Springer Theses
Recognizing Outstanding Ph.D. Research

Philipp Gubler

A Bayesian Analysis of QCD Sum Rules

Springer

Springer Theses

Recognizing Outstanding Ph.D. Research

For further volumes:
http://www.springer.com/series/8790

Aims and Scope

The series "Springer Theses" brings together a selection of the very best Ph.D. theses from around the world and across the physical sciences. Nominated and endorsed by two recognized specialists, each published volume has been selected for its scientific excellence and the high impact of its contents for the pertinent field of research. For greater accessibility to non-specialists, the published versions include an extended introduction, as well as a foreword by the student's supervisor explaining the special relevance of the work for the field. As a whole, the series will provide a valuable resource both for newcomers to the research fields described, and for other scientists seeking detailed background information on special questions. Finally, it provides an accredited documentation of the valuable contributions made by today's younger generation of scientists.

Theses are accepted into the series by invited nomination only and must fulfill all of the following criteria

- They must be written in good English.
- The topic should fall within the confines of Chemistry, Physics, Earth Sciences, Engineering and related interdisciplinary fields such as Materials, Nanoscience, Chemical Engineering, Complex Systems and Biophysics.
- The work reported in the thesis must represent a significant scientific advance.
- If the thesis includes previously published material, permission to reproduce this must be gained from the respective copyright holder.
- They must have been examined and passed during the 12 months prior to nomination.
- Each thesis should include a foreword by the supervisor outlining the significance of its content.
- The theses should have a clearly defined structure including an introduction accessible to scientists not expert in that particular field.

Philipp Gubler

A Bayesian Analysis of QCD Sum Rules

Doctoral Thesis accepted by
Tokyo Institute of Technology, Tokyo, Japan

Springer

Author (Current Address)
Dr. Philipp Gubler
Strangeness Nuclear Physics Laboratory
RIKEN, Nishina Center
Wako, Saitama
Japan

Supervisor
Prof. Makoto Oka
Department of Physics
Tokyo Institute of Technology
Tokyo
Japan

ISSN 2190-5053 ISSN 2190-5061 (electronic)
ISBN 978-4-431-54317-6 ISBN 978-4-431-54318-3 (eBook)
DOI 10.1007/978-4-431-54318-3
Springer Tokyo Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012956304

Printed on acid-free paper

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

Parts of this thesis have been published in the following articles:

- P. Gubler and M. Oka, "A Bayesian Approach to QCD Sum Rules", Progress of Theoretical Physics **125**, 995–1018 (2010).

- P. Gubler, K. Morita and M. Oka, "Charmonium Spectra at finite Temperature from QCD Sum Rules with the Maximum Entropy Method", Physical Review Letters **107**, 092003 (2011).

- K. Ohtani, P. Gubler and M. Oka, "A Bayesian Analysis of the Nucleon QCD Sum Rules", The European Physical Journal A **47**, 114 (2011).

- K. Suzuki, P. Gubler, K. Morita and M. Oka, "Thermal Modification of Bottomonium Spectra from QCD Sum Rules with the Maximum Entropy Method", Nuclear Physics **A897**, 28–41 (2013).

- K. Ohtani, P. Gubler and M. Oka, "Parity Projection of QCD Sum Rules for the Nucleon", Physical Review D **87**, 034027 (2013).

Denn unser Wissen ist Stückwerk,
und unser Weissagen ist Stückwerk.

(1. Korinther 13, 9)

Supervisor's Foreword

Quantum chromodynamics (QCD) has been established as the fundamental theory of the strong interaction of elementary particles, as a part of the standard model. The elements of QCD are quarks and gluons, which have a special internal degree of freedom, called "color" and interact with one another under the principle of the quantum gauge theory. Although the theory looks very simple as written down in one line (Lagrangian), it generates mysteriously a large variety of phenomena through the interactions of hadrons, such as proton, neutron, pion, and so on, which are made of quarks and gluons. This field of research, called hadron physics, has been developed in the past several decades through the close collaboration of experimentalists and theorists.

This volume of Springer Theses Series is devoted to present a newly developed analysis method of hadron spectra from the first principle of QCD. One of the serious difficulties in describing structures and interactions of hadrons from QCD is that the interactions of quarks and gluons are too strong at low energies to be treated by perturbation theory. As a consequence, quarks and gluons are always "confined" into a colorless (color-neutral) entity at low density and temperature and thus cannot be isolated or directly examined in the laboratory. Not many non-perturbative methods of analyzing QCD are known, and the QCD sum rule is one such semi-analytic method, while lattice QCD is a popular numerical method that requires huge computer power.

Dr. Philipp Gubler, in collaboration with a few members of Tokyo Institute of Technology and Kyoto University, has developed a new method of computing the hadron spectrum using the QCD sum rule approach. The QCD sum rule method, invented by M. A. Shifman, A. I. Vainshtein, V. I. Zakharov in 1979, has been successful in extracting masses of hadrons. The conventional analysis method, however, requires an assumption about the form of the spectral function and thus cannot be applied to cases where the shape of the spectral function is not known. The new method is based on the Bayesian inference theory and called maximum entropy method (MEM), which provides us with the most probable spectral function from given information by QCD. Dr. Gubler has applied the method to analyses of various hadron spectra and confirmed that the method works well and

indeed is superior to the conventional method of extracting the hadron spectrum from the QCD sum rule.

One important application happens to be temperature dependence of the spectra of heavy quarkonia, i.e., bound states of a heavy quark and a heavy antiquark. The subject is related to a phase transition of QCD at high temperature. There, quarks and gluons are supposed to become de-confined and form a plasma-like matter. Such matter may have been created at the beginning of our Universe, just after the Big Bang, while on the Earth, high-energy collisions of heavy ions will produce such matter for a short period. In 1986, T. Matsui and H. Satz proposed that the spectrum of heavy quarkonia is drastically modified in the plasma-like matter so that the formation of such matter can be detected by observing the dilepton spectrum in heavy ion collisions. Thus, theoretical study of heavy quarkonia in QCD at finite temperature is very important. According to the present analysis, QCD sum rules show that the quarkonia peaks in the spectral functions dissolve at finite temperature consistently with the Matsui-Satz prediction, while the dissociation temperatures are found to depend on the individual states. This volume contains all the details of this analysis and also other applications of this new method. We expect to have further applications and developments of this method, some of which have been already published after the thesis is accepted.

Dr. Philipp Gubler completed the doctoral course at the Graduate School of Science and Engineering, Tokyo Institute of Technology, in March 2012. Tokyo Institute of Technology is the leading Japanese National University in the fields of science and engineering and marked its 130th anniversary in 2011. The Department of Physics is one of the largest department with about 70 faculty members, 200 undergraduate students, and 180 graduate students. We joined the Springer Theses project in 2011 and Dr. Gubler is the first winner of the honor of being selected from the 18 successful doctoral theses in the academic year from Department of Physics at Tokyo Institute of Technology. I am very happy to introduce his achievement for the doctoral degree and also feel very honored to have supervised his 5-year study at our graduate school.

Tokyo, October 2012 Makoto Oka

Acknowledgments

First, I would like to express my gratitude to my academic advisor at the Tokyo Institute of Technology, Prof. Makoto Oka for allowing me to study such an interesting subject and for sharing with me his ideas, which were crucial for the development of the method presented in this thesis. I also thank him for his continuous encouragement, especially at times when I doubted whether this whole enterprise was really going in the right direction, for many inspiring discussions and for his endurance of my countless sarcastic comments.

Furthermore, I am grateful to the other senior members of the nuclear theory group Prof. Kazuo Muto, Prof. Tetsuo Hyodo and the postdocs Dr. Kiyoshi Sasaki, Dr. Yoichi Ikeda, and Dr. Takayasu Sekihara for the open-minded atmosphere and interesting discussions on physics and non-physics subjects. I also thank some of them and Dr. Hideaki Iida for studying with me about several subjects such as the basics of the quark-gluon plasma and chiral perturbation theory and Dr. Kenji Morita for collaborating with me and sharing his insights on the behavior of quarkonium at finite temperature. I also thank Prof. Su Houng Lee of Yonsei University for helpful discussions about QCD sum rules in general.

