energy of an isolated molecule.

A density matrix $\hat{\Gamma} \in \mathcal{L}$ is a combination of subdensity matrices[1]:

$$\hat{\Gamma} = w_0\left|0\right\rangle\left\langle 0\right| + w_1\hat{\Gamma}_1 + w_2\hat{\Gamma}_2 + \cdots,$$

where

$$\hat{\Gamma}_M = \sum_{\mathbf{k}} w_{M,\mathbf{k}}\left|\psi_{M,\mathbf{k}}\right\rangle\left\langle\psi_{M,\mathbf{k}}\right|;\ \sum_{\mathbf{k}} w_{M,\mathbf{k}} = 1.$$

This corresponds to expressing $\mathcal{L} = \mathcal{L}_0 \oplus \mathcal{L}_1 \oplus \mathcal{L}_2 \oplus \cdots$. The set $\mathcal{L}_M$ is the space of M-electron density matrices. If $\hat{\Gamma}_M \in \mathcal{L}_M$, then $\left[\hat{N},\hat{\Gamma}_M\right] = 0$. Consider the following functional:

$$\widetilde{e}_0\left[\hat{\Gamma};v\right] = \mathrm{Tr}\left[\hat{H}_v\hat{\Gamma}\right], \tag{5}$$

which also describes the energy of a system with any number of particles. Expanding $\hat{\Gamma}$ we find that

$$\widetilde{e}_0\left[\hat{\Gamma};v\right] = \sum_{M\in\mathbb{N}} w_M\widetilde{e}_0\left[\hat{\Gamma}_M;v\right].$$

The above equation is similar to (3); the difference is that $\widetilde{e}_0\left[\hat{\Gamma}_M;v\right]$ is a linear combination of ensemble energies of systems with a discrete number of electrons, whereas (3) is a linear combination of pure-state energies. By the variational principle, the infimum of $\widetilde{e}$ over density matrices in $\mathcal{L}_M$ for a given potential v is equal to $E_0(M;v)$. Therefore, the ground-state energy can be alternatively calculated using

$$\widetilde{E}_0(N;[v]) = \inf\left\{\widetilde{e}_0\left[\hat{\Gamma};v\right] \middle| \mathrm{Tr}\left[\hat{N}\hat{\Gamma}\right] = N, \hat{\Gamma} \in \mathcal{L}\right\}. \tag{6}$$

For a real, non-integer, number of electrons, the solution of the minimization of $\widetilde{e}$, under the constraints shown in (4), is a linear interpolation of the energy of the states with the nearest integer number of electrons. Hence, if $J < N < J+1$ ($J \in \mathbb{N}$), the energy is given by:

[1] In this work by "subdensity-matrix" we mean a density matrix corresponding to a state with strictly an integer number of electrons.

$$E_0(N;[v]) = (1-\omega)E_0(J;[v]) + \omega E_0(J+1;[v]), \tag{7}$$

where $\omega = N - J$. In general, the above result for the energy can be re-expressed as

$$E_0(N;[v]) = \sum_{M\in\mathbb{N}} y(N-M)E_0(M;[v]).$$

The definition of the function y is

$$y(x) = \begin{cases} 1 & x = 0, \\ 1-x & 0 < x < 1, \\ 1+x & -1 < x < 0, \\ 0 & \text{otherwise.} \end{cases}$$

The averaging formula (7) holds if $E_0(N;[v])$ is convex. For two particles $N_1, N_2 \in \mathbb{R}^+$ with $N_2 > N_1$:

$$E_0(tN_1 + (1-t)N_2;[v]) \le tE_0(N_1;[v]) + (1-t)E_0(N_2;[v]) \quad 0 < t < 1.$$

For some $J \in \mathbb{N}$, if $N_1 = J-1$, $N_2 = J+1$, and $t = 1/2$, the above relation turns into

$$E(J;[v]) \le \frac{1}{2}(E_0(J-1;[v]) + E_0(J+1;[v])).$$

This is referred to as the convexity assumption. No counterexample to this assumption is known for electronic systems under Coulomb repulsions. In general, any observable of the system becomes a linear interpolation; for example, if $\{n_M\}$ are the ground-state densities of the system under v for $M \in \mathbb{N}$, then the average density is

$$n(\mathbf{r}) = \sum_{M\in\mathbb{N}} y(N-M)n_M(\mathbf{r}).$$

The chemical potential is given by

$$\mu_0(N;[v]) = \frac{dE_0}{dN}(N;[v]).$$

This function shows a step-like behavior with respect to the number of electrons:

$$\mu_0(N;[v]) = E(\lceil N\rceil;[v]) - E(\lfloor N\rfloor;[v]) \quad N \notin \mathbb{N},$$

$\lceil\ \rceil$ and $\lfloor\ \rfloor$ are the ceiling and floor functions, respectively. In the special case where the molecule is neutral for J electrons, the chemical potential around J is the ionization of the system, or its affinity; the discontinuity of the chemical potential is

$$\lim_{\Delta N \downarrow 0} \mu_0 (J + \Delta N; [v]) - \mu_0 (J - \Delta N; [v]) = I[v] - A[v],$$

where $I[v] = E_0(J-1;\ [v]) - E_0(J;\ [v])$ is the ionization, and $A[v] = E_0(J;[v]) - E_0(J+1;\ [v])$ is the affinity. The quantity $I - A$ is the fundamental gap of the system [17].

The energy obtained in (7) corresponds to the average energy of a molecule in the grand-canonical ensemble. An ensemble of molecules that are distantly separated is considered (the molecular geometry can be regarded as static). The molecules are also in contact with a distant reservoir of electrons, i.e., a metal surface. When the temperature of this ensemble of molecules is 0^+ K, and the chemical potential of the system is $-A[v]$, each member of the ensemble has either J or $J+1$ electrons. On average, the number of electrons per molecule is a real number between J and $J+1$. An alternative interpretation is the following. A *single*, neutral molecule is placed in contact with a metallic reservoir. If the work function is tuned to the affinity of the molecule [10], then at 0^+ K the molecule exchanges electrons with the surface.[2] Time-averaging of the number of electrons yields a real number between J and $J+1$; the averaged energy of the system, in the ensemble and temporal cases, is thus given by (7).

4 Restricted Spaces

The σ-density (n_σ) corresponding to a state with $J_\uparrow$ spin-up and $J_\downarrow$ spin-down electrons derives from a normalized wavefunction $\psi \in \mathcal{H}_{J_\uparrow + J_\downarrow}$:

$$n_\sigma(\mathbf{r}) = J_\sigma \sum_{\sigma 2, \ldots, \sigma J,} \int \mathrm{d}^3\mathbf{r_2} \ldots \mathrm{d}^3\mathbf{r}_J \left|\psi(\mathbf{r}\sigma, \mathbf{r_2}\sigma_2, \ldots, \mathbf{r}_J\sigma_J)\right|^2.$$

This implies $n_\sigma(\mathbf{r}) \geq 0$, $\forall \mathbf{r} \in \mathbb{R}^3$, and $\int \mathrm{d}^3\mathbf{r}\ n_\sigma(\mathbf{r}) = J_\sigma$. For a given pair of spin-densities n_σ, there always exists an antisymmetric wave function yielding such pair of densities; such a wavefunction is a Slater determinant formed by spin-orbitals $\{|\phi_{k,\sigma}\rangle\}$, $k \in \mathbb{N}$ satisfying [18]:

$$\phi_{k,\sigma}(\mathbf{r}) = \sqrt{\frac{n_\sigma(\mathbf{r})}{\int n_\sigma}} \exp\left(\mathrm{i}k\, f(x)\right), \tag{8}$$

where f is a function chosen to satisfy orthogonality when integration in spatial coordinates is taken. If the J_σ are integer numbers, then the spin-orbitals of the form

[2] The described ensemble is not a truly zero-temperature system because the spin-interactions are neglected.

shown above can be used to construct a single Slater determinant. If the spin-densities do not integrate to an integer number, or the total number of electrons is not an integer, a proper combination of Slater determinants using excited spin-orbitals is required.

For the spin-orbitals of (8) to be physically meaningful, they must have a finite kinetic energy. This requires: $\int \mathrm{d}^3\mathbf{r}\left|\nabla\sqrt{n_\sigma}(\mathbf{r})\right| < \infty$, which is equivalent to demanding that $|\nabla\sqrt{n}|$ is $L^1\left(\mathbb{R}^3\right)$. Hence, the set of physically acceptable electronic densities is

$$\mathcal{D} = \left\{ n \in L^1\left(\mathbb{R}^3\right) \middle| n > 0,\ \left|\nabla n^{1/2}\right| \in L^1\left(\mathbb{R}^3\right),\ \int n \in \mathbb{R}^+ \right\}.$$

To define density functionals, in general, the Fock and Liouville spaces are restricted to sets of states corresponding to a given density. Let us introduce

$$\begin{aligned} \mathcal{G}_{\mathcal{F}}([n],\sigma) &= \left\{ |\psi\rangle \in \mathcal{F} \,\langle\psi|\hat{n}_\sigma(\mathbf{r})|\psi\rangle. = n(\mathbf{r}), \forall \mathbf{r} \in \mathbb{R}^3,\, n \in \mathcal{D} \right\}, \\ \mathcal{G}_{\mathcal{L}}([n],\sigma) &= \left\{ \hat{\Gamma} \in \mathcal{F} \middle| \mathrm{Tr}\left\{\hat{\Gamma}\,\hat{n}_\sigma(\mathbf{r})\right\} = n(\mathbf{r}), \forall \mathbf{r} \in \mathbb{R}^3,\, n \in \mathcal{D} \right\}, \end{aligned} \tag{9}$$

where in the spin-unpolarized case we denote $\mathcal{G}_{\mathcal{U}}([n],\circ), \mathcal{U} = \mathcal{F}, \mathcal{L}$ as the restricted set for the total density; the operator $\hat{n}_\sigma$ is replaced by $\hat{n}$ when $\sigma = \circ$. For a pair of spin-densities $n_\uparrow, n_\downarrow \in \mathcal{D}$ we define

$$\hat{\mathcal{G}}_{\mathcal{U}}[n_\uparrow, n_\downarrow] = \mathcal{G}_{\mathcal{U}}\left([n_\uparrow],\uparrow\right) \cap \mathcal{G}_{\mathcal{U}}\left([n_\downarrow],\downarrow\right); \quad \mathcal{U} = \mathcal{F}, \mathcal{L}\,.$$

5 Density Functionals

5.1 Spin Unpolarized Case

For a given density $n \in \mathcal{D}$, the infimum of the average value of $\hat{H}_0$ over the space $\mathcal{G}_{\mathcal{F}}([n], \circ)$ is a density functional, F_{u}. When evaluated at n, it reads

$$F_{\mathrm{u}}[n] = \inf\left\{ \langle\psi|\hat{H}_0|\psi\rangle \middle| |\psi\rangle \in \mathcal{G}_{\mathcal{F}}([n],\circ) \right\}. \tag{10}$$

If the infimum is sought over $\mathcal{G}_{\mathcal{L}}([n], \circ)$, the result is given by $\widetilde{F}_{\mathrm{u}}[n]$:

$$\widetilde{F}_{\mathrm{u}}[n] = \inf\left\{ \mathrm{Tr}\left[\hat{H}_0\hat{\Gamma}\right] \middle| \hat{\Gamma} \in \mathcal{G}_{\mathcal{L}}([n],\circ) \right\}. \tag{11}$$

Both functionals, F_{u} and $\widetilde{F}_{\mathrm{u}}$, are valid for any positive, real number of electrons. Nevertheless, the functional $\widetilde{F}_{\mathrm{u}}$ is a search over the density matrices of the grand-canonical ensemble. The functional F_{u}, on the other hand, is a mixed quantum-ensemble search: first one searches over the wave-functions corresponding to integer number of electrons; then one searches over the positive weights of the

states with integer numbers of electrons. The functional F_{u} is an extension of the Levy functional [19]:

$$F_{\mathrm{L}}[n] = \inf\left\{\langle\psi|\hat{H}_0|\psi\rangle \,\middle|\, \langle\psi|\hat{n}(\mathbf{r})|\psi\rangle = n(\mathbf{r}), \forall\mathbf{r}\in\mathbb{R}^3, \psi\in\mathcal{H}_N\right\},$$

where $N = \int n \in \mathbb{N}$. The functional $\widetilde{F}_{\mathrm{u}}$, on the other hand, extends the Lieb–Valone functional [20, 21]:

$$F_{\mathrm{V}}[n] = \inf\left\{\hat{\Gamma}\in\mathcal{L}_{\int n} \,\middle|\, \mathrm{Tr}\left[\hat{H}_0\hat{\Gamma}\right], \int \mathrm{d}^3\mathbf{r}\, n(\mathbf{r}) \in \mathbb{N}\right\}.$$

5.2 *Functional Derivatives and* v*-Representability*

The inner product for any pair of real-valued functions defined over $\mathbb{R}^3$ is obtained as follows. Let f and g be two real-valued functions. Then

$$\langle f, g\rangle = \int \mathrm{d}^3\mathbf{r}\, f(\mathbf{r})\, g(\mathbf{r}).$$

Suppose U is some space of functions and $G : U \to \mathbb{R}$. Consider two different functions f and h in U. The differential in the h direction, $\delta_h G[f]$, at f is defined as [22]

$$\delta_h G[f] = \lim_{\epsilon\to 0} \frac{G[f+\epsilon h] - G[f]}{\epsilon}.$$

If this limit exists for any $h \in U$ then the function is Gateaux-differentiable at f. Additionally, the space U must be equipped with a locally convex topological vector space (a set S is convex if for two points in S, x, and y, the point $tx + (1-t)y$ for any $t \in [0, 1]$ is in S as well). The convexity requirement is necessary for a Gateaux derivative to exist [23]. The functional derivative $\delta G/\delta f$ exists if we can write

$$\delta_h G\left[\widetilde{f}\right] = \left\langle \frac{\delta G}{\delta f}\left[\widetilde{f}\right], h\right\rangle = \int \mathrm{d}^3\mathbf{r}\, \frac{\delta G}{\delta f(\mathbf{r})}\left[\widetilde{f}\right] h(\mathbf{r})$$

for any $\widetilde{f}$, h in U.