I am indebted to all the younger members of our lab as well, namely Mr. Toshitaka Uchino, Mr. Akira Yokota, Mr. Keisuke Ohtani, Mr. Junya Nasu, Mr. Yu Naganuma, and Mr. Kei Suzuki, for sharing many enjoyable hours. Among them I would like to thank Mr. Toshitaka Uchino for organizing many parties and other gatherings and Mr. Junya Nasu for keeping the computing facilities of the lab intact. Moreover, I especially thank Mr. Keisuke Ohtani and Mr. Kei Suzuki for collaborating with me, carrying out several challenging calculations, and coping with my demanding impatience.

It was possible for me to fully concentrate on my research during the 3 years of the doctor course owing to the financial support of the Japan Society for the Promotion of Science for Young Scientists, which also supported my visits to several international conferences abroad.

Last but not least, I want to thank my family and friends, both back in Switzerland and Japan, for their moral support, love, and friendship.

Contents

Part I
Introduction and Review

Chapter 1
Introduction

1.1 Describing Hadrons from QCD

The theory of Quantum Chromodynamics (QCD) is by now firmly established to be the correct framework to describe the phenomena of the strong interaction, which governs the dynamics of quarks and gluons and, indirectly, of hadrons (Han and Nambu 1965; 't Hooft 1971a, b; Fritzsch et al. 1973; Gross and Wilczek 1973; Politzer 1973). Especially for processes occurring at a high energy scale, perturbative calculations based on QCD provide a very accurate description due to asymptotic freedom (Muta 1998). However, for systems where low energy scales are dominant (such a hadrons), it is much more difficult to make quantitative predictions form QCD, as perturbation theory is of no use here because of the large coupling constant at low energies. Therefore, it is necessary to implement non-perturbative methods to extract the properties of hadrons from QCD, and many such methods have been developed and applied to hadronic systems during the last decades. These methods can be generally divided into two groups. Firstly, there are those which make use of the techniques of effective field theories and employ the lowest lying mesons and baryons as their fundamental degrees of freedom. The second group directly computes the hadronic observables from the Lagrangian of QCD and hence relies on the language of quarks and gluons.

One representative example of the first group is chiral perturbation theory (Weinberg 1979; Gasser and Leutwyler 1984) with its various branches. In the most simple version of this theory, one considers the chiral symmetry and its dynamical breaking to be most fundamental and constructs the most general effective Lagrangian from the Nambu-Goldstone bosons (the pions), which occur due to the breaking of the chiral symmetry. The most general Lagrangian which satisfies chiral symmetry in principle contains an infinite number of terms, but as has been shown by Weinberg (1979), it is possible to introduce a power counting scheme in the energy scale p of the theory which allows to order the various terms of the general Lagrangian and to carry out a systematic expansion in terms of the scale p. If one

P. Gubler, *A Bayesian Analysis of QCD Sum Rules*, Springer Theses,
DOI: 10.1007/978-4-431-54318-3_1, © Springer Japan 2013

however takes the energy to be too large, this expansion breaks down, which is natural from the nature of the theory being an effective theory.

As for the second group, which makes use of quarks and gluons as their degrees of freedom, the most powerful method nowadays is lattice QCD (Wilson 1974; Creutz 1983), which approximates the continuous space-time by a discrete lattice of finite extent and calculates the expectation value of any desired quantity by numerical evaluation of the path integrals of the gluonic fields defined on the lattice. This procedure is completely non-perturbative and does not contain any truncated expansion such as the one mentioned above for chiral perturbation theory. However, numerical lattice simulations have turned out to be numerically very costly and it is only recently that simulations at realistically small quark masses and with dynamical quarks taken into account have become possible (Aoki et al. 2010; Durr et al. 2011). Furthermore, to compare the results obtained on the lattice with experiment, one in principle has to take the continuum and large volume limit of the lattice values which requires calculations at various lattice spacings and volumes, which is a highly time consuming task. These difficulties will however eventually be overcome once more powerful computers and more efficient algorithms are available and one can thus expect substantial and continuous progress in this field in the coming years.

Another method which is directly based on the QCD Lagrangian is the QCD sum rule approach (Shifman et al. 1979a, b), which will be the basic tool employed in this thesis. It can be considered to be of a complementary character compared to lattice QCD because it is analytic and can be used in continuum space-time in an infinite volume with arbitrary quark masses. On the other hand, it uses certain approximations which limit its precision to about $10 \sim 20\,\%$. Therefore, its predictions are only of a half-qualitative—half-quantitative nature, but they still contain many important physical insights. Among them, probably the most interesting one is the feature of providing non-trivial relations between hadronic observables and the properties of the QCD vacuum characterized by various condensates. This aspect will be discussed in detail in the course of this thesis.

Finally let us also briefly mention the approach based on the Schwinger-Dyson equations (Dyson 1949; Schwinger 1951; Roberts and Williams 1994), which also makes direct use of the QCD Lagrangian. In this approach, resummation techniques of Feynman diagrams are employed to incorporate non-perturbative effects into the calculation, but as it is the case for QCD sum rules, one necessarily needs to make some approximations in order to carry out the analysis in a finite amount of time, and thus can only get results with a limited precision.

Furthermore, note that there are also other methods, which do not fall into the two categories discussed above. The most prominent examples are non-relativistic QCD (NRQCD) (Caswell and Lepage 1986) and potential non-relativistic QCD (pNRQCD) (Brambilla et al. 2005), which are effective theories, but still use (heavy) quarks as their fundamental degrees of freedom. Another such case is the Nambu-Jona-Lasinio (NJL) model (Nambu and Jona-Lasinio 1961a, b), which uses an effective theory of quarks to model the breaking of chiral symmetry.

In all, one can understand from the above description that there are nowadays many numerical and analytical approaches available for the description of hadrons,

each with strengths and weaknesses and each providing some specific viewpoint of the numerous phenomena in which hadrons are involved. Hence it at present does not seem to be appropriate to just rely on one method for obtaining a thorough understanding of the hadronic world, but rather to have in mind all the various approaches with their differing degrees of freedom to grasp the whole richness of the processes governed (on the most fundamental level) by QCD. However, from a practical point of view, one can of course not do everything at once and thus has to concentrate on one specific method. For this thesis, this method will be QCD sum rules.

1.2 QCD Sum Rules and Its Ambiguities

The method of QCD sum rules was developed by the members of the ITEP group in Moscow, Russia (Shifman et al. 1979a, b) and was widely used to investigate the properties of many hadronic ground states in the eighties (Reinders et al. 1985). It was especially popular back then because lattice QCD simulations were still in their early stages of development and thus not yet as precise as they are today. Later, the sum rules were generalized to study the behavior of hadrons at finite temperature (Bochkarev and Shaposhnikov 1986; Hatsuda et al. 1993) and density (Drukarev and Levin 1988a, b; Cohen et al. 1991; Hatsuda and Lee 1992). This is still an active field of study because of the relations mentioned in the last section between the properties of hadrons and the condensates, which QCD sum rules provide. Extending these relations to finite temperature or density therefore makes it possible to interpret the changes of, for instance, hadron masses in matter as shifts of a specific combination of condensates and thus to changes of the QCD vacuum in a hot or dense environment.

As will be discussed in detail in Chap. 3, QCD sum rules can essentially be described as a certain class of integrals over some hadronic spectral function $\rho(\omega)$:

$$G(x) = \int_0^\infty d\omega K(x, \omega)\rho(\omega). \tag{1.1}$$

Here, $K(x, \omega)$ is some kernel, whose concrete analytic forms can be found in Chap. 3. In principle, $\rho(\omega)$ contains all information on the physical states, but all we can directly calculate is the left-hand side of Eq. (1.1), $G(x)$, which is obtained by the so-called operator product expansion. However, $G(x)$ can not be evaluated with full accuracy as it contains several condensates at the higher orders of the operator product expansion, for which precise information is usually not available. Hence, the left-hand side of Eq. (1.1) involves considerable uncertainties, which makes the evaluation of the spectral function $\rho(\omega)$ a very difficult task. Precisely speaking, this is in fact an ill-posed problem and cannot be solved analytically. Therefore, is has been common practice in QCD sum rule studies to make some simple assumption about the functional form of $\rho(\omega)$, parametrize it with a small number of parameters and fit it with the help of Eq. (1.1). This approach has been successful, especially for

channels, for which one already has some prior information on the spectrum and can thus make reasonable guesses about the form of $\rho(\omega)$.

However, if one does not know much about the form of the spectral function of interest (for example, about its behavior at finite temperature and density), a fit of $\rho(\omega)$ using some specific ansatz, can lead to ambiguous or even misleading results, and it thus would be helpful if one could analyze Eq. (1.1) in a more general and unbiased manner. We propose in this thesis that such an analysis is possible by the use of the maximum entropy method (MEM), based on Bayesian inference theory. The main advantage of such an approach is the possibility of analyzing the sum rule of Eq. (1.1), without any assumptions about its functional form. This feature is especially useful for the investigations of hadrons at finite temperature or density as in this case one does usually not know much about the general behavior of the spectral function and thus should better be able to analyze it without strong constraints.

1.3 Hadrons in a Hot and/or Dense Environment

In recent years, the probably most actively investigated topic related to QCD has been the problem of understanding and specifying the phase diagram of QCD (Fukushima and Hatsuda 2011). During the last two decades, these studies have been stimulated by the enormous effort that has been invested into the dedicated heavy-ion collision experiments at CERN and at the Brookhaven National Laboratory and this stimulation will continue as the LHC at CERN has just started its operation and first exciting results are becoming available. While these experiments probe matter at high temperature and low density, there are also other facilities which rather aim to create and investigate matter at high density and low temperature. One example is J-PARC in Japan, at which, among other scientific goals, the behavior of various hadrons in a dense environment will be studied. Moreover, the planned FAIR facility at GSI, will hopefully provide valuable information on the behavior of matter at very high density, which has so far not been accessible at past and present experiments.

At the same time, there has been much progress in the theoretical description of the QCD phase diagram, on one hand based on advancements of numerical techniques used for lattice QCD (whose applications are, however, still largely confined to the zero- or low-density region of the phase diagram, due to the "sign problem" (Nakamura 1984)), on the other hand based on the developments in the understanding of effective theories of QCD at finite temperature and density.

There is, however, a considerable gap between these experimental and theoretical efforts, because all one can measure in an experiment are hadronic and other physical observables that appear as a result of the experiment and not the properties of matter itself. Therefore, one somehow has to deduce these properties from, for instance, the abundances and distributions of certain hadrons in a specific experiment. A large amount of work has been and is still being done to fill this gap and reviewing all of it is beyond the scope of this thesis and of my present knowledge. However, one important part of these efforts is to understand how hadrons behave in a hot and dense

environment, and it is my personal hope that the results presented in this thesis will provide a novel and useful tool for investigating this sort of problem.

This then leads me directly to the motivation for my thesis, which is outlined in the next section.

1.4 Motivation and Purpose of this Thesis

The motivation for the studies presented in this thesis is based on the considerations of the preceding two sections. First of all, our aim is to investigate whether the maximum entropy method can be used to analyze QCD sum rules, and if yes, what structures of the spectral function can be extracted. As this kind of approach has never been tried before, this step is crucial for making sure that the method works properly and for correctly interpreting the output of the analyses. Secondly, it is our purpose to illustrate the use of the developed method in a concrete example of physical interest and to show that the MEM analysis of QCD sum rules really is able to provide interesting results on the behavior of the spectral function in hot or dense matter, which could not have been obtained with the conventional analysis.

To test the ability of MEM to extract the spectral function from the sum rules, we first analyze two well known channels in the vacuum: the ρ meson and the nucleon channel. As is shown in detail in Chaps. 5 and 6, we find in these investigations that the MEM procedure can be applied to QCD sum rules and that the positions and residues (value of the integrated spectral function in the peak-region) can be approximately reproduced by our method. We, however, also find that details of the spectral function, such as the width of the peaks, can not be properly reproduced due to the limited resolution of the MEM analysis.

Next, we go on and analyze various quarkonium sum rules at zero and finite temperature (see Chap. 7 for details). Charmonia (in particular J/ψ) are considered to be a suitable probe for the experimental determination of the properties of the quark-gluon plasma (QGP) created in heavy ion collisions (Matsui and Satz 1986), and have thus been intensively studied for many years. However, their behavior at finite temperature is still not completely understood at present, and our goal is to provide new and independent information on this system based on QCD. We show in our analysis that charmonium peaks of the vector, pseudoscalar, scalar and axialvector channel experience a sudden change above the deconfinement temperature T_c and melt quickly into the continuum around $1.0\ T_c \sim 1.2\ T_c$. Furthermore, due to the heavy bottom quark mass, bottomonia are expected to persist as bound states up to higher temperatures, as they are more tightly bound. Thus, they can be useful for the study the QGP at higher temperatures, which has been experimentally produced in the recent heavy ion collisions at the LHC. We therefore have examined several bottomonium channels and have deduced the corresponding melting temperatures which roughly lie in the region of $2.0\ T_c \sim 3.0\ T_c$.

Another goal of the present thesis is of somewhat more technical nature. Namely, we want to investigate which kind of kernel is most suitable for the MEM analysis

of the sum rules of Eq. (1.1). As we will show in the following chapters, in QCD sum rule studies, one is not restricted to one unique form of $K(x, \omega)$, but in principle has an infinite number of choices. Up to now, the so-called Borel kernel has been the one used most frequently, but this is not necessarily the best one for an MEM analysis. Indeed, as we will see in our study of a toy model and the nucleon sum rule, the Gaussian kernel has turned out to be more effective. On the other hand, for our study of quarkonium at finite temperature, the traditional Borel kernel was more suitable, as the Gaussian kernel had problems in reproducing a flat continuum structure. Therefore, our results indicate that there is no single best choice and one thus should choose the kernel depending on what channel one wants to study. This will be our attitude during this thesis and we hence will use both the Borel and Gaussian kernels in the following investigations.

1.5 Outline of the Thesis

The thesis is organized as follows.

Part I deals with general issues of QCD and the methods we use for the analysis. After the introduction of this chapter, the properties of QCD are briefly reviewed in Chap. 2. Here basic concepts such as asymptotic freedom, chiral symmetry and the phases of QCD are explained, as they are fundamental for the understanding of the following chapters. Then, we recapitulate the ideas of QCD sum rules in Chap. 3. Because this method is the main tool used in this thesis, we discuss it in some detail, including several appendices with concrete calculations. In the last section of this chapter, we furthermore discuss the problem of parity projection of baryonic sum rules and clarify some technical problems that have been recently discussed in the literature (Jido et al. 1996; Kondo et al. 2006; Ohtani et al. 2013). Next, we explain the details of the maximum entropy method in Chap. 4, especially focusing on its application to QCD sum rules. This discussion is partly based on Gubler and Oka (2010). Moreover, in the last part of this chapter, the MEM analysis of certain sum rules based on a non-relativistic potential model, which have some resemblance to the real QCD sum rules, is discussed for illustration and as a first check of the validity of the method.

Part II is then devoted to applications of MEM to the actual QCD sum rules of several channels. In Chap. 5, the ρ meson case is discussed on the basis of Gubler and Oka (2010). As a second application, the nucleon channel is investigated in Chap. 6. The calculations of this chapter were done mainly by Mr. Keisuke Ohtani, a younger colleague at our lab at TokyoTech. The related publication is Ohtani et al. (2011). The last application then deals with the behavior of quarkonia at finite temperature and is presented in Chap. 7, which is work done partly in collaboration with Mr. Kei Suzuki of the same TokyoTech lab and Dr. Kenji Morita from the Yukawa Institute for Theoretical Physics, Kyoto University. The content of this chapter is based on Gubler et al. (2011) and Suzuki et al. (2013).

Finally, part III contains the concluding remarks. Here, we will also try to give an outlook for possible future work related to the studies presented in this thesis.