Let $\mathcal{D}_M \subset \mathcal{D}$ be a subspace of densities integrating to M electrons, where $M \in \mathbb{N}$. The functional F_V over $\mathcal{D}_M$ is $+\infty$ if the density at which F_V is evaluated is not in $\mathcal{D}_M$. Because of this property, F_V is not Gateaux differentiable because if n is a density in $\mathcal{D}_M$, then one can always find a direction in b such that $n + \epsilon b$ is not in $\mathcal{D}_M$ regardless of how small ϵ is, implying that the differential $\delta F_V/\delta n$ does not exist

[24]. On the other hand, for perturbations of the form $n+\epsilon b$ that are in $\mathcal{D}_M$, the directional derivative does exist, which indicates that the term $\delta F_V/\delta n$ must be interpreted under exceptional restrictions.

A density n is said to be pure-state v representable [25, 26] if there is a potential, v, giving rise to a non-degenerate ground-state ket $|\psi\rangle$, such that $n(\mathbf{r}) = \langle\psi|\hat{n}(\mathbf{r})|\psi\rangle$ and $\hat{H}_v|\psi\rangle = E|\psi\rangle$. This definition can be extended to the Fock space for densities integrating to real numbers. If there is a $|\psi\rangle \in \mathcal{G}_{\mathcal{F}}([n], \circ)$ yielding n, and for which $\langle\psi|\hat{H}_0|\psi\rangle$ is an infimum over $\mathcal{G}_{\mathcal{F}}([n], \circ)$ *and* $\langle\psi|\hat{H}_v|\psi\rangle$ is an infimum over $\mathcal{F}$, then we refer to n as Fock-space v-representable (FSVR). On the other hand, if there is a density matrix in $\mathcal{G}_{\mathcal{L}}([n], \circ)$ that gives n such that $\mathrm{Tr}\left[\hat{H}_0\hat{\Gamma}\right]$ is infimized with respect to $\mathcal{G}_{\mathcal{L}}([n], \circ)$ and $\mathrm{Tr}\left[\hat{H}_v\hat{\Gamma}\right]$ is an infimum over $\mathcal{L}([n], \circ)$, then we say n is ensemble-v-representable (EVR). An FSVR density is EVR because one can write the density matrix as a combination of pure-state density submatrices. Nevertheless, an EVR density may not be FSVR.

There is a map u from the set of FSVR densities into the set of physical potentials. If n is FSVR, then $u[n]$ is the potential representing such density. Now define the following functional, valid for FSVR densities:

$$f[n] = E_0(N; [u]) - \int \mathrm{d}^3\mathbf{r}\, n(\mathbf{r})u(\mathbf{r}), \tag{12}$$

where N and u on the right-hand-side of the above equation are evaluated at n ($N = \int n$). Let

$$n' = n + \epsilon m$$

be a perturbed density, where ϵ is a small real number and m is a perturbation satisfying $\int m = 1$, and let $v = u(n), \overline{N} = N[n]$. We *assume* there is a potential $v' = v + \epsilon w + o(\epsilon^2)$ [24] that represents the density n'. The differential of f at n in the direction m is

$$\delta_m f[n] = \delta_m E\left(\overline{N}; [v]\right) - \delta_m \int \mathrm{d}^3\mathbf{r}\, n(\mathbf{r})v(\mathbf{r}). \tag{13}$$

Following the steps in the Appendix, we find

$$\begin{aligned}\delta_m f[n] &= \mu(N; [v]) - \int \mathrm{d}^3\mathbf{r}\, m(\mathbf{r})v(\mathbf{r}),\\ &= \int \mathrm{d}^3\mathbf{r}\left(\mu\left(\overline{N}; [v]\right) - v(\mathbf{r})\right)m(\mathbf{r}).\end{aligned} \tag{14}$$

This result suggests that under the v-representability condition in the Fock space we obtain

$$\frac{\delta f}{\delta n(\mathbf{r})}[n] = \mu\big(\overline{N};[v]\big) - v(\mathbf{r})$$

and

$$\frac{\delta E_0}{\delta n(\mathbf{r})} = \mu\big(\overline{N};[u]\big).$$

This result is valid if $\overline{N}$ is not an integer, and implies that the functional derivative of E_0 is discontinuous (non-analytic) with respect to the number of electrons.

The functional derivative of the functional F_V evaluated at a certain EVR density, under the special restrictions mentioned above, is a constant minus the external potential which represents the density. Notwithstanding, not every density is EVR nor is F_V differentiable at every density. Alternative formulations to avoid the differentiability, and/or representability problem, are reported in the literature. Chayes et al. [27] showed that discretizing the coordinate space into a lattice and formulating quantum mechanics in such a lattice solves the v-representability problem. Examination of the limit when the grid spacing tends to zero leads to recovered derivatives and potentials. Ayers [28] suggested that any non-v-representable density can be approached from a sequence of v-representable ones. Kvaal et al. [29] added a small term to the energy that penalizes largely oscillating potentials. For an arbitrary strength of the penalty, the differentiability is recovered, and by analyzing the limit when the coupling goes to zero one recovers the physical quantities of interest. The price for differentiability and/or v-representability is an extra operation that maps the system out of the troublesome regime and then brings it back by some limiting procedure (or extra operation). This also applies to time-dependent density-functional theory [30–32].

In the exact framework discussed here, there is thus a problem of representability and differentiability. In practice, however, neither exotic densities nor the exact energy functional is accessible, except in very special cases. The space of densities is severely constrained by computational limitations. The problem solved in practice is a minimization over a set of expansion coefficients, $\mathbf{C}_\uparrow$, $\mathbf{C}_\downarrow$. A differentiable energy functional is constructed (i.e., the LSDA energy) and then the KS orbitals are expanded in terms of a truncated basis of Gaussian functions. The resulting problem is that of minimizing $\mathcal{E}(\mathbf{C}_\uparrow, \mathbf{C}_\downarrow)$ subject to a normalization condition, a nonlinear optimization problem. In principle, one would search using a convenient algorithm using derivatives of the function $\mathcal{E}$, requiring functional differentiation as an intermediate step through the chain rule. The search for the best expansion coefficients, however, is performed in practice by self-consistent solution of the KS equations; convergence within this scheme is challenging [33, 34]. Symmetry-breaking must be properly addressed. In addition, near avoided crossing points the HOMO-LUMO gap might be small, causing convergence problems.

6 Spin-Polarized Density Functionals

6.1 The 2-Potential

For a system of electrons interacting with a static magnetic field **B**, the Hamiltonian is written as

$$\hat{H}[v, \mathbf{B}] = \hat{H}_0 + \int \mathrm{d}^3\mathbf{r}\,[\hat{\mathbf{m}}(\mathbf{r}) \cdot \mathbf{B}(\mathbf{r}) + \hat{n}(\mathbf{r})v(\mathbf{r})],$$

where $\hat{\mathbf{m}}(\mathbf{r})$ is the local magnetization operator: $\hat{\mathbf{m}}(\mathbf{r}) = \frac{1}{2}\sum_{\sigma\sigma'} \hat{\psi}^{\dagger}_{\sigma}(\mathbf{r})\boldsymbol{\sigma}_{\sigma\sigma'}\hat{\psi}_{\sigma'}(\mathbf{r})$, where $\boldsymbol{\sigma}$ is the Pauli Spin-matrix vector. The conditions of the HK theorem stating that there is a one-to-one map between ground-state densities and external potentials can be extended and analyzed for systems under static magnetic fields. The constrained-search density functional, analog to F, is defined over the set of kets that yield the local magnetization $\mathbf{m}(\mathbf{r})$ and density $n(\mathbf{r})$. The resulting functional is then employed to introduce the total-energy functional depending on $\mathbf{m}$ and n. Although there is an energy functional depending on the ($\mathbf{m}$, n) pair, the 1–1 correspondence between ($\mathbf{m}$, n) and (v, $\mathbf{B}$) is absent if the ground-state ket $|\psi\rangle$ is common to another system subject to a different combined magnetic-scalar field pair [35]. In other words, there is not a 1–1 map if one can find a perturbation $\Delta\mathbf{B}$, Δn such that $\int[\hat{\mathbf{m}}(\mathbf{r}) \cdot \Delta\mathbf{B}(\mathbf{r}) + \hat{n}(\mathbf{r})\Delta v(\mathbf{r})]|\psi\rangle = C|\psi\rangle$, $|\psi\rangle$ remains the ground state of the system under the potentials $\mathbf{B} + \Delta\mathbf{B}$ and $v + \Delta v$, and C is a constant. Therefore, the energy functional regarding the local magnetization and electronic density is many-valued. Nevertheless, in the presence of non-collinear magnetic fields, the 1–1 correspondence is preserved [36].

Suppose $\mathbf{B}(\mathbf{r}) = (0, 0, B_z(\mathbf{r}))$, and set

$$v_{\uparrow}(\mathbf{r}) = v(\mathbf{r}) + \frac{1}{2}B_z(\mathbf{r}),$$
$$v_{\downarrow}(\mathbf{r}) = v(\mathbf{r}) - \frac{1}{2}B_z(\mathbf{r}).$$

These potentials allow us to write $v = v_{\uparrow} + v_{\downarrow}$ and $B_z = 2(v_{\uparrow} - v_{\downarrow})$. We define the Hamiltonian for the electrons in the collinear case as

$$\hat{H}[\mathbf{v}] = \hat{H}_0 + \int \mathrm{d}^3\mathbf{r}\,\mathbf{v}(\mathbf{r}) \cdot \hat{\mathbf{n}}(\mathbf{r}),$$

where $\mathbf{v} = (v_{\uparrow}, v_{\downarrow})$ and $\hat{\mathbf{n}} = (\hat{n}_{\uparrow}, \hat{n}_{\downarrow})$. $\hat{H}[\mathbf{v}]$ is the Hamiltonian of main interest and we introduce spin-density functionals based on it in the next sections.

6.2 Lowest Energy in Terms of the Number of Spin-Electrons

Let $\mathbf{N}$ be a vector of two components, the number of spin-up ($N_\uparrow$) and spin-down ($N_\downarrow$) electrons. For $\mathbf{N} = (N_\uparrow, N_\downarrow)$, $N_\uparrow, N_\downarrow \in \mathbb{R}^+$, let E be energy:

$$E(\mathbf{N}; [\mathbf{v}]) = \inf\left\{\widetilde{e}\left[\hat{\Gamma}; \mathbf{v}\right] \middle| \left\langle \psi \middle| \hat{\mathbf{N}} \middle| \psi \right\rangle = \mathbf{N}, \hat{\Gamma} \in \mathcal{L}\right\}, \tag{15}$$

where $\widetilde{e}\left[\hat{\Gamma}; \mathbf{v}\right]$ has the same form as $\widetilde{e}_0$ except that $\hat{H}_v$ is replaced by $\hat{H}[\mathbf{v}]$ in the r.h.s. of (5). To calculate $E(\mathbf{N}; [\mathbf{v}])$ one first minimizes over the subdensity matrices, and then over the weights, leading to the linear-programming problem:

$$\begin{aligned} &\min \sum_{\mathbf{M}} w_{\mathbf{M}} E_{\mathbf{M}}[\mathbf{v}] \\ &\text{s.t.} \sum_{\mathbf{M}} w_{\mathbf{M}} = 1 \\ &\sum_{\mathbf{M}} \mathbf{M} w_{\mathbf{M}} = \mathbf{N} \\ &0 \leq w_{\mathbf{M}} \leq 1. \end{aligned} \tag{16}$$

Upon specification of $\mathbf{N}$ and $\mathbf{v}$, solution of the above linear programming problem yields the average energy of a system in the ensemble as a function of $\mathbf{N}$ and $\mathbf{v}$. Unlike the case of spin-unpolarized systems, there is no simple convexity assumption allowing us to derive an analytical expression for the energy as a function of the average number of spin-electrons, $\mathbf{N}$. For example, introduce $N = N_\uparrow + N_\downarrow$ and $S_z = 1/2(N_\uparrow - N_\downarrow)$, and suppose N is fixed. For a neutral atom such as carbon the energy along the pairs $(N_\uparrow, N_\downarrow)$ is not strictly convex because the energy tends to be minimized according to Hund's rules [37, 38],[3] giving a flat line between states with $N_\downarrow = 5, 6, 7$. For the carbon anion [40–42] the ground state is a quartet; if N is fixed as 7 then there is a flat line between $N_\downarrow = 2, 3, 4, 5$.

Balawender [43] performed an analysis in the grand-canonical ensemble of spin-electrons, and Malek and Balawender [44] studied the bi-spline interpolations of a set of molecules using several spin-polarized DFAs. They found that the energy as a function of S_z and N is a series of triangles with derivative discontinuities at the edges. The pattern in which the triangles form varies depending on how the energies of the pure states distribute. Malek and Balawender [44] offer explicit formulas for the chemical potentials in the different types of triangles. These formulas are useful for the explicit calculation of the derivative discontinuities. Because the discontinuities are system-dependent, we approach the problem in a more general fashion, such that the formalism is compact.