References

Aoki S et al (2010) PACS-CS collaboration. Phys Rev D 81:074503
Bochkarev AI, Shaposhnikov ME (1986) Nucl Phys B268:220
Brambilla N, Pineda A, Soto J, Vairo A (2005) Rev Mod Phys 77:1423
Caswell WE, Lepage GP (1986) Phys Lett B 167:437
Cohen T, Furnstahl RJ, Griegel DK (1991) Phys Rev Lett 67:961
Creutz M (1983) Quarks, gluons and lattices. Cambridge Monographs on Mathematical Physics, Cambridge
Drukarev EG, Levin EM (1988a) Pis'ma. Zh Eksp Teor Fiz 48:307
Drukarev EG, Levin EM (1988b) JETP Lett 48:338
Durr S et al (2011) Phys Lett B 701:265
Dyson FJ (1949) Phys Rev 75:1736
Fritzsch H, Gell-Mann M, Leutwyler H (1973) Phys Lett B 47:365
Fukushima K, Hatsuda T (2011) Rep Prog Phys 74:014001
Gasser J, Leutwyler H (1984) Ann Phys 158:142
Gross DJ, Wilczek F (1973) Phys Rev Lett 30:1343
Gubler P, Oka M (2010) Prog Ther Phys 125:995
Gubler P, Morita K, Oka M (2011) Phys Rev Lett 107:092003
Han MY, Nambu Y (1965) Phys Rev 139:1006
Hatsuda T, Lee SH (1992) Phys Rev C 46:34
Hatsuda T, Koike Y, Lee SH (1993) Nucl Phys B394:221
Jido D, Kodama N, Oka M (1996) Phys Rev D 54:4532
Kondo Y, Morimatsu O, Nishikawa T (2006) Nucl Phys A764:303
Matsui T, Satz H (1986) Phys Lett B 178:416
Muta T (1998) Foundations of quantum chromodynamics, 2nd edn. World Scientific, Singapore
Nakamura A (1984) Phys Lett B 149:391
Nambu Y, Jona-Lasinio G (1961a) Phys Rev 122:345
Nambu Y, Jona-Lasinio G (1961b) Phys Rev 124:246
Ohtani K, Gubler P, Oka M (2011) Eur Phys J A 47:114
Ohtani K, Gubler P, Oka M (2013) Phys Rev D 87:034027
Politzer HD (1973) Phys Rev Lett 30:1346
Reinders LJ, Rubinstein H, Yazaki S (1985) Phys Rep 127:1
Roberts CD, Williams AG (1994) Prog Part Nucl Phys 33:477
Schwinger JS (1951) Proc. Nat. Acad. Sci. 37(452):455
Shifman MA, Vainshtein AI, Zakharov VI (1979a) Nucl Phys B147:385
Shifman MA, Vainshtein AI, Zakharov VI (1979b) Nucl Phys B147:448
Suzuki K, Gubler P, Morita K, Oka M (2013) Nucl Phys A897:28
't Hooft G, (1971a) Nucl Phys B33:173
't Hooft G, (1971b) Nucl Phys B35:167
Weinberg S (1979) Physica A 96:327
Wilson KG (1974) Phys Rev D 10:2445

Chapter 2
Basic Properties of QCD

2.1 The QCD Lagrangian

In this and the following sections, the basic properties of QCD will be briefly discussed. A more detailed exposure of the many facets of QCD can be found in various textbooks, such as Yndurain (1983), Greiner et al. (1994), Ioffe et al. (2010).

QCD is a non-abelian $SU(3)$ gauge theory with color charges as the generators of the gauge group. Its Lagrangian can be given as

$$\mathscr{L}_{QCD} = \sum_f \overline{q}_f (i \not{D} - m_f) q_f - \frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu}, \tag{2.1}$$

where q_f are the quark fields with flavor f, which runs over all presently known six flavors u, d, s, c, b, t with the corresponding masses m_f. The quark fields in fact have two more indices, which are omitted above for simplicity. One is a spinor index running from 1 to 4, showing that the quark (and its anti-particle) are spin-1/2 particles, the other is a color index with values from 1 to 3, meaning that the quarks live in the fundamental representation of the $SU(3)$ gauge group. The covariant derivative $\not{D}$ contains the coupling between the quarks and gauge fields and is defined as

$$\not{D} = \gamma^\mu (\partial_\mu - i g A_\mu). \tag{2.2}$$

Here, the gluon field A_μ is a 3×3 matrix and lives in the adjoint representation of the $SU(3)$ gauge group. Using the Gell-Mann matrices λ^a, it can be expanded as $A_\mu = 1/2 \sum_a \lambda^a A^a_\mu$ $(a = 1 \sim 8)$. Furthermore, g stands for the gauge coupling constant. Finally, the last term of Eq. (2.1) represents the dynamics of the gluonic fields only. It can be expressed in terms of the field strength tensor $G^a_{\mu\nu}$, which is obtained from the gluon fields as

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu, \tag{2.3}$$

where f^{abc} are the structure constants of the $SU(3)$ gauge group.

P. Gubler, *A Bayesian Analysis of QCD Sum Rules*, Springer Theses,
DOI: 10.1007/978-4-431-54318-3_2, © Springer Japan 2013

Note that in the simple description of this section, we have omitted ghost fields and possible gauge fixing terms, which are introduced during the quantization of the theory (Faddeev and Popov 1967).

2.2 Asymptotic Freedom

One important piece of evidence, suggesting that QCD is indeed the true theory to describe the strong interaction, was provided by the discovery of asymptotic freedom (Gross and Wilczek 1973; Politzer 1973). This property, which essentially means that the coupling constant g appearing in Eqs. (2.2) and (2.3) becomes small at large energies, can be derived through the renormalization procedure of QCD. As in any field theory, the perturbative quantum (loop) corrections in QCD contain ultra-violet divergences, which have to be renormalized for the theory to produce meaningful results. If (as it is the case for QCD) the theory is renormalizable, all these divergences can be absorbed into a redefinition of the bare coupling constant g, the bare masses m_f and the fields q and A. However, this redefinition will depend on the energy scale μ, at which the renormalization is carried out, therefore introducing some dependence of the parameters of the theory on μ. As μ is an arbitrary parameter, which has been introduced by hand, the observables calculated from the theory should not depend on it. This requirement leads to several renormalization group equations (Callan 1970; Symanzik 1970), in which the part dealing with the coupling constant g is given as,

$$\mu \frac{\partial g}{\partial \mu} = \beta(g), \tag{2.4}$$

where $\beta(g)$ is the β-function, which can be perturbatively calculated for small g. In QCD, this function has the following form (as an expansion in g):

$$\beta(g) = -\beta_0 g^3 - \beta_1 g^5 + \cdots, \tag{2.5}$$

$$\beta_0 = \frac{1}{(4\pi)^2}\left(11 - \frac{2}{3}N_f\right), \tag{2.6}$$

$$\beta_1 = \frac{1}{(4\pi)^4}\left(102 - \frac{38}{3}N_f\right), \tag{2.7}$$

N_f being the number of flavors. The fact that β_0 is positive and that therefore $\beta(g)$ is negative for sufficiently small values of g has been revealed in the papers of Gross and Wilczek (1973), Politzer (1973) and is equivalent to asymptotic freedom, as will be seen below.

Solving Eq. (2.4) and keeping for simplicity only the leading β_0 term, the following result can be obtained:

$$\alpha_s(\mu) = \frac{1}{4\pi \ln(\mu^2/\Lambda^2_{\mathrm{QCD}})}. \tag{2.8}$$

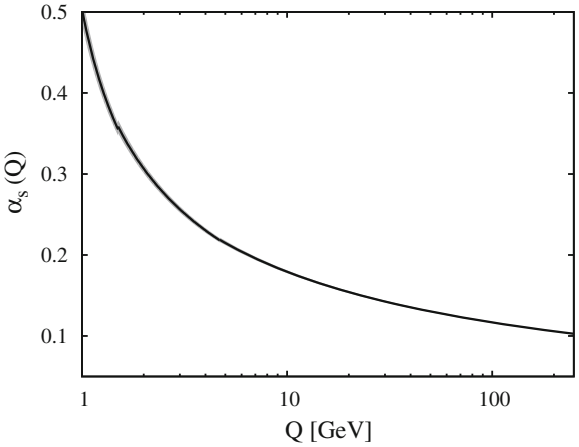

Fig. 2.1 The value of α_s as a function of the energy Q, obtained from the β-function including corrections up to 4-loops (Bethke 2009). The shaded region shows its corresponding numerical uncertainty. The discontinuity seen in the plot at around 1.5 GeV stems from the matching between the 3- and 4-flavor β-function, which has to be implemented in this energy region

Here, α_s stands for $g^2/(4\pi)$ and the integration constant Λ_{QCD} is known as the QCD scale parameter. Its actual value is about 200–300 MeV, depending on how many flavors one considers to be active. From this equation, one sees that the value α_s decreases with larger energy μ. Nowadays, the β-function is known up to 4-loops (g^9) (Bethke 2009), giving values of α_s as shown in Fig. 2.1. As can be observed from this figure, as long as the energy scale Q is much larger than $\sim$1 GeV, α_s is small and a perturbative treatment is meaningful, while for energy scales around or below $\sim$1 GeV, α_s becomes so large that a perturbative expansion will eventually break down. It is therefore clear that non-perturbative methods are necessary for studying low-energy QCD processes.