The minimum finite element for interpolation in 2d surfaces is a triangle; one needs three different points to form a plane. We thus assume that the energy $E(\mathbf{N}; [\mathbf{v}])$ is formed by averaging the energies at the vertices of the corresponding

[3] Here we use Hund's rules as a guide. Exceptions to these rules are known [39].

triangle. Let Δ_k be a triangle used in the interpolation. The domain of $E(\mathbf{N}; [\mathbf{v}])$ is then $\cup_k \Delta_k$. Denoting $\mathbf{A}_k$, $\mathbf{B}_k$, and $\mathbf{C}_k$ as the vertices of the triangle we find that

$$E(\mathbf{N}; [\mathbf{v}]) = w_{\mathbf{A}_k} E(\mathbf{A}_k; [\mathbf{v}]) + w_{\mathbf{B}_k} E(\mathbf{B}_k; [\mathbf{v}]) + w_{\mathbf{C}_k} E(\mathbf{C}_k; [\mathbf{v}]).$$

To simplify the notation let $V(\Delta_k) = \{\mathbf{A}_k, \mathbf{B}_k, \mathbf{C}_k\}$. The average number of spin-electrons is

$$\mathbf{N} = \sum_{\mathbf{M} \in V(\Delta_k)} w_{\mathbf{M}} \mathbf{M}.$$

An important feature of the triangles is that they are determined by a functional of the 2-potential $\mathbf{v}$.

Suppose $\mathbf{N}$ lies in a triangle whose vertex with the lowest number of electrons is $\mathbf{A}_k$, a singlet state ($A_{\uparrow,k} = A_{\downarrow,k}$), and let

$$\omega_\sigma = N_\sigma - A_{\sigma,k}.$$

If $\mathbf{B}_k = (A_{\downarrow,k}, A_{\uparrow,k} + 1)$ and $\mathbf{C}_k = (A_{\downarrow,k} + 1, A_\uparrow)$, then the energy of the system becomes

$$E(\mathbf{N}; [\mathbf{v}]) = \big(1 - \omega_\uparrow - \omega_\downarrow\big) E(\mathbf{A}; [\mathbf{v}]) + \omega_\uparrow E\big(\mathbf{A} \pm \mathbf{e}_\uparrow; [\mathbf{v}]\big) + \omega_\downarrow E\big(\mathbf{A} \pm \mathbf{e}_\downarrow; [\mathbf{v}]\big).$$

Now assume $v_\uparrow = v_\downarrow$. If we let

$$\omega = \omega_\uparrow + \omega_\downarrow,$$

we recover the PPLB form:

$$E(\mathbf{N}; [\mathbf{v}]) = (1 - \omega) E(\mathbf{A}; [\mathbf{v}]) + \omega E\big(\mathbf{A} \pm \mathbf{e}_\uparrow; [\mathbf{v}]\big).$$

It should be noted that one can add $\mathbf{e}_\uparrow$ or $\mathbf{e}_\downarrow$ to $\mathbf{A}$ in the argument of the last term of the above equation and the result remains the same because $v_\uparrow = v_\downarrow$. In general, the energy surface as a function of the spin-electron numbers is piecewise bilinear, where the surface is a union of triangular elements.

6.3 Spin DFT

The spin-densities are used instead of the total electronic density to perform calculations in SDFT, hence the name. Denote the 2-density as a vector $\mathbf{n}(\mathbf{r}) = (n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r}))$. The energy is expressed as

$$E[\mathbf{n};\mathbf{v}] = F[\mathbf{n}] + \int \mathrm{d}^3\mathbf{r}\left(v_\uparrow(\mathbf{r})n_\uparrow(\mathbf{r}) + v_\downarrow(\mathbf{r})n_\downarrow(\mathbf{r})\right). \tag{17}$$

In a similar fashion to the spin-unpolarized case, the functional $F[\mathbf{n}]$ is defined as

$$F[\mathbf{n}] = \inf\left\{\mathrm{Tr}\left[\hat{H}_0\hat{\Gamma}\right] \middle| \hat{\Gamma} \in \widetilde{\mathcal{G}}_{\mathcal{L}}[n]\right\}.$$

For a given 2-potential $(v_\uparrow, v_\downarrow)$, (17), when minimized over the spin-densities, yields the ground state of the system. However, it should be noted that the Hohenberg–Kohn theorem does not apply for the map between 2-potentials and 2-densities because there can be two different sets of spin-densities that yield the same total electronic density [35].

If there is a 2-potential $\mathbf{u}$, yielding $\mathbf{n}$, for which the following quantity exists:

$$\mathcal{E}[\mathbf{n}] = \inf\left\{\mathrm{Tr}\left[\hat{H}[\mathbf{u}]\hat{\Gamma}\right] \middle| \mathrm{Tr}\left[\hat{\mathbf{N}}\hat{\Gamma}\right] = \int \mathbf{n}, \hat{\Gamma} \in \mathcal{L}\right\},$$

then $\mathbf{n}$ is ensemble 2-v representable (E2VR). The infimization to calculate $\mathcal{E}$ is performed only over density matrices yielding $\int \mathbf{n}$ pairs of spin-electrons.

The 2-potential $\mathbf{u}$ is a functional of the E2VR spin-densities. Let $\hat{\Gamma}^*[\mathbf{n}]$ (for E2VR densities) be defined by

$$\hat{\Gamma}^*[\mathbf{n}] = \mathrm{arginf}\left\{\mathrm{Tr}\left[\hat{H}[\mathbf{u}]\hat{\Gamma}\right] \middle| \mathrm{Tr}\left[\hat{\mathbf{N}}\hat{\Gamma}\right] = \int \mathbf{n}, \hat{\Gamma} \in \mathcal{L}\right\},$$

where $\hat{\Gamma}^*[\mathbf{n}]$ is a density matrix that minimizes the quantity $\mathrm{Tr}\left[\hat{H}[\mathbf{u}]\hat{\Gamma}\right]$ under the mentioned restrictions. Around the pair $\mathbf{M}$, $\hat{\Gamma}^*[\mathbf{n}]$ obeys

$$\hat{\Gamma}^*[\mathbf{n}] = \sum_{\mathbf{M}\in V(\Delta[\mathbf{n}])} w_{\mathbf{M}}\hat{\Gamma}^*_{\mathbf{M}}[\mathbf{n}]. \tag{18}$$

The submatrices $\hat{\Gamma}^*_{\mathbf{M}}[\mathbf{n}]$, $\mathbf{M} \in V(\Delta[\mathbf{n}])$, give rise to three densities at the vertices, $\{\mathbf{n}_\mathrm{M}\}$ $\mathbf{M} \in V(\Delta[n])$. Now define the projector $\hat{I}_\mathbf{J}$, for $\mathbf{J} \in \mathbb{N}^2$, as follows:

$$\hat{I}_\mathbf{J} = \sum_{\mathbf{k}} |\psi_{\mathbf{J},\mathbf{k}}\rangle\langle\psi_{\mathbf{J},\mathbf{k}}|,$$

where $\{|\psi_\mathbf{k}\rangle\}$ is a complete basis set with $\langle\psi_{\mathbf{J},\mathbf{k}}|\hat{\mathbf{N}}|\psi_{\mathbf{J},\mathbf{k}}\rangle = \mathbf{J}$. We extract the densities of the discrete states (those with strictly integer number of spin-electrons) of $\hat{\Gamma}[\mathbf{n}]$:

$$\mathbf{n}_{\mathbf{J}}(\mathbf{r}, [\mathbf{n}]) = \mathrm{Tr}\left[\hat{I}_{\mathbf{J}}\hat{\mathbf{n}}(\mathbf{r})\hat{\Gamma}^{*}[\mathbf{n}]\right].$$

The energy of a discrete state in the ensemble is obtained in a similar fashion:

$$\mathcal{E}_{\mathbf{J}}[\mathbf{n}] = \mathrm{Tr}\left[\hat{I}_{\mathbf{J}}\hat{H}[\mathbf{u}]\hat{\Gamma}^{*}[\mathbf{n}]\right]$$

We limit the free-energy functional to E2VR 2-densities:

$$F[\mathbf{n}] = \mathcal{E}[\mathbf{n}] - \int \mathrm{d}^3\mathbf{r}\ \mathbf{u}[\mathbf{n}](\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}).$$

Insertion of (20) into the above equation yields the recursive relation:

$$F[\mathbf{n}] = \sum_{\mathbf{M}\in V(\Delta[n])} w_{\mathbf{M}} F[\mathbf{n}_{\mathbf{M}}]. \tag{19}$$

For simplicity, suppose that $\omega_{\uparrow} + \omega_{\downarrow} < 1$ and $\omega_{\uparrow} > 0$, $\omega_{\downarrow} > 0$. Following the steps shown in Sect. 5.2, it can be shown that

$$\frac{\delta F}{\delta n_{\sigma}(\mathbf{r})}[\mathbf{n}] = \mu_{\sigma}[\mathbf{n}] - u_{\sigma}[\mathbf{n}](\mathbf{r})$$

For this result to hold, the vertices of the triangle must be the same upon perturbation of the 2-density.

The chemical potential $\mu_{\sigma}[\mathbf{n}]$ is a function of the energies at the vertices. In the simple case where the triangle is $\mathbf{B}_k = (A_{\downarrow}, A_{\uparrow} + 1)$ and $\mathbf{C}_k = (A_{\downarrow} + 1, A_{\uparrow})$:

$$\mu_{\sigma}[\mathbf{n}] = \mathcal{E}_{\mathbf{A}+\mathbf{e}_{\sigma}}[\mathbf{n}] - \mathcal{E}_{\mathbf{A}}[\mathbf{n}],$$

which is the energy released when a spin-electron is added to the system in the lowest energy state $\mathbf{A}$.

7 Integer Discontinuities

The results presented in this section generalize the work of [15]. Let us consider the variation of F:

$$\delta_{\mathbf{m}} F[\mathbf{n}] = \lim_{\epsilon\to 0} \frac{1}{\epsilon}[F[\mathbf{n} + \epsilon\mathbf{m}] - F[\mathbf{n}]] = \int \mathrm{d}^3\mathbf{r}\ \frac{\delta F}{\delta \mathbf{n}}[\mathbf{n}](\mathbf{r}) \cdot \mathbf{m}(\mathbf{r}),$$

where

$$\frac{\delta F}{\delta \mathbf{n}(\mathbf{r})}[\mathbf{n}] = \left(\frac{\delta F}{\delta n_{\uparrow}(\mathbf{r})}[\mathbf{n}], \frac{\delta F}{\delta n_{\downarrow}(\mathbf{r})}[\mathbf{n}]\right).$$

From the results of the previous section we infer the following. The variation $\delta_{\mathrm{m}} F$ displays a discontinuity at the points where non-coplanar triangles join. In some cases the discontinuity is present at the points where the spin-electron numbers are integers. $\delta F/\delta n_{\sigma}$ also displays discontinuities.

Suppose $\int \mathbf{n}$ is located at a joint-point, $\mathbf{u}$ is the potential representing $\mathbf{n}$, and $\mathbf{N} = \int \mathbf{n}$. We define the following the DD measure:

$$\Delta_{F,\sigma}[\mathbf{n}](\mathbf{r}) = \lim_{\epsilon \downarrow 0} \frac{\delta F}{\delta n_{\sigma}(\mathbf{r})}\left[\mathbf{n}_{\sigma}^{+\epsilon}\right] - \frac{\delta F}{\delta n_{\sigma}(\mathbf{r})}\left[\mathbf{n}_{\sigma}^{-\epsilon}\right],$$

where $\mathbf{n}_{\sigma}^{\pm\epsilon}$ is determined by the potential $\mathbf{u}$, fixed at the joint point, and $\mathbf{N} \pm \epsilon$. The discontinuity is thus taken along a path of ensemble ground-state densities. In this case we set the external 2-potential as fixed and only vary the number of electrons. Therefore

$$\Delta_{F,\sigma}[\mathbf{n}](\mathbf{r}) = \mu_{\sigma}^{+}[\mathbf{n}] - \mu_{\sigma}^{-}[\mathbf{n}],$$

where

$$\mu_{\sigma}^{\pm}[\mathbf{n}] = \mu_{\sigma}\left[\mathbf{n} \pm \mathbf{n}_{\sigma}^{\pm\epsilon}\right].$$

The discontinuity of the functional derivative of F is independent of the position $\mathbf{r}$. In the special case where $\mathbf{u}_{\uparrow} = \mathbf{u}_{\downarrow}$, and $N_{\uparrow} = N_{\downarrow}$, or $|N_{\uparrow} - N_{\downarrow}| = 1$, the DDs of F are each the fundamental gap of the system. In contrast with spin-unpolarized DFT, the above results for the functional derivative of F carry a caution: there can be another 2-potential $\mathbf{u}'$ that differs from $\mathbf{u}$ by more than a 2-component constant vector, yielding the same density $\mathbf{n}$.

One can define a DD that depends on the directions used to perturb the density. For example, if two E2VR 2-densities $\mathbf{m}^{+}$ and $\mathbf{m}^{-}$ are considered, the generalized discontinuity $\delta F/\delta n_{\sigma}[\mathbf{n} + \epsilon \mathbf{m}^{+}] - \delta F/\delta n_{\sigma}[\mathbf{n} - \epsilon \mathbf{m}^{-}]$, as $\epsilon \downarrow 0$, is not related to the fundamental gap of the system at $\mathbf{n}$ because $\mathbf{m}^{+}$ and $\mathbf{m}^{-}$ are arbitrary variations with arbitrary density decay in the asymptotics; the generalized DD of F is a function that depends on position.

7.1 KS SDFT

In spin-unpolarized KS-DFT, a system of non-interacting electrons is subject to an effective potential such that its electronic density matches the ground-state

electronic density of the "real system". The extension of this formalism considering spin-densities is thanks to von Barth and Hedin [45].

The kinetic energy of a system of non-interacting electrons is

$$T_{\mathrm{s}}[\mathbf{n}] = \inf\left\{\mathrm{Tr}\left[\hat{T}\,\hat{\Gamma}\right] \middle| \hat{\Gamma} \in \mathcal{L}_{\Delta}[\mathbf{n}]\right\}.$$

This kinetic energy functional averages over the triangle of the density-functional F. It is convenient to introduce the following space:

$$\mathcal{L}_{\Delta}[\mathbf{n}] = \left\{ \hat{\Gamma} \in \widetilde{\mathcal{G}}_{\mathcal{L}}[\mathbf{n}] \middle| \hat{\Gamma} = \sum_{\mathbf{M}\in\Delta[\mathbf{n}]} w_{\mathbf{M}} \hat{\Gamma}_{\mathbf{M}} \right\}.$$

A density is said to be non-interacting ensemble 2-v representable if there is a potential $\mathbf{u}_{\mathrm{s}}$ such that there is a density matrix satisfying

$$\begin{aligned} \hat{\Gamma}^{*}_{\mathrm{s}}[\mathbf{n}] = \mathrm{arginf}\left\{ \mathrm{Tr}\left[\hat{H}_{\mathrm{s}}[\mathbf{u}_{\mathrm{s}}]\hat{\Gamma}\right] \middle| \mathrm{Tr}\left[\hat{\mathbf{N}}\,\hat{\Gamma}\right] = \int \mathbf{n},\, \hat{\Gamma} \in \mathcal{L}_{\Delta} \right\}, \\ \mathrm{Tr}\left[\hat{\Gamma}^{*}_{\mathrm{s}}[\mathbf{n}]\hat{\mathbf{n}}\,(\mathbf{r})\right] = \mathbf{n}(\mathbf{r}), \end{aligned}$$

where $\hat{H}_{\mathrm{s}}$ is the Hamiltonian of the non-interacting system of electrons:

$$\hat{H}_{\mathrm{s}}[\mathbf{u}_{\mathrm{s}}] = \hat{T} + \int \mathrm{d}^3\mathbf{r}\, \mathbf{u}_{\mathrm{s}}(\mathbf{r}) \cdot \hat{\mathbf{n}}\,(\mathbf{r}).$$

The density matrix can be decomposed as

$$\hat{\Gamma}^{*}_{\mathrm{s}}[\mathbf{n}] = \sum_{\mathbf{M}\in\Delta[\mathbf{n}]} w_{\mathbf{M}} \hat{\Gamma}^{*}_{\mathrm{s},\mathbf{M}}[\mathbf{n}]. \tag{20}$$

Note, however, that the density matrix $\hat{\Gamma}^{*}_{\mathrm{s}}$ comes from a different constrained search than $\hat{\Gamma}^{*}$, and the projected densities of the form

$$\mathbf{n}_{\mathbf{J},\mathrm{s}}[\mathbf{n}](\mathbf{r}) = \mathrm{Tr}\left\{\hat{I}_{\mathbf{J}}\hat{\mathbf{n}}\,(\mathbf{r})\hat{\Gamma}^{*}_{\mathrm{s}}[\mathbf{n}]\right\}$$

are not necessarily the same as those extracted from $\hat{\Gamma}^{*}(\mathbf{n})$. In the case where the spin-electron density $\mathbf{n}$ integrates to integer number of spin-electrons, for example $\int \mathbf{n} = \mathbf{M}$, $\mathbf{n}_{\mathbf{M}} = \mathbf{n}_{\mathrm{s},\mathbf{M}}$.

For a non-interacting E2VR density n, the KS equations are

$$\left(-\frac{1}{2}\nabla^2_{\mathbf{r}} + u_{\mathrm{s},\sigma}[\mathbf{n}](\mathbf{r})\right)\varphi_{\mathbf{k},\sigma}[\mathbf{n}](\mathbf{r}) = \epsilon_{\mathbf{k},\sigma}[\mathbf{n}]\varphi_{\mathbf{k},\sigma}[\mathbf{n}](\mathbf{r}),$$

the spin-densities of the system being obtained from

$$n_\sigma(\mathbf{r}) = \sum_{\mathbf{M}\in\Delta[n]} w_{\mathbf{M}} n_{\sigma,\mathbf{M}}(\mathbf{r}).$$

This averaging of densities is similar to that of Ullrich and Kohn [46], who presented an extension of the KS theory for densities that come from degenerate orbitals, for example, those of atoms with non-zero total angular momentum. In such a case, the HOMO level is manifold, where each orbital is assigned an occupation number. In our case, we are assigning weights to spin-densities that, if required, are already averaged over spatial degeneracy, according to the formalism of Ullrich and Kohn [46].

Suppose that the number of spin-electrons lies in a doublet-singlet triangle, such as the one used to recover the PPLB averaging in Sect. 6.2. To minimize the kinetic energy of the system, each level is fully occupied ($f_{\mathbf{k},\sigma} = 1$), in increasing order, and the highest occupied level is assigned an occupation number ω_σ (the remainder levels are not occupied; this is valid as long as the highest occupied energy level is non-degenerate). Therefore

$$\sum_{\mathbf{k}} f_{\mathbf{k},\sigma} = M_\sigma + \omega_\sigma.$$

The non-interacting projected densities are

$$\mathbf{n}_{\mathrm{s},\mathbf{M}}(\mathbf{r}) = \sum_{\text{fully occ.}} \left(\left|\varphi_{\mathbf{k},\uparrow}(\mathbf{r})\right|^2, \left|\varphi_{\mathbf{k}',\downarrow}(\mathbf{r})\right|^2 \right),$$
$$\mathbf{n}_{\mathrm{s},\mathbf{M}+\mathbf{e}_\sigma}(\mathbf{r}) = \mathbf{n}_{\mathrm{s},\mathbf{M}}(\mathbf{r}) + \left|\varphi_{\mathrm{HOMO},\sigma}\right|^2(\mathbf{r})\mathbf{e}_\sigma.$$

At non-interacting E2VR densities (and directions), the functional derivative of T_s becomes

$$\frac{\delta T_\mathrm{s}}{\delta n_\sigma(\mathbf{r})}[\mathbf{n}] = \mu_{\mathrm{s},\sigma}[\mathbf{n}] - u_{\mathrm{s},\sigma}(\mathbf{r}),$$

where $\mu_{\mathrm{s},\sigma}$ is the chemical potential of the channel σ. It is a function of the HOMO orbitals of the states at the triangle vertices; for the single-doublet triangle the chemical potential is the HOMO energy of the state $\mathbf{M}+\mathbf{e}_\sigma$. Assuming that $\mathbf{u}_\mathrm{s}$ gives bound states and that it goes to zero in asymptotics, we obtain the discontinuity

$$\begin{aligned}\Delta_{T_\mathrm{s},\sigma}(\mathbf{r},[\mathbf{n}]) &= \lim_{\epsilon\downarrow 0} \frac{\delta T_\mathrm{s}}{\delta n_\sigma(\mathbf{r})}\left[\mathbf{n}+\mathbf{n}_\sigma^{+\epsilon}\right] - \frac{\delta T_\mathrm{s}}{\delta n_\sigma(\mathbf{r})}\left[\mathbf{n}-\mathbf{n}_\sigma^{-\epsilon}\right] \\ &= \mu^+_{\mathrm{s},\sigma}[\mathbf{n}] - \mu^-_{\mathrm{s},\sigma}[\mathbf{n}].\end{aligned}$$

The functional F is expanded as follows:

$$F[\mathbf{n}] = T_{\mathrm{s}}[\mathbf{n}] + E_{\mathrm{HXC}}[\mathbf{n}], \tag{21}$$

where E_{HXC} is the Hartree-exchange-correlation spin-density functional. By definition,

$$E_{\mathrm{HXC}}[\mathbf{n}] = E_{\mathrm{H}}[\mathbf{n}] + E_{\mathrm{XC}}[\mathbf{n}],$$

where

$$E_{\mathrm{H}}[\mathbf{n}] = \frac{1}{2}\int \mathrm{d}^3\mathbf{r}\mathrm{d}^3\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|},$$

and E_{XC} is the exchange-correlation energy. The Hartree energy does not display derivative discontinuity, implying that the integer discontinuity of the XC energy functional is

$$\begin{aligned}\Delta_{\mathrm{XC},\sigma}[\mathbf{n}] &= \Delta_{\mathcal{F}_\sigma}[\mathbf{n}] - \Delta_{T_s,\sigma}[\mathbf{n}] \\ &= \left(\mu_\sigma^+[\mathbf{n}] - \mu_\sigma^-[\mathbf{n}]\right) - \left(\mu_{\mathrm{s},\sigma}^+[\mathbf{n}] - \mu_{\mathrm{s},\sigma}^-[\mathbf{n}]\right).\end{aligned}$$

By inserting (21) into (19) we arrive at

$$T_{\mathrm{s}}[\mathbf{n}] + E_{\mathrm{HXC}}[\mathbf{n}] = \sum_{\mathbf{M}\in V(\Delta[n])} w_{\mathbf{M}}(T_{\mathrm{s}}[\mathbf{n}_{\mathbf{M}}] + E_{\mathrm{HXC}}[\mathbf{n}_{\mathbf{M}}]).$$

Expanding T_{s} we find the formula for the HXC energy:

$$E_{\mathrm{HXC}}[\mathbf{n}] = \sum_{\mathbf{M}\in V(\Delta[n])} w_{\mathbf{M}}(T_{\mathrm{s}}[\mathbf{n}_{\mathbf{M}}] - T_{\mathrm{s}}[\mathbf{n}_{\mathrm{s},\mathbf{M}}] + E_{\mathrm{HXC}}[\mathbf{n}_{\mathbf{M}}]). \tag{22}$$

This is a recursion relation between the HXC energy of the ensemble density, the HXC energy of the systems with discrete number of electrons, and a residual term produced by the different density matrices used (i.e., those involved in the constrained searches defining $T_{\mathrm{s}}[\mathbf{n}]$ and $F[\mathbf{n}]$). The densities $\{\mathbf{n}_{\mathrm{M}}\}$ are obtained by projections based on the minimization of $\hat{H}_0$. However, for practical computer calculations the explicit use of $\hat{W}$ is quite expensive. Hence we define

$$G[\mathbf{n}] = \inf_{\mathbf{n}_{\mathbf{M}}\to\mathbf{n}} \sum_{\mathbf{M}\in\Delta[\mathbf{u}]} w_{\mathbf{M}} G[n_{\mathbf{M}}],$$

where

$$G[\mathbf{n_M}] = \inf_{\hat{\Gamma}_{\mathbf{M}} \to \mathbf{n_M}} Tr\left[\hat{\Gamma}_{\mathbf{M}}\left(\hat{T} + \int \mathrm{d}^3\mathbf{r}\, \mathbf{u}_{\mathrm{HXC}}[\mathbf{n_M}](\mathbf{r}) \cdot \hat{\mathbf{n}}\,(\mathbf{r})\right)\right].$$

The optimal densities at the vertices are now obtained in the search defining $G[\mathbf{n}]$, without the use of $\hat{H}_0$, provided an estimation to $\mathbf{u}_{\mathrm{HXC}}$ is given.

An approximation to E_{HXC} for states with integer number of electrons and the density-relaxation terms $T_{\mathrm{s}}[\mathbf{n_M}] - T_{\mathrm{s}}[\mathbf{n}_{\mathrm{s},\mathbf{M}}]$ need to be inserted on the r.h.s. of (22) to recover the discontinuity of any density-functional approximation. In dissociation of symmetric radicals, fractional average numbers of electrons emerge on each fragment. The XC energy averages over the XC energy of the pure states and over the relaxation KS kinetic energy. The latter is a term often ignored in the literature. Relaxation of the orbitals is required to yield the electron densities at the vertices. This can be particularly challenging for small atoms, where self-interaction and correlation errors can lead to unphysical unbound anions; recall, the densities are usually sought through a single self-consistent solution of the KS equations [4] (although non-self consistent densities seem to be useful in these cases [4]).

So far, the various spin-orbit interactions have been omitted. Inclusion of these interactions in the internal Hamiltonian is allowed. We must say, however, that the HXC energy accounting for these relativistic interactions is different from the traditional one. Even if the exact HXC functional in the absence of these couplings is employed, the results are inexact; a proper XC energy functional must be used for the magnetic couplings. The total spin is not a constant of motion if relativistic interactions are considered. Therefore, the ensemble-averaging does not necessarily involve integer states, but fractional ones. In atoms, alternative constant of motions are more adequate such as the J quantum number.

8 Molecular Dissociation

Suppose two fragments are isolated from one another. The total density of the system is the density of the sum of two localized densities $\mathbf{n}_{\mathrm{L}} + \mathbf{n}_{\mathrm{R}}$ such that $\int \mathbf{n}_{\mathrm{L}} \cdot \mathbf{n}_{\mathrm{R}} = 0$. Then the energy obtained using a local or semi-local XC energy functional is size-consistent:

$$E^{\mathrm{A}}[\mathbf{n}_{\mathrm{L}} + \mathbf{n}_{\mathrm{R}}] = E^{\mathrm{A}}[\mathbf{n}_{\mathrm{L}}] + E^{\mathrm{A}}[\mathbf{n}_{\mathrm{R}}],$$

where E^{A} is an approximation to the energy. The LDA and GGA functionals suffer from self-interaction error (SIE). When $N = 1$, the XC energy does not cancel the Hartree energy; the electron density repels itself.