2.3 Symmetries of QCD

2.3.1 Gauge Symmetry

Gauge symmetry is in a certain sense the most important symmetry of QCD, as the QCD Lagrangian was indeed constructed on the basis of gauge invariance. It demands that the theory should be invariant under the following gauge transformation.The fermionic fields change as

$$q'(x) = U(x)q(x), \tag{2.9}$$

where $U(x)$ is a 3×3 unitary matrix in color space, which generally depends on the space-time point x. Because of this dependence, Eq. (2.9) is a local transformation and the corresponding gauge symmetry a local symmetry. At the same time, the gauge fields $A_\mu(x)$ are transformed as

$$A'_\mu(x) = U(x)A_\mu(x)U^\dagger(x) + \frac{i}{g}U(x)\partial_\mu U^\dagger(x), \tag{2.10}$$

with the same $U(x)$ as in Eq. (2.9). It is not difficult to show that these transformations act on the covariant derivative of Eq. (2.2) in the following way:

$$D'_\mu(x) = U(x)D_\mu(x)U^\dagger(x). \tag{2.11}$$

This then immediately shows that the first term of the Lagrangian in Eq. (2.1), which describes the motion of the quark fields, is gauge invariant. For the second term, involving only gluonic fields and their mutual interactions, it is convenient to note that the field strength tensor of Eq. (2.3) (contracted with $\frac{1}{2}\lambda^a$) can be expressed as

$$G_{\mu\nu}(x) = \frac{i}{g}[D_\mu, D_\nu], \tag{2.12}$$

from which, together with Eq. (2.11), it follows that this object transforms as

$$G'_{\mu\nu}(x) = U(x)G_{\mu\nu}(x)U^\dagger(x). \tag{2.13}$$

As the last term of Eq. (2.1), can be written down as the trace of two $G_{\mu\nu}(x)$'s with contracted Lorentz indices, one can see from Eq. (2.13) that it is also gauge invariant, as it should be.

In actual calculations, one often makes use of the freedom of choosing a gauge to simplify the algebraic manipulations. For calculations of the operator product expansion in QCD sum rules, for instance, the so-called Fock-Schwinger gauge is a convenient choice, as we will discuss in Chapter 3 and Appendix B. On the other hand, any physical result obtained from QCD should not depend on the gauge, in which it was calculated. Therefore, to verify the gauge independence of some result can serve as a useful check of the calculation. Furthermore, for formulating QCD on a space-time lattice in order to carry out Monte-Carlo simulations, gauge invariance also has provided essential guidance (Wilson 1974).

2.3.2 Chiral Symmetry

As will be explained below, chiral symmetry is not an exact symmetry of the QCD Lagrangian of Eq. (2.1), but is only valid in the limit of small quark masses. However, because the masses of the u- and d-quarks are much smaller than $\Lambda_{\rm QCD}$, the typical

scale of QCD, these can be treated as small perturbations. Therefore chiral symmetry becomes a useful concept, at least for the u- and d-quark sector, which plays the most prominent role in almost all low-energy hadronic processes.

To discuss chiral symmetry, one first has to introduce left-handed and right-handed quarks, which are defined as follows

$$q_{\mathrm{L}} = P_{\mathrm{L}} q, \quad q_{\mathrm{R}} = P_{\mathrm{R}} q, \tag{2.14}$$

with

$$P_{\mathrm{L}} = \frac{1}{2}(1 - \gamma_5), \quad P_{\mathrm{R}} = \frac{1}{2}(1 + \gamma_5). \tag{2.15}$$

Here, it is clear that the projection operators $P_{\mathrm{L,R}}$ satisfy the necessary conditions $P_{\mathrm{L,R}}^2 = P_{\mathrm{L,R}}$, $P_{\mathrm{L}} P_{\mathrm{R}} = 0$, $P_{\mathrm{L}} + P_{\mathrm{R}} = 1$. Rewriting the QCD Lagrangian of Eq. (2.1) with the help of the left- and right-handed quark fields of Eq. (2.14), we obtain

$$\mathscr{L}_{QCD} = \overline{q}_{\mathrm{L}} i \not{D} q_{\mathrm{L}} + \overline{q}_{\mathrm{R}} i \not{D} q_{\mathrm{R}} - \overline{q}_{\mathrm{L}} m q_{\mathrm{R}} - \overline{q}_{\mathrm{R}} m q_{\mathrm{L}} - \frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu}, \tag{2.16}$$

where we have omitted the sum over the flavors for simplicity of notation. It is seen in the above equation that, if the quark mass m approaches 0, the left- and right-handed quarks completely decouple and behave as independent degrees of freedom. Ignoring the mass terms for a moment, it is also observed that this Lagrangian has a global symmetry, corresponding to certain unitary transformations of the quark fields:

$$q'_{\mathrm{L}} = U_{\mathrm{L}} q_{\mathrm{L}}, \quad U_{\mathrm{L}} \in U(N_f)_{\mathrm{L}}, \tag{2.17}$$

$$q'_{\mathrm{R}} = U_{\mathrm{R}} q_{\mathrm{R}}, \quad U_{\mathrm{R}} \in U(N_f)_{\mathrm{R}}. \tag{2.18}$$

Here, $U_{\mathrm{L,R}}$ are unitary $N_f \times N_f$ matrices, operating in the flavor space of the quark fields. As only the u- and d-quarks (and, to a lesser degree, the s-quarks) can be considered to be light, N_f here is usually taken to be 2 or 3. Among the symmetries contained in Eqs. (2.17) and (2.18), two have a somewhat special character. Firstly, $U(1)_{\mathrm{V}}$, standing for the case, in which U_{L} and U_{R} are diagonal and equal, represents the quark number conservation in the strong interaction. This symmetry even holds when the finite quark masses are taken into account and is valid at both the classic and quantum level. Secondly, the symmetry of $U(1)_{\mathrm{A}}$, in which U_{L} and U_{R} are diagonal as well, but represent rotations in the opposite direction, is violated by quantum corrections, leading to the axial anomaly (Bell and Jackiw 1969; Adler 1969). Therefore, even if the quark masses are exactly 0, this symmetry is broken.

Removing the two subgroups $U(1)_{\mathrm{V}}$ and $U(1)_{\mathrm{A}}$ discussed above, we are left with the symmetries corresponding to $SU(N_f)_{\mathrm{L}} \times SU(N_f)_{\mathrm{R}}$, which are usually referred to as chiral symmetry. The respective transformations can be parametrized as

$$q'_{\rm L} = e^{i\theta^a_{\rm L} t^a} q_{\rm L}, \quad e^{i\theta^a_{\rm L} t^a} \in SU(N_f)_{\rm L}, \tag{2.19}$$

$$q'_{\rm R} = e^{i\theta^a_{\rm R} t^a} q_{\rm R}, \quad e^{i\theta^a_{\rm R} t^a} \in SU(N_f)_{\rm R}, \tag{2.20}$$

with $a = 1 \sim N_f^2 - 1$ and t^a being the generators of $SU(N_f)$. $\theta^a_{\rm L,R}$ are arbitrary real parameters.

Even though the QCD Lagrangian possesses chiral symmetry as described in Eqs. (2.19) and (2.20), this symmetry is not fully realized in the QCD vacuum. Specifically, it is instead believed to be dynamically broken (Nambu and Jona-Lasinio 1961a, b), the symmetry breaking pattern being

$$SU(N_f)_{\rm L} \times SU(N_f)_{\rm R} \to SU(N_f)_{\rm V}, \tag{2.21}$$

where the transformation corresponding to $SU(N_f)_{\rm V}$ is such that both rotations of Eqs. (2.19) and (2.20) are the same ($\theta^a_{\rm L} = \theta^a_{\rm R}$). The simplest order parameter for such a partial breaking of chiral symmetry is the quark condensate, expressed as

$$\langle \overline{q} q \rangle = \langle \overline{q}_{\rm L} q_{\rm R} \rangle + \langle \overline{q}_{\rm R} q_{\rm L} \rangle. \tag{2.22}$$

Here, contraction and summation over Dirac-, color- and flavor-indices are implicitly assumed. It is clear that the quark condensate (if it has a finite value) generally changes its form under arbitrary transformations of Eqs. (2.19) and (2.20), but is invariant under $SU(N_f)_{\rm V}$, making it an appropriate order parameter for the symmetry breaking pattern of Eq. (2.21).