Let us analyze the case of a simple one-electron system, H_2^+. Assume that the nuclei are well separated and the ground-state density is a linear combination of the kets $|\psi_{\mathrm{L}}\rangle$ and $|\psi_{\mathrm{R}}\rangle$, which represent the electron being around the left and the right nucleus, respectively. The state of the system at infinite separation is

$$|\psi\rangle = c_{\mathrm{L}}|\psi_{\mathrm{L}}\rangle + c_{\mathrm{R}}|\psi_{\mathrm{R}}\rangle,$$

where the coefficients only need to satisfy $|c_{\mathrm{L}}|^2 + |c_{\mathrm{R}}|^2 = 1$. On the other hand, if the separation is taken from the equilibrium distance step by step until approaching the infinite separation, the coefficients tend to $|c_{\mathrm{L}}|^2 = |c_{\mathrm{R}}|^2 = 1/2$. The energy at dissociation reads

$$E[\mathbf{n}] \to \frac{1}{2}E_{\mathrm{L}}\left[\mathbf{n}_{\mathrm{L}}^1\right] + \frac{1}{2}E_{\mathrm{R}}\left[\mathbf{n}_{\mathrm{R}}^1\right],$$

where $n_{\alpha,\sigma}^1(\mathbf{r}) = \left|\psi_{\alpha,\sigma}\right|^2(\mathbf{r})$, $\alpha = \mathrm{L}, \mathrm{R}$. The LDA functional, for example, does not satisfy the above condition. If one scales the density of a fragment as $\int (\nu_\alpha n_{\alpha,\sigma})^{4/3} = \nu_\alpha^{4/3} \int n_{\alpha,\sigma}^{4/3}$, one finds that the relation is not linear; the correlation energy also displays a nonlinear dependency.

At infinite separation the total energy splits into a *weighted* sum of energy functionals of the isolated atoms (or fragments). This result is obtained if the energy functional is defined in terms of density matrices or kets. The use of density matrices, however, has several advantages: (1) there is no need to constrain the analysis to dissociating molecules and (2) ensemble-v-representability is stronger than the pure-state one. Nevertheless, the results are the same at the ground state of the molecule.

Vydrov et al. [47] have illustrated the spurious minimization of energy with the XC PBE functional, spin-polarized case, for lithium hydride. Any neutral diatomic molecule dissociates adiabatically to produce two neutral atoms. The PBE spin-polarized functional, when used in self-consistent KS-SDFT calculations, does not yield two neutral species, but two partially charged atoms, i.e., $\mathrm{Li}^{+0.4} + \mathrm{H}^{-0.4}$. This result is unphysical, and calls for a proper correction. Long-range corrected functionals achieve this task by allowing the local exchange to act only around each nucleus, although non-local exact exchange accounts for the long-range interaction. This eliminates the fractional charges via symmetry breaking. Furthermore, when the PBE XC energy functional is assumed to be applicable in the ensemble regime, it is observed that the average energy is a nonlinear function in terms of the average number of electrons that is convex between integers. This piecewise convexity is responsible for the spurious behavior of functionals at dissociation.

The rule for constructing spin-polarized XC density functionals is usually the following: given a functional f define $f'[n_\uparrow, n_\downarrow] = 1/2(f[2n_\uparrow] + f[2n_\downarrow])$, where f' is the spin-polarized density functional. This rule is applied to obtain approximations to the exchange and non-interacting kinetic energies. Even though the rule is simple and appealing, one can rotate the KS orbitals while preserving the total ground-state density and can observe that the energy varies depending on the rotation parameter. For this simple reason, in molecules, the LSDA yields dissociation states whose densities do not correspond to those of pure states (states with integer numbers of electrons and half-integer azimuthal spin number).

If an approximation to the exchange correlation energy displayed the correct limit at a fractional number of electrons, (22), the erroneous behavior would be replaced by a piecewise planar dependency (or a piecewise-linear dependency in the unpolarized case). Nonetheless, we have highlighted the correct form at dissociation and not the whole picture along the dissociation. The question is thus: how can we connect the regimes where the regular DFA works and the regime at complete dissociation?

9 Energy Spin-Density Functionals and Approximations

The energy of a non-relativistic system of electrons where the nuclei are fixed is

$$E_v[\mathbf{n}] = F[\mathbf{n}] + \int \mathrm{d}^3\mathbf{r}\left(n_\uparrow(\mathbf{r}) + n_\downarrow(\mathbf{r})\right)v(\mathbf{r}). \tag{23}$$

Its functional derivatives are

$$\frac{\delta E_v}{\delta n_\sigma(\mathbf{r})}[\mathbf{n}] = \mu_\sigma[\mathbf{n}] + v(\mathbf{r}) - u_\sigma[\mathbf{n}](\mathbf{r}),$$

and if we demand that the r.h.s. of the above equation is μ_σ we obtain the ground-state problem

$$u_\sigma[\mathbf{n}](\mathbf{r}) = v(\mathbf{r}).$$

This problem remains the same when the number of electrons of the system is integer (where the derivative discontinuity is present). Let us denote $\mathbf{n}^*$ as the ground-state density of the system. The ground state of each isolated molecule is usually (but not always) singlet or doublet. For the former, the ground-state spin-densities are equal, $n_\uparrow^* = n_\downarrow^*$. When the system is a doublet, however, there are infinitely many spin-densities that yield the same ground-state energy.

The Kohn–Sham potential is split into two components, $\mathbf{u}_\mathrm{s} = \mathbf{u} + \mathbf{u}_\mathrm{HXC}$. The ground-state density problem, using non-interacting E2VR potentials, is

$$u_{\mathrm{s},\sigma}\left[\mathbf{n}^*\right] = u_{\mathrm{HXC},\sigma}\left[\mathbf{n}^*\right] + v. \tag{24}$$

Comparing the above equation with (23) we identify, in general,

$$\frac{\delta E_{\mathrm{HXC},\sigma}}{\delta n_\sigma(\mathbf{r})}[\mathbf{n}] = u_{\mathrm{HXC},\sigma}[\mathbf{n}](\mathbf{r}).$$

Equation (24) represents a fixed-point problem, and requires the solution of the KS equations self-consistently. As mentioned previously, the functional derivatives are

not unique. Hence, we must set the potential that represents our system, such that the functional derivatives can be taken. This is accomplished in approximate SDFT by choosing an explicit differentiable form of the XC energy. The formulation of approximated XC energy functionals is mainly based on the following criteria: (1) the approximate XC energy functional must satisfy known formal properties (scaling, bounds, size-consistency, etc.) or (2) a suitable functional with adjustable parameters is defined for the XC functional, and the parameters are fit by minimizing the error in the estimation of molecular properties with respect to experiments and/or ab initio calculations.

The problem of representing the 2-density with a system of non-interacting electrons can also be written in terms of the KS equations:

$$\left(-\frac{1}{2}\nabla_{\mathbf{r}}^2 + u_{\mathrm{HXC},\sigma}(\mathbf{r},\ [\mathbf{n}^*]) + v(\mathbf{r})\right)\phi_{\mathbf{k},\sigma}(\mathbf{r},\ [\mathbf{n}^*]) = \epsilon_{\mathbf{k},\sigma}[\mathbf{n}^*]\,\phi_{\mathbf{k},\sigma}(\mathbf{r},\ [\mathbf{n}^*]).$$

Using the decomposition of the exact energy, (23) and (21), and taking the functional derivative of E_v, we identify the HOMO energy as

$$\mu_\sigma[\mathbf{n}] = \epsilon_{\mathrm{H},\sigma}[\mathbf{n}],$$

But

for this result to hold the numbers of spin-electrons must be non-integer. The exchange-correlation potential displays a discontinuity with respect to the number of electrons. It is, however, always convenient to set the KS potential as zero in the asymptotics. In Coulomb systems this guarantees that the numerical value of the HOMO has a physical meaning. For example, when the number of electrons is an integer and the ground state is a singlet, the HOMOs in both the $\uparrow$ and $\downarrow$ channels yield the ionization of the system. For a doublet ground state, the HOMO of the mostly populated channel is the ionization of the system, and the HOMO of the other channel is an ionization with respect to an excited ion, the triplet of lowest energy [48]. The ionization of the less-populated channel is not the one given by the chemical potential of such a channel. Therefore, a constant must be added to the KS potential of the less populated channel to compensate. In general, the use of ensembles eliminates the arbitrary constant in the potential if the numbers of electrons are non-integers [15].

A simple form of the total energy is

$$E_v[\mathbf{n}] = \sum_{\mathrm{M}\in V(\Delta[n])} w_{\mathrm{M}} E_v[\mathbf{n}_{\mathrm{M}}].$$

This form does not use E_{HXC} evaluated at the density $\mathbf{n}$ which integrates to non-integer number of electrons. If some density-functional approximation is used for the discrete-electron states described by $E[\mathbf{n}_{\mathbf{M}}]$, then the piecewise linearity and the proper derivative discontinuity are recovered. Nevertheless, this replacement does not solve the dissociation problem.

The main argument in support of the inclusion of the piecewise linearity is the example provided in the work of Perdew et al. [8] (PPLB). They considered two separated atoms, where the energy of the system, in practice, is additive, and the energy as a function of the charge transfer amount displays a *piecewise linearity* and a *derivative discontinuity*. If an approximation such as the local-density approximation were used to approximate the energy of each atom, one would obtain a deviation from the piecewise linearity. With the methods shown here, however, we can recover the linearity and the derivative discontinuity in such a hypothetical case of infinitely separated atoms. How can we thus extend this idea of using nearly additive energies to describe bond stretching? Partition-DFT [16, 49, 50] answers this question.

In the presence of two nuclei, the 1-body external potential is of the form $v_{\mathrm{L}} + v_{\mathrm{R}}$. The 2-density, when both nuclei are well separated, is the sum of two 2-densities, $\mathbf{n}_{\mathrm{L}} + \mathbf{n}_{\mathrm{R}}$; as the distance increases the energy becomes more additive. In general, one can split the 1-body external potential as the sum of potentials, that is:

$$v(\mathbf{r}) = \sum_{\alpha} v_{\alpha}(\mathbf{r}).$$

Define an energy functional of the form

$$E^{+}[\mathbf{n};\ \{v_{\alpha}\}] = \inf\left\{ \sum_{\alpha} E_{v_{\alpha}}[\mathbf{n}_{\alpha}] \,\middle|\, \sum_{\alpha} \mathbf{n}_{\alpha} = \mathbf{n} \right\}.$$

This is an additive energy, under a constraint: the sum of all the $\mathbf{n}_{\alpha}$ must yield a preset total 2-density, i.e., that of the ground state. The functionals $\{E_{v_{\alpha}}[\mathbf{n}_{\alpha}]\}$ display the required non-analytic properties. The functional E^{+} derives from the PPLB picture of adiabatic charge transfer between atoms, and is applicable to the molecular fragments. Using the formalism outlined in this chapter, E^{+} can be calculated. The functional E^{+} is not the true energy of the system, except when the distance between the fragments is very large. Because it is a functional of the total density, comparison with E_v leads to the partition energy functional:

$$E_{\mathrm{P}}[\mathbf{n};\ v, \{v_{\alpha}\}]\ := E_v[\mathbf{n}] - E^{+}[\mathbf{n};\ \{v_{\alpha}\}].$$

This functional accounts for the correlation effects between the fragments defined across the molecule, and it is the energy required to place the fragments in contact while maintaining the total density fixed.

When the distance between fragments is finite, the averaging of energies is taken over configurations where the fragments are discrete, in terms of the number of electrons. Hence, the non-analytic, additive energy is constructed, E^{+}, by ensemble-averaging; the partition energy functional is the missing object. An alternative route is thus offered: approximate the partition energy. This requires analysis of formal limits, comparison with multiconfigurational theories, and tuning of the coupling

strength of the fragments, using available, accurate ab initio calculations. For the case of H_2 this idea has been successful. Because it is easier to restrict the configurations to be discrete with PDFT, the Kohn–Sham potential displays naturally the desired barrier in the bonding region [51].

10 Conclusions

We analyzed the non-analyticity of density functionals for the spin-polarized case. In agreement with unpolarized DFT, the integer discontinuity also arises when the energy is minimized for systems under a collinear, position-dependent, magnetic field, and is expressed in terms of the number of spin-electrons. It must be borne in mind that the total and KS spin-density functionals present non-unique functional derivatives with respect to the spin-densities. Nevertheless, around a chosen 2-potential that represents the spin-densities of integer spin-electrons there are derivative discontinuities with physical interpretation. The construction of approximations with piecewise linearity which we have illustrated here can be used as exact conditions to develop functionals within SDFT, and extensions where the system of interest is decomposed into composite fragments.

Acknowledgements We thankfully acknowledge support by the National Science Foundation CAREER program under Grant No. CHE-1149968. A.W. also acknowledges support from an Alfred P. Sloan Foundation Research Fellowship and a Camille-Dreyfus Teacher Scholar award.

Appendix. Calculation of Functional Derivatives

The first term on the r.h.s. of (13) can be calculated using the chain rule:

$$\delta_m E\big(\overline{N},\,[v]\big) = \left.\frac{\partial E}{\partial N}\right|_u \delta_m N + \delta_w E\big(\overline{N},\,[v]\big)\big|_N,$$

where $|_u$ and $|_N$ mean that these objects are fixed. The second term of the r.h.s. is interpreted as

$$\begin{aligned}\lim_{\varepsilon\to 0}\frac{E\big(\overline{N},\,\big[u\big[n'\big]\big]\big) - E\big(\overline{N},\,[u[n]]\big)}{\varepsilon} &= \lim_{\varepsilon\to 0}\frac{E\big(\overline{N},\,[v+\varepsilon]\big) - E\big(\overline{N},\,[v]\big)}{\varepsilon}\\ &= \delta_w E\big(\overline{N},\,[v]\big).\end{aligned}$$

The differential of N is trivial:

$$\delta_m N = \lim_{\varepsilon \to 0} \frac{N[n + \varepsilon m] - N[m]}{\varepsilon} = \int m = 1.$$

When N is fixed, $|\psi\rangle$ is only determined by u, then

$$\delta_w E|_N = \delta_w \langle \psi | \hat{H}_0 | \psi \rangle \big|_N + \delta_w \int \mathrm{d}^3\mathbf{r}\, n(\mathbf{r}) u(\mathbf{r}) \big|_N$$
$$= \left(\delta_w \langle \psi | \right) \hat{H}_0 | \psi \rangle + \text{c.c.} + \int \mathrm{d}^3\mathbf{r}\, n(\mathbf{r}) w(\mathbf{r}) + \int \mathrm{d}^3\mathbf{r}\, u(\mathbf{r}) \delta_w n(\mathbf{r}).$$

Note that

$$\delta_w n(\mathbf{r})|_N = \sum_{M \in \mathbb{N}} y\left(\overline{N} - M\right) \left[\left(\delta_w \langle \psi_M | \right) \hat{n}(\mathbf{r}) | \psi_M \rangle \big|_N + \text{c.c.} \right].$$

This allows us to express $\delta_w E|_N$ as

$$\delta_w E\big|_N = \sum_{M \in \mathbb{N}} y\left(\overline{N} - M\right) \left[\left(\delta_w \left\langle \psi | \hat{H}_0 + \int \mathrm{d}^3\mathbf{r}\, \hat{n}(\mathbf{r}) u(\mathbf{r}) | \psi \right\rangle + \text{c.c.} \right] + \int \mathrm{d}^3\mathbf{r}\, n(\mathbf{r}) w(\mathbf{r}) \right.$$
$$= \sum_{M \in \mathbb{N}} y\left(\overline{N} - M\right) E_0(M, [u]) \delta_w \left(\langle \psi_M | \psi_M \rangle \right) + \int \mathrm{d}^3\mathbf{r}\, n(\mathbf{r}) w(\mathbf{r})$$
$$= \int \mathrm{d}^3\mathbf{r}\, n(\mathbf{r}) w(\mathbf{r}).$$

Noting that $\delta_m u[n] = w$, we get (14).

References

1. Kohn W, Sham LJ (1965) Self-consistent equations including exchange and correlation effects. Phys Rev 140(4A):A1133
2. Perdew J, Kurth S (2003) Density functionals for non-relativistic Coulomb systems in the new century. In: Fiolhais C, Nogueira F, Marques M (eds) A primer in density functional theory, volume 620 of lecture notes in physics, chapter 1. Springer, Berlin, pp 1–55. ISBN 978-3-540-03083-6
3. Elliott P, Lee D, Cangi A, Burke K (2008) Semiclassical origins of density functionals. Phys Rev Lett 100:256406
4. Lee D, Furche F, Burke K (2010) Accuracy of electron affinities of atoms in approximate density functional theory. J Phys Chem Lett 1(14):2124–2129
5. Csonka GI, Perdew JP, Ruzsinszky A (2010) Global hybrid functionals: a look at the engine under the hood. J Chem Theory Comput 6(12):3688–3703
6. Ruzsinszky A, Perdew JP (2011) Twelve outstanding problems in ground-state density functional theory: a bouquet of puzzles. Comput Theor Chem 963(1):2–6

7. Gunnarsson O, Lundqvist BI (1976) Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism. Phys Rev B 13(10):4274
8. Perdew JP, Parr RG, Levy M, Balduz JL Jr (1982) Density-functional theory for fractional particle number: derivative discontinuities of the energy. Phys Rev Lett 49(23):1691
9. Cohen AJ, Mori-Sánchez P, Yang W (2012) Challenges for density functional theory. Chem Rev 112(1):289–320
10. Perdew J (1985) What do the Kohn-Sham orbital energies mean? How do atoms dissociate? In: Dreizler RM, da Providência J (eds) Density functional methods in physics. Plenum, New York
11. Tempel DG, Martinez TJ, Maitra NT (2009) Revisiting molecular dissociation in density functional theory: a simple model. J Chem Theory Comput 5(4):770–780
12. Miller JS, Epstein AJ, Reiff WM (1988) Ferromagnetic molecular charge-transfer complexes. Chem Rev 88(1):201–220
13. Kollmar C, Kahn O (1993) Ferromagnetic spin alignment in molecular systems: an orbital approach. Acc Chem Res 26(5):259–265
14. Bogani L, Wernsdorfer W (2008) Molecular spintronics using single-molecule magnets. Nat Mater 7(3):179–186
15. Mosquera MA, Wasserman A (2014) Integer discontinuity of density functional theory. Phys Rev A 89(5):052506
16. Mosquera MA, Wasserman A (2013) Partition density functional theory and its extension to the spin-polarized case. Mol Phys 111(4):505–515
17. Perdew JP (1985) Density functional theory and the band gap problem. Int J Quantum Chem 28(S19):497–523
18. Harriman JE (1981) Orthonormal orbitals for the representation of an arbitrary density. Phys Rev A 24(2):680
19. Levy M (1979) Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. Proc Natl Acad Sci U S A 76(12):6062–6065
20. Lieb EH (1983) Density functionals for coulomb systems. Int J of Quantum Chem 24:243–277
21. Valone SM (1980) Consequences of extending 1-matrix energy functionals from pure-state representable to all ensemble representable 1 matrices. J Chem Phys 73(3):1344–1349
22. Hille E, Phillips RS (1957) Functional analysis and semi-groups, vol 31. American Mathematical Society, Providence
23. Hormander L (1994) Notions of convexity, vol 12. Springer, Birkhäuser, Boston
24. Lammert PE (2007) Differentiability of Lieb functional in electronic density functional theory. Int J Quantum Chem 107(10):1943–1953
25. Englisch H, Englisch R (1984) Exact density functionals for ground-state energies. I. General results. Phys Status Solidi B 123(2):711–721
26. Englisch H, Englisch R (1984) Exact density functionals for ground-state energies II. Details and remarks. Phys Status Solidi B 124(1):373–379
27. Chayes JT, Chayes L, Ruskai MB (1985) Density functional approach to quantum lattice systems. J Stat Phys 38(3–4):497–518
28. Ayers PW (2006) Axiomatic formulations of the Hohenberg-Kohn functional. Phys Rev A 73 (1):012513
29. Kvaal S, Ekström U, Teale AM, Helgaker T (2014) Differentiable but exact formulation of density-functional theory. J Chem Phys 140(18):18A518
30. van Leeuwen R (1998) Phys Rev Lett 80(6):1280
31. Vignale G (2008) Phys Rev A 77:062511
32. Mosquera MA (2013) Action formalism of time-dependent density-functional theory. Phys Rev A 88:022515
33. Rabuck AD, Scuseria GE (1999) Improving self-consistent field convergence by varying occupation numbers. J Chem Phys 110(2):695–700

34. Cancès E (2001) Self-consistent field algorithms for Kohn–Sham models with fractional occupation numbers. J Chem Phys 114(24):10616–10622
35. Capelle K, Vignale G (2001) Nonuniqueness of the potentials of spin-density-functional theory. Phys Rev Lett 86(24):5546
36. Gidopoulos NI (2007) Potential in spin-density-functional theory of noncollinear magnetism determined by the many-electron ground state. Phys Rev B 75:134408
37. Kutzelnigg W, Morgan JD III (1996) Hund's rules. Z Phys D Atom Mol Cl 36(3–4):197–214
38. Boyd RJ (1984) A quantum mechanical explanation for Hund's multiplicity rule. Nature 310:480–481
39. Kollmar H, Staemmler V (1978) Violation of Hund's rule by spin polarization in molecules. Theor Chim Acta 48(3):223–239
40. Botch BH, Dunning TH Jr (1982) Theoretical characterization of negative ions. Calculation of the electron affinities of carbon, oxygen, and fluorine. J Chem Phys 76(12):6046–6056
41. Feller D, Davidson ER (1985) Ab initio multireference CI determinations of the electron affinity of carbon and oxygen. J Chem Phys 82(9):4135–4141
42. Cramer CJ, Truhlar DG (2009) Density functional theory for transition metals and transition metal chemistry. Phys Chem Chem Phys 11(46):10757–10816
43. Balawender R (2012) Thermodynamic extension of density-functional theory. III. Zero-temperature limit of the ensemble spin-density functional theory. arXiv preprint arXiv:1212.1367
44. Malek AM, Balawender R (2013) Discontinuities of energy derivatives in spin-density functional theory. arXiv preprint arXiv:1310.6918
45. von Barth U, Hedin L (1972) A local exchange-correlation potential for the spin polarized case. J Phys C Solid State Phys 5(13):1629
46. Ullrich CA, Kohn W (2001) Kohn-Sham theory for ground-state ensembles. Phys Rev Lett 87 (9):093001
47. Vydrov OA, Scuseria GE, Perdew JP (2007) Tests of functionals for systems with fractional electron number. J Chem Phys 126(15):154109
48. Almbladh C-O, Von Barth U (1985) Exact results for the charge and spin densities, exchange-correlation potentials, and density-functional eigenvalues. Phys Rev B 31(6):3231
49. Cohen MH, Wasserman A (2007) On the foundations of chemical reactivity theory. J Phys Chem A 111(11):2229–2242
50. Elliott P, Burke K, Cohen MH, Wasserman A (2010) Partition density-functional theory. Phys Rev A 82(2):024501
51. Nafziger J, Wasserman A (2012) Delocalization and static correlation in partition density-functional theory. arXiv preprint arXiv:1305.4966

Top Curr Chem (2015) 365: 175–186
DOI: 10.1007/128_2014_581

Published online: 3 December 2014

Fractional Kohn–Sham Occupancies from a Strong-Correlation Density Functional

Axel D. Becke

Abstract It is not always possible in Kohn–Sham density-functional theory for the non-interacting reference state to have integer-only occupancies. Cases of "strong" correlation, with very small HOMO-LUMO gaps, involve fractional occupancies. At the transition states of symmetric avoided-crossing reactions, for example, representation of the correct density requires a 50/50 mixing of degenerate HOMOs. In a recent paper (Becke, J Chem Phys 139:021104, 2013) the "B13" strong-correlation density functional of Becke (J Chem Phys 138:074109, 2013 and 138:161101, 2013) was shown to give excellent barrier heights in symmetric avoided-crossing reactions. However, the calculations were performed only at reactant and transition-state geometries, where the fractional HOMO-LUMO occupancies in the latter are 50/50 by symmetry. In the present chapter, we compute full reaction curves for avoided crossings in $H_2 + H_2$, ethylene (twisting around the double bond), and cyclobutadiene (double-bond automerization) by determining fractional occupancies variationally. We adopt a practical strategy for doing so which does not involve self-consistent B13 computations (not yet possible) and involves minimal cost. Single-bond dissociation curves for H_2 and LiH are also presented.

Keywords Density-functional theory · Electronic structure · Strong correlation

Contents

A.D. Becke (✉)
Department of Chemistry, Dalhousie University, 6274 Coburg Rd., P.O. Box 15000, Halifax NS, B3H 4R2, Canada
e-mail: axel.becke@dal.ca

1 Introduction

Kohn–Sham density-functional theory (KS-DFT) is based on a non-interacting, single-Slater-determinant reference state having the *same total density* ρ as the "real" interacting system [1–3]. The KS reference orbitals ψ_i satisfy the single-particle Schroedinger equation

$$-\frac{1}{2}\nabla^2\psi_i + v_{KS}\psi_i = e_i\psi_i \tag{1}$$

where v_{KS} is the non-interacting potential such that

$$\rho = \sum_i n_i\psi_i^2, \quad n_i = 0,\ 1,\ \text{or } 2 \tag{2}$$

equals the density of the real system. The KS potential is given by [2]

$$v_{KS} = v_{ext} + v_{el} + v_{XC}, \quad v_{el} = \int \frac{\rho(r_2)}{r_{12}} d^3r_2, \quad v_{XC} = \frac{\delta E_{XC}}{\delta\rho} \tag{3}$$

where v_{ext} is the external (one-body) potential in the real system, v_{el} is the classical Coulomb potential arising from the electron density, and v_{XC} is the functional derivative [3] with respect to the density of the "exchange-correlation" energy E_{XC} defined by

$$E_{total} = -\frac{1}{2}\sum_i n_i \int \psi_i \nabla^2\psi_i + \int \rho v_{ext} + \frac{1}{2}\iint \frac{\rho(r_1)\rho(r_2)}{r_{12}} + E_{XC} \tag{4}$$

i.e., that part of the total energy containing all the quantum and correlation effects.

We restrict ourselves in this work to singlet states. Thus the occupation numbers n_i in (2) and (4) are 0 or 2 if the reference state is a *single* Slater determinant. Not all quantum systems, however, can be referred to a single determinant or "configuration". Consider a four-electron system in a perfectly square D_{4h} nuclear framework: square H_4, for example, or the pi-electrons in square cyclobutadiene C_4H_4. The orbital energies are as sketched in Fig. 1, with a doubly degenerate HOMO set, ψ_a and ψ_b. Two equivalent singlet configurations (determinants) are possible: $|\psi_a^2\rangle$ and $|\psi_b^2\rangle$ in the figure. These interact strongly to produce the two-configuration mixtures $\Psi_\pm$:

Fig. 1 D_{4h} Slater determinants

$$\Psi_{\pm} = \frac{1}{\sqrt{2}}\left(|\psi_{\mathrm{a}}^2\rangle \pm |\psi_{\mathrm{b}}^2\rangle\right) \tag{5}$$

with Ψ_- being the ground state. Neither of the individual determinants, $|\psi_{\mathrm{a}}^2\rangle$ or $|\psi_{\mathrm{b}}^2\rangle$, is an adequate reference state. Indeed, neither has a density having even the correct D_{4h} *symmetry* of the actual ground-state density (if we constrain ourselves to *real* orbitals). The actual ground-state ρ is represented by

$$\rho = 2\sum_{i=1}^{\mathrm{HOMO}-1} \psi_{\mathrm{i}}^2 + \psi_{\mathrm{a}}^2 + \psi_{\mathrm{b}}^2 \tag{6}$$

involving occupancies *other than* 2 for the degenerate HOMO orbitals. In each of ψ_{a} and ψ_{b}, there is *half* an electron of spin up and *half* an electron of spin down.