The value of the quark condensate nowadays can be obtained from first principle lattice QCD calculations (Fukaya et al. 2010). The earliest estimation, however, relied on the Gell-Mann-Oakes-Renner relation (Gell-Mann et al. 1968), based on general considerations of chiral symmetry. It gives (with $N_f = 2$),

$$f_\pi^2 m_\pi^2 = -m_q \langle \overline{u} u + \overline{d} d \rangle, \tag{2.23}$$

where f_π and m_π are the pion decay constant and pion mass, respectively. The values of these parameters can be obtained from experiment. m_q stands for the averaged quark mass of u- and d-quarks, which can not be directly extracted from experimental studies, but has to be estimated by other methods (Gasser and Leutwyler 1982). Finally, assuming that the value for the condensate of the u- and d-quarks is the same, one arrives at the following number:

$$\langle \overline{u} u \rangle = \langle \overline{d} d \rangle \simeq -(240\ {\rm MeV})^3. \tag{2.24}$$

In addition, QCD sum rule studies provide estimates for the s-quark condensate $\langle \overline{s} s \rangle$, which gives an about 20 % reduced value, compared to Eq. (2.24) (Reinders et al. 1985).

2.3.3 Dilatational Symmetry

The dilatational symmetry, similarly to chiral symmetry, only holds in the limit of vanishing quark masses. In this limit, the QCD Lagrangian involves no explicit energy scale and the theory is therefore scale invariant. This means that the energy dependence of any physical quantity is fixed by its dimension,

$$F(E, p_1, p_2, \dots) = E^{d_F} f\Big(\frac{p_1}{E}, \frac{p_2}{E}, \dots\Big), \tag{2.25}$$

where d_F stands for the dimension of the quantity F and f is a dimensionless function.

Infinitesimal scale transformations can be parametrized by the following coordinate redefinition

$$x'_\mu = x_\mu + \varepsilon x_\mu, \tag{2.26}$$

with the infinitesimal parameter ε. The N other current derived from this transformation reads as

$$j_\mu = x_\nu T^\nu_\mu, \tag{2.27}$$

T^ν_μ being the energy momentum tensor. This gives

$$\partial^\mu j_\mu = T^\mu_\mu. \tag{2.28}$$

Therefore, the dilatational symmetry is valid if the right hand side of the above equation vanishes. Classically, the trace of the energy momentum tensor T^μ_μ only receives non-zero contributions from terms involving finite quark masses. However, quantum fluctuations lead to additional effects due to the so-called trace anomaly (Crewther 1972; Chanowitz and Ellis 1972; Collins et al. 1977). Taking this contribution into account, one obtains

$$T^\mu_\mu = \frac{\beta}{2g} G^a_{\mu\nu} G^{a\mu\nu} + \sum_f m_f \overline{q}_f q_f, \tag{2.29}$$

where the first term originates from the trace anomaly. β is the β-function of QCD, which has already appeared in Eq. (2.4). It therefore follows from Eqs. (2.28) and (2.29), that the dilatational symmetry is not only broken by the finite quark masses, but also by quantum effects. This can be understood from the fact that in quantum field theory, a renormalization point μ has to be introduced, thus leading to a new scale that violates the symmetry of Eq. (2.25).

2.3.4 Center Symmetry

The center (or $Z(N_c)$) symmetry of QCD has a somewhat different character from the ones discussed so far, as it is a symmetry of QCD at finite temperature, and is only exactly valid when all quarks are infinitely heavy. As will be discussed below, this symmetry is related to the confinement-deconfinement transition of QCD at finite temperature (McLerran and Svetitsky 1981; Svetitsky and Yaffe 1982).

For discussing the center symmetry, it first has to be remembered that in quantum field theory at finite temperature one takes the time axis to be imaginary ($t \to -i\tau$, τ being a real parameter) and all bosonic fields have to satisfy periodic boundary conditions with respect to this axis, the period being $1/T$ (Le Bellac 1996). Now, it is noticed that one can gauge transform the periodic gluon field according to Eq. (2.10) with a transformation matrix $U(x)$, which does not necessarily need to be periodic:

$$U(\tau + 1/T, \mathbf{x}) = \mathbf{z}U(\tau, \mathbf{x}). \tag{2.30}$$

Here, $\mathbf{z}$ must be an element of $SU(N_c)$. Substituting this gauge transformation matrix into Eq. (2.10), one obtains

$$A'_\mu(\tau + 1/T, \mathbf{x}) = \mathbf{z}A'_\mu(\tau, \mathbf{x})\mathbf{z}^\dagger + \frac{i}{g}\mathbf{z}\partial_\mu\mathbf{z}^\dagger. \tag{2.31}$$

In order for this transformed gauge field to be periodic, the right-hand side of the above equation should be equal to $A'_\mu(\tau, \mathbf{x})$, and this can only happen if $\mathbf{z}$ can be interchanged with any other $SU(N_c)$ matrix and does not depend on the space-time coordinates at all. Therefore, $\mathbf{z}$ has to be proportional to the identity matrix $\mathbf{1}$ with a constant coefficient. As $\mathbf{z}$ is an element of $SU(N_c)$, its possible realizations turn out to be

$$\mathbf{z} = e^{2\pi i n/N_c}\mathbf{1}, \quad (n = 0, 1, \ldots, N_c - 1). \tag{2.32}$$

These matrices commute with any member of the $SU(N_c)$ group, are called the center of $SU(N_c)$ and are denoted as $Z(N_c)$.

While the action for the pure $SU(N_c)$ theory is invariant under the center symmetry, one can consider other gauge invariant operators constructed from gluonic fields, for which this is not the case. Among them, the Polyakov loop (Polyakov 1978), defined as the path-ordered product the gauge field, directed in the imaginary time direction from 0 to $1/T$, is most simple:

$$L(\mathbf{x}) = \frac{1}{N_c}\mathrm{Tr}\left\{\mathrm{P}\exp\left[ig\int_0^{1/T} d\tau A_4(\tau, \mathbf{x})\right]\right\}. \tag{2.33}$$

It can be shown that the Polyakov loop transforms as

$$L'(\mathbf{x}) = zL(\mathbf{x}), \tag{2.34}$$

under the center symmetry transformation, where z stands for the constant factor in front of the identity matrix in Eq. (2.32). Hence, as long as it does not vanish, the Polyakov loop is not invariant under the $Z(N_c)$ transformation and therefore serves as an order parameter of the corresponding center symmetry. In other words, if $L(\mathbf{x})$ takes a finite value, the center symmetry is spontaneously broken.

Additionally, it should be mentioned here that the Polyakov loop of Eq. (2.33) also has significant implications related to deconfinement of quarks. Namely, the expectation value of $L(\mathbf{x})$ can be related to the free energy of a single quark (Polyakov 1978; Susskind 1979),

$$\langle L(\mathbf{x}) \rangle = e^{-\Delta F_q / T}, \tag{2.35}$$

where ΔF_q is the difference between the free energy of a system with and without a single deconfined quark. As long as the quarks are confined, such a free energy of a quark should be infinitely large, and the Polyakov loop should thus vanish. On the other hand, for a system in the deconfined phase, the respective quark free energy should have a finite value, meaning that the Polyakov loop will have a value larger than 0.

2.4 Phases of QCD

The phases of QCD at various values of temperature and density continue to be intensively studied both theoretically and experimentally. For a recent review of the current statues in theory, see Fukushima and Hatsuda (2011). However, despite of these efforts, there are still many open questions and fully established facts are rather rare. In this short introduction, we will not discuss all open issues in detail, but can only give a broad overview about what is known about the properties of QCD in a hot or dense medium.

In Fig. 2.2 a sketch of the QCD phase diagram is given. One can see in this figure that there are essentially three phases. Firstly, there is the hadron gas phase at low temperature and density, where both the vacuum in which we live in and nuclear matter are located. Secondly, the quark-gluon plasma phase is realized at high temperature (Cabibbo and Parisi 1975), in which quarks and gluons are deconfined and behave as weakly interacting particles. Thirdly, the color superconductor phase is expected to appear at high density and low temperature, where quarks are believed to form Cooper pairs, leading to color superconductivity (Barrois 1977; Bailin and Love 1984). It however has to be noted here, that many features of this phase diagram are not well understood. Especially for the region of moderately high chemical potential and low temperature, where the three phases meet, there is no conclusive picture available, yet. This is so because in this domain, neither perturbative methods nor lattice QCD calculations (Wilson 1974) can be reliably applied, and one therefore has to resort to model calculations.