The orbital occupancies in the above example are fixed by symmetry: i.e., a doubly degenerate HOMO level yielding two symmetry-equivalent reference configurations. In general, strongly correlated systems have densities of the form

$$\rho = 2\sum_{i=1}^{\mathrm{HOMO}-1} \psi_{\mathrm{i}}^2 + 2(1-f)\psi_{\mathrm{HOMO}}^2 + 2f\psi_{\mathrm{LUMO}}^2 \tag{7}$$

where f may have *any* value in the interval $0 \leq f \leq 1$, depending on the HOMO-LUMO gap and the relevant interaction matrix elements. The f parameter in (7), and the corresponding fractional occupancies in (4), must be determined variationally.

An analysis by Schipper et al. [4] of the $H_2 + H_2$ reaction is a beautiful illustration. These authors have computed accurate CI and MRCI wavefunctions at various points on the $H_2 + H_2$ potential-energy surface and, at each geometry, have extracted the exact Kohn–Sham potential v_{KS} and orbitals ψ_{i} using a robust Kohn–Sham inversion procedure [5]. For geometries close to perfect squares (D_{4h}), their inversion procedure fails to converge; i.e., occupation "holes" appear under the HOMO if single-determinant occupancies are enforced. Stable KS solutions are obtained *only if fractional* occupancies are allowed in (7). Geometries far from D_{4h} admit conventional integer-occupancy KS solutions. Standard exchange-correlation GGAs (Generalized Gradient Approximations) also produce fractional

occupancies near D_{4h} geometries, but with rather poor accuracy compared to the exact Kohn–Sham results [4].

In this chapter, a recently developed [6, 7] strong-correlation density functional, "B13", is benchmarked on the Schipper et al. [4] $H_2 + H_2$ data, and also on ethylene torsion and cyclobutadiene automerization data from Jiang et al. [8]. All three of these are challenging avoided-crossing problems involving fractional occupancies near their transition states. The B13 functional is reviewed in Sect. 2. Variational optimization of B13 fractional occupancies, and reaction profiles for our three avoided-crossing tests, are discussed in Sect. 3. Dissociation curves for the single bonds H_2 and LiH are presented in Sect. 4. Concluding remarks and future prospects are discussed in Sect. 5.

2 The B13 Strong-Correlation Density Functional

In recent papers [6, 7] we have introduced a correlation-energy density functional able to describe both moderately and strongly correlated systems. "Strong" correlation arises from a small HOMO–LUMO gap, resulting in strong mixing of the configurations

$$\left|\cdots\psi^2_{\mathrm{HOMO}}\psi^0_{\mathrm{LUMO}}\right\rangle \text{ and } \left|\cdots\psi^0_{\mathrm{HOMO}}\psi^2_{\mathrm{LUMO}}\right\rangle \tag{8}$$

as described above. The signature of strong correlation in KS-DFT is *fractional occupancies* in the minimum-energy KS density (see (7)). At the same time, Janak's theorem [3, 9] implies that

$$e_{\mathrm{HOMO}} = e_{\mathrm{LUMO}} \tag{9}$$

in any such strongly-correlated, fractionally-occupied, Kohn–Sham minimum.

Strong correlation is often discussed in the context of molecular bond dissociation; i.e., the correlation energy required to dissociate molecular bonds using *spin-restricted* orbitals. Small HOMO–LUMO gaps and strong mixing of configurations play a major role here as well. Spin-restricted dissociation of molecules implies

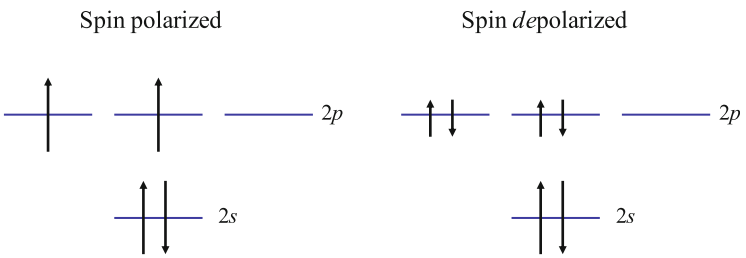


Fig. 2 Carbon atom spin states

that, for *all free atoms*, the usual Hund's-rule spin-polarized configuration has the *same energy* as the spin-*de*polarized configuration obtained by replacing every unpaired electron by half an electron of spin up and half an electron of spin down (as sketched for the carbon atom in Fig. 2). This is a stringent test of density functionals. Conventional GGAs fail this test [7] with errors of the order of 40 kcal/mol in the first three rows of the periodic table. Our B13 functional, however, is designed to capture this spin-invariance property [7] while maintaining good performance in standard thermochemical applications.

B13 is based on *exact* exchange and is hence a *pure correlation* theory. The total exchange-correlation energy has the form [6, 7]

$$E_{\mathrm{XC}}^{\mathrm{B13}} = E_{\mathrm{X}}^{\mathrm{exact}} + E_{\mathrm{C}}^{\mathrm{B13}} + \Delta E_{\mathrm{strong\,C}}^{\mathrm{B13}} \tag{10}$$

where $E_{\mathrm{X}}^{\mathrm{exact}}$ is the exactly-computed Kohn–Sham (or Hartree–Fock depending on the implementation) exchange energy, $E_{\mathrm{C}}^{\mathrm{B13}}$ is a sum of opposite- and parallel-spin *static* and *dynamical* correlation potential energies:

$$E_{\mathrm{C}}^{\mathrm{B13}} = a_{\mathrm{stat\,C}}^{\mathrm{opp}} U_{\mathrm{stat\,C}}^{\mathrm{opp}} + a_{\mathrm{stat\,C}}^{\mathrm{par}} U_{\mathrm{stat\,C}}^{\mathrm{par}} + a_{\mathrm{dyn\,C}}^{\mathrm{opp}} U_{\mathrm{dyn\,C}}^{\mathrm{opp}} + a_{\mathrm{dyn\,C}}^{\mathrm{par}} U_{\mathrm{dyn\,C}}^{\mathrm{par}} \tag{11}$$

and $\Delta E_{\mathrm{strong\,C}}^{\mathrm{B13}}$ is a *strong-correlation correction* given by

$$\Delta E_{\mathrm{strong\,C}}^{\mathrm{B13}} = \sum_{n=2}^{N} c_n \int x^n u_{\mathrm{C}} d^3 r. \tag{12}$$

To an excellent first approximation, the four prefactors in (11) are all equal to each other with optimum value 0.62 (see Becke [6]). Nevertheless, greater accuracy can be achieved by fitting these independently. In (12), u_{C} is the sum of the *integrands* of the four terms in (11):

$$u_{\mathrm{C}} = u_{\mathrm{stat\,C}}^{\mathrm{opp}} + u_{\mathrm{stat\,C}}^{\mathrm{par}} + u_{\mathrm{dyn\,C}}^{\mathrm{opp}} + u_{\mathrm{dyn\,C}}^{\mathrm{par}}, \tag{13}$$

namely the static + dynamical correlation potential-energy *density*, and x is the following ratio of the *static* correlation potential-energy density to the total:

$$x = \frac{u_{\mathrm{stat\,C}}^{\mathrm{opp}} + u_{\mathrm{stat\,C}}^{\mathrm{par}}}{u_{\mathrm{C}}}. \tag{14}$$

This is a reasonable, local measure of strong correlation. In atoms, where strong correlation is insignificant, $x \to 0$. In stretched, spin-restricted H_2, on the other hand, the quintessential case of strong correlation, $x \to 1$. We assume that intermediate situations are well characterized by intermediate values of x. The c_n in (12) are polynomial expansion coefficients fit to the atomic spin-depolarization condition of the previous paragraph.

Our fit set in Becke [7] consisted of the "G2/97" atomization energies of Curtiss, Raghavachari, and Pople [10] plus the spin-depolarization condition on all open-shell atoms $Z < 36$, including transition metal atoms. Two terms were deemed optimum in the strong-correlation correction, (12). The best-fit coefficients are

$$\begin{array}{ll} c_{\text{stat C}}^{\text{opp}} = 0.552 & c_{\text{stat C}}^{\text{par}} = 0.844 \\ c_{\text{dyn C}}^{\text{opp}} = 0.640 & c_{\text{dyn C}}^{\text{par}} = 0.559 \\ c_2 = 0.825 & c_3 = -0.380 \end{array} \tag{15}$$

and these are used in the present work. Mean absolute B13 errors on the G2/97 atomization energies and the atomic spin-depolarization tests are 3.8 and 5.7 kcal/mol, respectively [7].

It should be noted that the B13 correlation model is based on a *single* Slater determinant as the reference state [6]. The model adds multi-center correlations to *single*-determinantal pair densities as the starting point. How then, can B13 be justified in fundamentally two-configuration problems such as avoided crossings? Consider again the 50/50 mixtures at symmetric avoided-crossing transition states [11]. Imagine replicating the system of interest by an identical system at infinite distance from the first. The degenerate orbitals ψ_{a} and ψ_{b} on the system (A) and its replicant (B) combine to form the degenerate "super" system orbitals

$$\psi_{\text{a}\pm}^{\text{super}} = \frac{1}{\sqrt{2}}\left(\psi_{\text{a}}^{\text{A}} \pm \psi_{\text{a}}^{\text{B}}\right) \text{ and } \psi_{\text{b}\pm}^{\text{super}} = \frac{1}{\sqrt{2}}\left(\psi_{\text{b}}^{\text{A}} \pm \psi_{\text{b}}^{\text{B}}\right). \tag{16}$$

Now consider the *single* supersystem Slater determinant with occupancy

$$\left|\cdots \psi_{\text{a}+}^{\text{super}}(\uparrow\downarrow)\psi_{\text{b}+}^{\text{super}}(\uparrow\downarrow)\right\rangle. \tag{17}$$

In each *sub*system, this determinant places *one* electron in ψ_{a} (half spin-up and half spin-down) and similarly *one* electron in ψ_{b}, precisely the density of (6) corresponding to a 50/50 configuration mixture.

The more general densities of (7) can be reproduced by single Slater determinants spanning supersystems of *more* than two replicas. Appropriate numbers of replicants and appropriate occupations of supersystem orbitals can reproduce *any* two-configuration density. Because B13 *maximizes* multi-center correlation, it delivers the *ground state* energy corresponding to any such two-configuration density, *not* an excited-state energy or an ensemble energy. Excited states supported by the HOMO–LUMO orbitals will be studied in future work.

3 Avoided-Crossing Reaction Curves

In Becke [11], B13 was tested on barrier heights of four symmetric avoided-crossing reactions: H_4, ethylene double-bond twisting, and double-bond automerization in cyclobutadiene and cyclooctatetraene. At the transition states of these reactions, there is 50/50 HOMO–LUMO mixing by symmetry. Variational determination of fractional occupancies was not needed. In this chapter, however, we draw full reaction curves for the first three of these reactions. Thus fractional occupancies must be computed.

Self-consistent B13 calculations are not yet possible, though prospects for SCF-B13 are good; see Proynov et al. [12–14] for an SCF implementation, and Arbuznikov and Kaupp [15] for an OEP implementation of an earlier B13 variant known as B05 [16]. Our current calculations are "post-LDA." All total energies are evaluated using LDA orbitals computed by the grid-based NUMOL program [17, 18]. The optimum post-LDA fractional occupancies in (7) are determined by searching in the range $0 \leq f \leq 1$ for the minimum-energy f value. We adopt the simple strategy of minimizing the energy of a three-point quadratic interpolation with calculations performed at $f = 0$, 1/2, and 1. For comparison purposes, the same post-LDA three-point approach is applied to the B88 [19] + PBE [20] exchange-correlation GGA.

In our first avoided-crossing reaction, $H_2 + H_2$, we consider the nine H_4 geometries in Table II of Schipper et al. [4]. Each geometry is a rectangle with the longer side denoted R and the shorter side denoted r. Table 1 lists these geometries, along with the fractional occupancies f for "exact" Kohn–Sham, for the B88 + PBE exchange-correlation GGA, and for B13. We find qualitative agreement between our B88 + PBE GGA results and the B88 + LYP [21] GGA results in Schipper et al. [4] (also recorded in Table1), but both GGA sets agree poorly with exact Kohn–Sham. There is qualitative agreement, however, between our B13 fractional occupancies and the exact KS values, although B13 appears to be somewhat too strongly correlated. Note that the region of fractional occupancies near the square

Table 1 Fractional occupancies as in (7) for the $H_2 + H_2$ reaction

R	r	Exact KS	B88 + LYP	B88 + PBE	B13
2.32	2.32	0.50	0.50	0.50	0.50
2.35	2.29	0.32	0.10	0.07	0.45
2.40	2.21	0.07	0.00	0.00	0.33
2.50	2.06	0.00	0.00	0.00	0.11
2.75	1.68	0.00	0.00	0.00	0.00
3.00	1.44	0.00	0.00	0.00	0.00
4.00	1.41	0.00	0.00	0.00	0.00
5.00	1.40	0.00	0.00	0.00	0.00
10.0	1.40	0.00	0.00	0.00	0.00

Geometry parameters R and r in bohr. Exact KS and B88 + LYP results from Schipper et al. [4]. B88 + PBE and B13 results from the present work

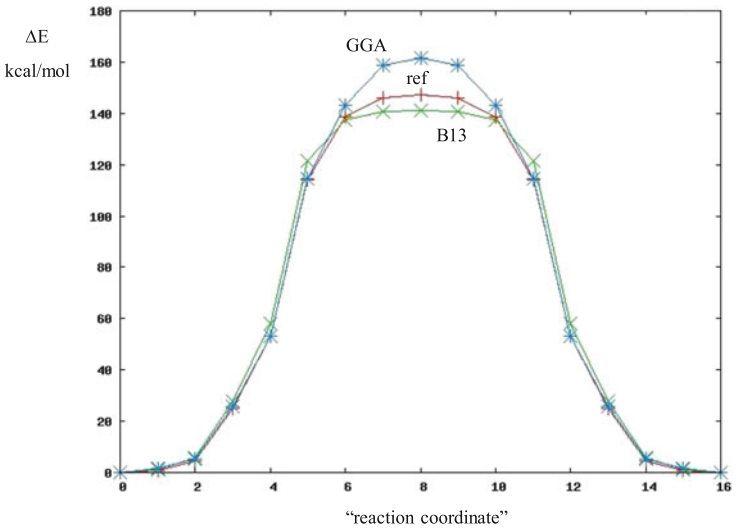


Fig. 3 Reaction profile for $H_2 + H_2$

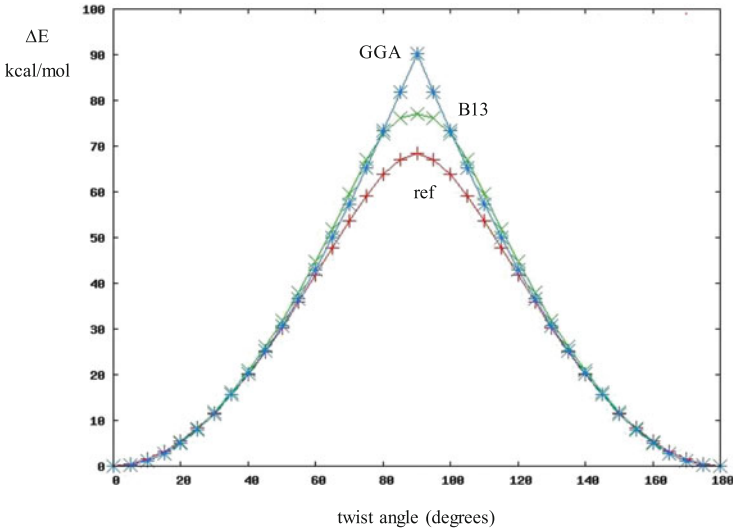


Fig. 4 Reaction profile for ethylene twist

transition state is smaller for the GGA than for exact KS and B13. The exact KS, GGA, and B13 energy curves are plotted in Fig. 3. The energy zero for each curve is the energy at geometry $R = 10.0$ and $r = 1.40$ bohr. The B13 reaction barrier of 141.2 kcal/mol is in fair agreement with the MRCI barrier of 147.6 kcal/mol.

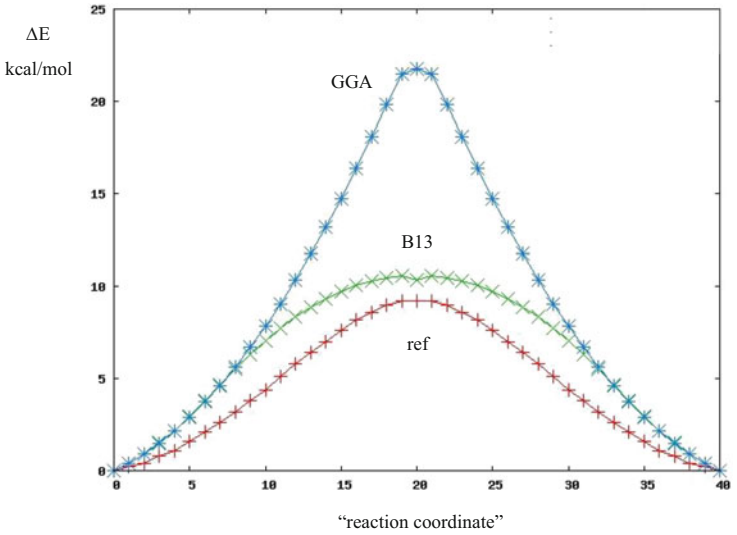


Fig. 5 Reaction profile for cyclobutadiene automerization

Our second avoided-crossing test is twisting around the double bond in ethylene. We took MR-ccCA (Multi-Reference Correlation Consistent Composite Approach) reference data from Jiang, Jeffrey, and Wilson [8] and have performed all calculations at the geometries specified in their paper. Figure 4 plots reference, B88 + PBE GGA, and B13 reaction curves. Zero energy for each method is the planar ethylene equilibrium geometry. In this case, the GGA energies always minimize at $f = 0$ (no HOMO–LUMO mixing) producing an unphysical cusp at the top of the GGA curve. The B13 curve, on the other hand, is smooth throughout. The B13 barrier of 77.1 kcal/mol agrees better with the MR-ccCA barrier of 68.3 kcal/mol than does the GGA barrier, 90.2 kcal/mol.

Double-bond automerization in cyclobutadiene is an especially challenging avoided-crossing test. The reaction profile is plotted in Fig. 5, with MR-ccCA reference data and geometries again from Jiang, Jeffrey, and Wilson [8]. Zeroes of all curves correspond to the rectangular equilibrium geometry. The GGA curve is cuspless in this case but has a barrier more than twice too large (21.8 kcal/mol compared to 9.2 kcal/mol for MR-ccCA). The B13 barrier of 10.4 kcal/mol compares quite well with the 9.2 kcal/mol MR-ccCA barrier.

4 H_2 and LiH Dissociation Curves

Spin-restricted KS-DFT dissociation curves, even employing sophisticated GGA and hybrid functionals, have asymptotes well above the exact dissociation limits. Spin-*un*restricted calculations ("UKS") can capture the correct limits. UKS

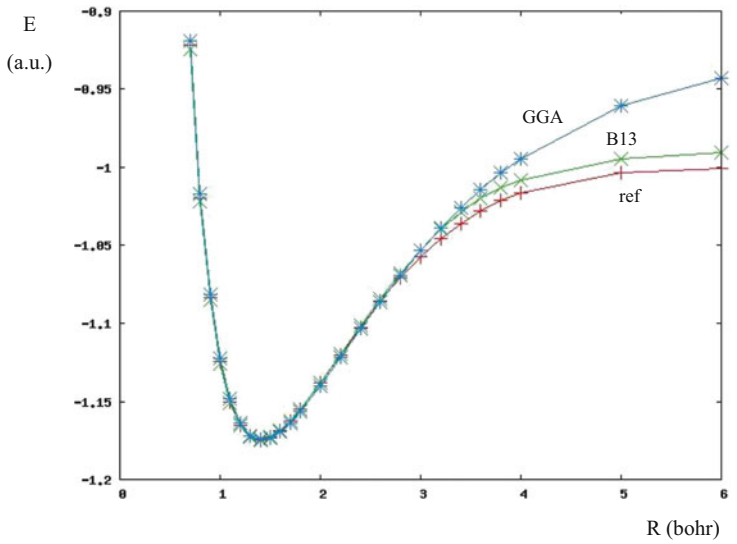


Fig. 6 H_2 dissociation

computations are undesirable, however, as they break spin and space symmetries. To the best of our knowledge, no successful spin-restricted DFT dissociation curve has yet been reported in the literature. We therefore apply B13 to this fundamental problem.

In Fig. 6 the H_2 dissociation curve is plotted for highly accurate Hylleraas variational reference data from Sims and Hagstrom [22] for the B88 + PBE GGA and for B13. Fractional mixing of the σ_g and the σ_u orbitals begins at $R \sim 3.4$ bohr for B13. The GGA does not mix these orbitals at all, and the GGA curve is significantly above the exact asymptote. B13, on the other hand, exhibits a good dissociation limit and, quite interestingly, an asymptotic f of 0.50 (computed at $R = 30.0$ bohr), in accord with the asymptotic CI wavefunction

$$\frac{1}{\sqrt{2}}\left(\left|\sigma_g^2\right\rangle - \left|\sigma_u^2\right\rangle\right). \tag{18}$$

That B13 is sensitive enough to mix these MOs, and with the correct mixing fraction no less, is intriguing.

The heteronuclear dissociation of LiH is even more interesting. At large separation, the HOMO of the system is the 1s orbital of the H atom, and the LUMO is the 2s orbital of the Li atom. Thus the two configurations

$$\left|1s_{Li}^2 \cdot \mathrm{HOMO}^2\right\rangle \text{ and } \left|1s_{Li}^2 \cdot LUMO^2\right\rangle \tag{19}$$

have charges Li^+H^- and Li^-H^+ and dissociation to neutral atoms requires $f = 0.50$. Because LDA calculations on Li^- and H^- are problematic, the $f = 0$ and $f = 1$

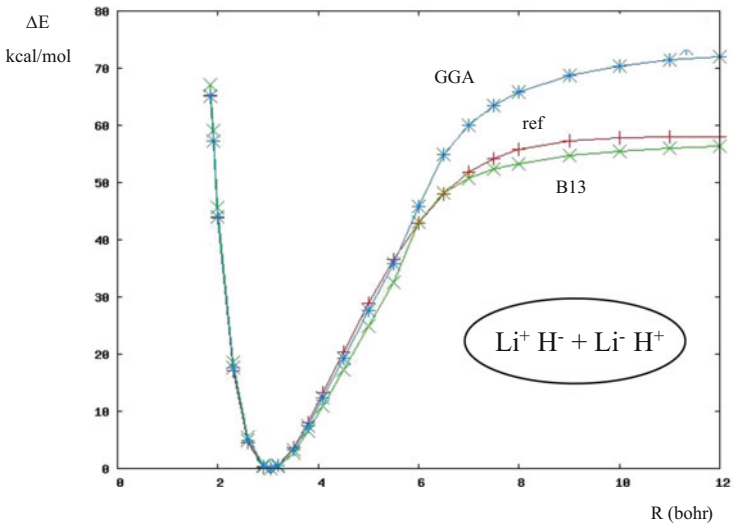


Fig. 7 LiH dissociation

endpoints of the quadratic minimization are unreliable at large separations. Instead, we scan over f values in steps of 0.05 in order to locate the energy minimum at each internuclear separation.

We plot LiH dissociation curves for MR-CISD (Multi-Reference Configuration Interaction with Singles and Doubles) reference data [23], the B88 + PBE GGA, and B13 in Fig. 7. All curves are zeroed at their minimum energy. The limit of the GGA dissociation curve is too high. The B13 curve has an excellent asymptotic energy, and an asymptotic f (computed at $R = 80.0$ bohr) of 0.45, very close to the required 0.50. The dissociation limit is roughly $Li^{+0.1}H^{-0.1}$, very close to neutral atoms. The manner in which B13 captures, to a very good approximation, the correct dissociation limit of this heteronuclear bond is fascinating. At play is a "resonance" of singlet ionic atomic states.

5 Summary and Conclusions

This work marks the first successful application of a spin-restricted Kohn–Sham density functional, "B13", to two-configuration mixing problems with densities as in (7). Moreover, B13 is *exact-exchange*-based. Therefore stretched radical systems such as H_2^+ and He_2^+, the bane of GGA and hybrid functionals [24], are correctly accommodated.

These calculations have been post-LDA and not self-consistent. It is difficult to predict how self-consistency might change their nature, and what significance would be lent to the orbital energies. Is the HOMO–LUMO mixing in H_2 and

LiH an artifact of the post-LDA approach here? Might mixing be obviated in an SCF approach? SCF implementations of the predecessor "B05" functional have been published [12–15]. SCF-B13 should be possible too and could be very interesting.

In future work we will also investigate *multiple* bond dissociations and other strongly-correlated reactions requiring variational optimization of one *or more* mixing fractions. Low-lying *excited* states of strongly-correlated systems, and how they might depend on B13 ground-state determinations, will be explored as well.

The author gratefully acknowledges the Natural Sciences and Engineering Research Council (NSERC) of Canada for research support, the Killam Trust of Dalhousie University for salary support through the Killam Chair program, and the Atlantic Computational Excellence network (ACEnet) for computing support.

References

1. Hohenberg P, Kohn W (1964) Phys Rev 136:B864
2. Kohn W, Sham LJ (1965) Phys Rev 140:A1133
3. Parr RG, Yang W (1989) Density-functional theory of atoms and molecules. Oxford University Press, New York
4. Schipper PRT, Gritsenko OV, Baerends EJ (1999) J Chem Phys 111:4056
5. Schipper PRT, Gritsenko OV, Baerends EJ (1997) Theor Chem Acc 98:16
6. Becke AD (2013) J Chem Phys 138:074109
7. Becke AD (2013) J Chem Phys 138:161101
8. Jiang W, Jeffrey CC, Wilson AK (2012) J Phys Chem A 116:9969
9. Janak JF (1978) Phys Rev B 18:7165
10. Curtiss LA, Raghavachari K, Redfern PC, Pople JA (1997) J Chem Phys 106:1063
11. Becke AD (2013) J Chem Phys 139:021104
12. Proynov E, Shao Y, Kong J (2010) Chem Phys Lett 493:381
13. Proynov E, Liu F, Kong J (2012) Chem Phys Lett 525:150
14. Proynov E, Liu F, Shao Y, Kong J (2012) J Chem Phys 136:034102
15. Arbuznikov AV, Kaupp M (2009) J Chem Phys 131:084103
16. Becke AD (2005) J Chem Phys 122:064101
17. Becke AD (1989) Int J Quantum Chem Quantum Chem Symp 23:599
18. Becke AD, Dickson RM (1990) J Chem Phys 92:3610
19. Becke AD (1988) Phys Rev A 38:3098
20. Perdew JP, Burke K, Ernzerhof M (1996) Phys Rev Lett 77:3865
21. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785
22. Sims JS, Hagstrom SA (2006) J Chem Phys 124:094101
23. Holka F, Szalay PG, Fremont J, Rey M, Peterson KA, Tyuterev VG (2011) J Chem Phys 134:094306
24. Becke AD (2003) J Chem Phys 119:2972

Index