The region of the QCD phase diagram that is perhaps best known is located around zero chemical potential, as here lattice QCD calculations are available. Particularly

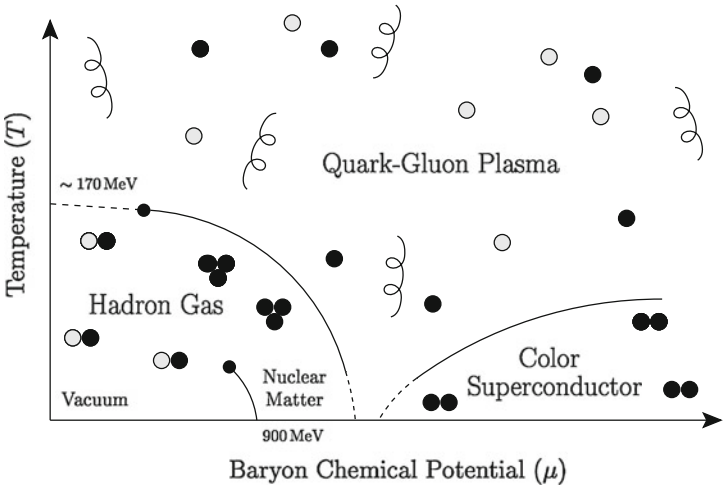

Fig. 2.2 Sketch of the various phases of QCD, as a function of the temperature and the baryon chemical potential, which is related to the baryon density

interesting is the transition region between the hadron and quark-gluon plasma phase, which has been investigated in detail in past studies. The behavior of this transition depends on the flavor content of the theory, as illustrated in the Columbia plot, shown in Fig. 2.3. For the purely gluonic case with no active flavors, which is most easily studied on the lattice because the quenched approximation can be used here, it is found that the transition is of first-order (Fukugita et al. 1990), with a critical temperature of about $T_c \simeq 260$–270 MeV. This situation corresponds to the top-right corner of Fig. 2.3.

On the other hand, due to technical difficulties related to the description of quarks on the lattice, massless quarks can at present not be reliably treated in lattice QCD and one has to consider other methods in this case. A quite general method for handling this problem is the Ginzburg-Landau approach, in which one writes down a general effective Lagrangian in terms of an appropriate order parameter of chiral symmetry (Pisarski and Wilczek 1984). Furthermore, taking into account the effect of the $U_A(1)$ or axial anomaly (Kobayashi and Maskawa 1970; 't Hooft 1976), one obtains

$$\begin{aligned}\mathscr{L}_{\text{eff}} = &\frac{1}{2}\text{Tr}\partial\Phi^\dagger\partial\Phi + \frac{a}{2}\text{Tr}\Phi^\dagger\Phi \\ &+ \frac{b_1}{4!}\left(\text{Tr}\Phi^\dagger\Phi\right)^2 + \frac{b_2}{4!}\text{Tr}\left(\Phi^\dagger\Phi\right)^2 \\ &- \frac{c}{2}\left(\det\Phi + \det\Phi^\dagger\right).\end{aligned} \tag{2.36}$$

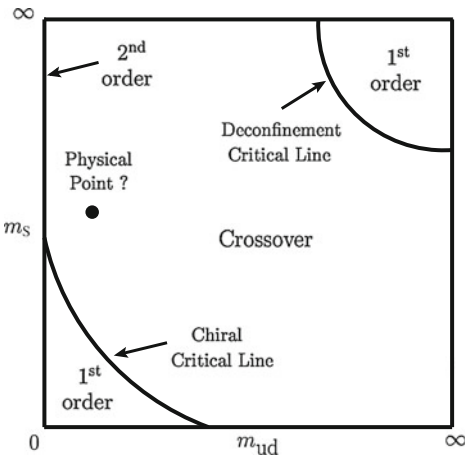

Fig. 2.3 The Columbia plot for QCD at zero density, depicting the nature of the confinement-deconfinement or chiral transition at finite temperature for various flavors. The *top-right* corner describes the purely gluonic transitions with infinitely heavy quarks, while the *bottom-left* corner stands for the chiral transitions with three massless quarks

Here, Φ is a $N_f \times N_f$ matrix and stands for the order parameter of chiral symmetry. Under a chiral $SU(N_f)_{\rm L} \times SU(N_f)_{\rm R}$ transformation, it changes as

$$\Phi' = V_{\rm L} \Phi V_{\rm R}, \tag{2.37}$$

while under the axial $U_{\rm A}(1)$ transformation, it changes as

$$\Phi' = e^{i\alpha} \Phi. \tag{2.38}$$

It can be seen from the above equations that the first four terms of the Lagrangian of Eq. (2.36) are invariant under both $SU(N_f)_{\rm L} \times SU(N_f)_{\rm R}$ and $U_{\rm A}(1)$ transformations, while the last terms breaks the $U_{\rm A}(1)$ symmetry. Therefore, this last term explicitly incorporates the effect of the chiral anomaly.

Analyzing now the thermal properties of Eq. (2.36), one finds that for $N_f = 2$ this Lagrangian is equivalent to the ϕ^4 model, which possesses the $O(4)$ symmetry. This model is known to have a second order phase transition. On the other hand, for $N_f = 3$, due to the cubic interaction introduced by the axial anomaly term, the model exhibits a first order transition. These considerations lead to the picture shown in Fig. 2.3, where on the right side the transition changes from first order at the bottom ($N_f = 3$) to second order on the top ($N_f = 2$). For a more detailed discussion of this issue, see Chap. 6 of Yagi et al. (2005).

Finally, let us consider more realistic cases, which lie close to the physical point, indicated by the black dot in Fig. 2.3. Even though still challenging due to the light u and d quark masses, lattice simulations are now at the stage of becoming possible in

such a regime. Most of these simulations employ the staggered fermions (Susskind 1977; Sharatchandra et al. 1981), while some also use the Wilson fermion formalism (Wilson 1975). Recent results of such studies suggest that the transition at the physical point is smooth crossover (Aoki et al. 2006). Furthermore, the value of the critical temperature has been evaluated by various groups, the latest results giving an averaged value of roughly 170 MeV, with a scatter of about 20 MeV (Aoki et al. 2009; Bazavov et al. 2009; Bornyakov et al. 2010; Cheng et al. 2009; Borsanyi et al. 2010).

References

Adler SL (1969) Phys Rev 177:2426
Aoki Y, Endrodi G, Fodor Z, Katz SD, Szabo KK (2006) Nature 443:675
Aoki Y, Borsanyi S, Durr S, Fodor Z, Katz SD, Krieg S, Szabo KK (2009) JHEP 06:088
Bailin D, Love A (1984) Phys Rep 107:325
Barrois BC (1977) Nucl Phys B 129:390
Bazavov A et al (2009) Phys Rev D 80:014504
Bell JS, Jackiw R (1969) Nuovo Cimento 60A:47
Le Bellac M (1996) Thermal field theory. Cambridge Monographs on Mathematical Physics, Cambridge
Bethke S (2009) Eur Phys J C 64:689
Bornyakov VG et al (2010) Phys Rev D 82:014504
Borsanyi S, Fodor Z, Hoelbling C, Katz SD, Krieg S, Ratti C, Szabo K (2010) JHEP 09:073
Cabibbo N, Parisi G (1975) Phys Lett B 59:67
Callan CG (1970) Phys Rev D 2:1541
Chanowitz M, Ellis J (1972) Phys Lett B 40:397
Cheng M et al (2009) Phys Rev D 81:054510
Collins J, Dunkan L, Joglekar S (1977) Phys Rev D 16:438
Crewther R (1972) Phys Rev Lett 28:1421
Faddeev LD, Popov VN (1967) Phys Lett B 25:25
Fukaya H et al (2010) JLQCD collaboration. Phys Rev Lett 104:122002
Fukugita M, Okawa M, Ukawa A (1990) Nucl Phys B337:181
Fukushima K, Hatsuda T (2011) Rep Prog Phys 74:014001
Gasser J, Leutwyler H (1982) Phys Rep 87:77
Gell-Mann M, Oakes RJ, Renner B (1968) Phys Rev 175:2195
Greiner W, Schramm S, Stein E (1994) Quntum chromodynamics. Springer-Verlag, Berlin
Gross DJ, Wilczek F (1973) Phys Rev Lett 30:1343
Ioffe BL, Fadin VS, Lipatov LN (2010) Quantum chromodynamics, perturbative and nonperturbative aspects. Cambridge Monographs on Particle Physics Nuclear Physics and Cosmology, Cambridge
Kobayashi M, Maskawa T (1970) Prog Theor Phys 44:1422
McLerran LD, Svetitsky B (1981) Phys Rev D 24:450
Nambu Y, Jona-Lasinio G (1961a) Phys Rev 122:345
Nambu Y, Jona-Lasinio G (1961b) Phys Rev 124:246
Pisarski RD, Wilczek F (1984) Phys Rev D 29:338
Politzer HD (1973) Phys Rev Lett 30:1346
Polyakov AM (1978) Phys Lett B 72:477
Reinders LJ, Rubinstein H, Yazaki S (1985) Phys Rep 127:1
Sharatchandra HS, Thun HJ, Weisz P (1981) Nucl Phys B192:205
Susskind L (1977) Phys Rev D 16:3031

Susskind L (1979) Phys Rev D 20:2610
Svetitsky B, Yaffe LG (1982) Nucl Phys B210:423
Symanzik K (1970) Comm Math Phys 18:227
't Hooft G (1976) Phys Rev Lett 37:8
Wilson KG (1975) Quarks and strings on a lattice. In: Arnowitt R, Nath P (ed) Gauge theories and modern field theory. MIT Press, Cambridge
Wilson KG (1974) Phys Rev D 10:2445
Yagi K, Hatsuda T, Miyake Y (2005) Quark-Gluon plasma. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Cambridge
Yndurain FJ (1983) The theory of quark and gluon interactions. Springer-Verlag, Berlin

Chapter 3
Basics of QCD Sum Rules

3.1 Introduction

The method of QCD sum rules was invented in the late seventies (Shifman et al. 1979a, b) and then further developed in the subsequent years (Ioffe 1981a, b; Reinders et al. 1985). It is a useful tool in hadron phenomenology, to obtain a qualitative (in some cases even quantitative) understanding of the properties of hadrons. To do this, one exploits the analytic properties of various correlation functions of certain interpolating fields to connect two separate limits of QCD. On one hand, we have the high energy (or short distance) limit, where asymptotic freedom makes it possible to use perturbative methods to describe the behavior of quarks and gluons, which are the active degrees of freedom in this regime. Concretely, the operator product expansion (OPE) can be employed to evaluate the correlators, explicitly calculating its perturbative parts, while the non-perturbative contributions are stored in the vacuum expectation values of gauge invariant operators. On the other hand, in the low energy (or large distance) limit, the relevant degrees of freedom are the hadrons, whose properties we aim to extract by the sum rules. By this procedure, we can then connect non-perturbative information of the QCD vacuum, parametrized as various vacuum condensates, with the physical quantities of actual hadrons. Discussions on early applications of QCD sum rules, their extensions to a broader range of systems as well as technical details of the actual calculations can be found in various reviews and books (Shifman 1998; Colangelo and Khodjamirian 2001; Leinweber 1997; Ioffe et al. 2010; Pascual and Tarrach 1984; Kojo 2008), on which I partly relied when writing this review.

As the most simple case, let us consider the two-point function (in the following referred to as "correlator") of an interpolating field $J(x)$, which is composed of quark- and/or gluon-fields and carries the quantum number of the hadronic state that we aim to investigate:

$$\Pi(q^2) = i \int d^4x e^{iqx} \langle 0|T[J(x)\bar{J}(0)]|0\rangle. \tag{3.1}$$

P. Gubler, *A Bayesian Analysis of QCD Sum Rules*, Springer Theses,
DOI: 10.1007/978-4-431-54318-3_3, © Springer Japan 2013

Starting from this expression, the analyticity of the function $\Pi(q^2)$ in the whole imaginary plane of the variable q^2 except the positive part of the real axis allows one to write down the dispersion relation

$$\begin{aligned}\Pi(q^2) &= \frac{1}{\pi}\int_0^\infty ds \frac{\mathrm{Im}\Pi(s+i\varepsilon)}{s-q^2} \\ &= \int_0^\infty ds \frac{\rho^J(s)}{s-q^2},\end{aligned} \tag{3.2}$$

where the spectral function related to the interpolating field $J(x)$ has been defined as $\frac{1}{\pi}\Pi(s+i\varepsilon) = \rho^J(s)$. The above equation connects the region of the correlator where calculations using perturbation theory are, in principle, possible ($|q^2| \to \infty$) with the region which is of actual physical interest ($q^2 \sim E^2_{\text{ground state}}$). A detailed derivation of Eq. (3.2) can be found in Appendix A.

3.1.1 The Theoretical Side

For extracting information on the spectral function $\rho^J(s)$ from the sum rule of Eq. (3.2), one first needs to calculate its left hand side, $\Pi(q^2)$ in the deep Euclidean region, where $-q^2$ goes to ∞. This is done by perturbative methods, but simple minded perturbation theory is not enough. Even though, due to asymptotic freedom, the coupling constant approaches 0 at $|q^2| \to \infty$, there are sizable non-perturbative contributions, which have to be taken into account. For this purpose, the authors of Shifman et al. (1979a, b) proposed to use the OPE, originally developed by Wilson Wilson (1969), to incorporate the non-perturbative effects of the low-energy vacuum fields interacting with the high-energy quarks and gluons induced by the external current $J(x)$.

This treatment leads to an expansion of the correlator into a various local operators O_d, ordered by their mass dimension d and the corresponding Wilson coefficients:

$$\Pi(q^2) = \sum_d C_d(q^2)\langle 0|O_d|0\rangle. \tag{3.3}$$

As the local operators are sandwiched between the vacuum state, only Lorentz- and Gauge-invariant operators with positive parity are allowed in the above expansion. The leading operator is the unit operator ($O_0 = 1$) and its Wilson coefficient stands for the correlator of Eq. (3.1), evaluated by standard perturbation theory. The next operators in the expansion can be constructed from all possible Lorentz- and Gauge-invariant local combinations of quark and gluon fields:

$$
\begin{aligned}
&\text{dim. 3}: \quad \langle 0|\overline{q}q|0\rangle, \\
&\text{dim. 4}: \quad \langle 0|G^a_{\mu\nu}G^{a\mu\nu}|0\rangle, \\
&\text{dim. 5}: \quad \langle 0|\overline{q}\sigma_{\mu\nu}\frac{\lambda^a}{2}G^{a\mu\nu}q|0\rangle, \\
&\text{dim. 6}: \quad \langle 0|\overline{q}q\overline{q}q|0\rangle, \langle 0|\overline{q}\gamma_5 q\overline{q}\gamma_5 q|0\rangle, \\
&\qquad\qquad \langle 0|\overline{q}\frac{\lambda^a}{2}q\overline{q}\frac{\lambda^a}{2}q|0\rangle, \langle 0|\overline{q}\gamma^\mu\frac{\lambda^a}{2}q\overline{q}\gamma_\mu\frac{\lambda^a}{2}q|0\rangle, \ldots \\
&\qquad\qquad \langle 0|g^3 f^{abc}G^{a\nu}_{\mu}G^{b\lambda}_{\nu}G^{c\mu}_{\lambda}|0\rangle, \\
&\qquad\qquad \ldots
\end{aligned}
\tag{3.4}
$$

The vacuum expectation values of these operators are usually referred to as "condensates". Among these shown above, only the ones up to dimension 5 have been thoroughly studied and more or less reliable estimates on their actual values exist. As for the many possible versions of the four-quark condensates, they can all be related to the square of the chiral condensate of dimension 3, if one uses Fierz transformation in combination with the vacuum saturation approximation (which will be discussed later). The dimension 6 gluon condensate can also be related to the dimension 4 gluon condensate, with the help of the dilute instanton gas model (Novikov et al. 1979). For condensates with even higher dimension, the uncertainty of their values is much larger, as the validity of the above mentioned approximations used for the dimension 6 condensates is not at all established. Therefore, to obtain reliable results it is important that the OPE is dominated by the first few dimensional terms, and that the contributions from condensates larger that 6 only give small corrections.

3.1.2 The Phenomenological Side

Coming back to the dispersion relation in Eq. (3.2), we now consider the right hand side. Using the optical theorem and inserting a complete set of intermediate hadronic states, we can write

$$
\mathrm{Im}\Pi(q^2) = \frac{1}{2}\sum_n \langle 0|J|n(p_n)\rangle\langle n(p_n)|\overline{J}|0\rangle d\tau_n (2\pi)^4\delta^{(4)}(q-p_n), \tag{3.5}
$$

where n is summed over all hadronic states coupling to J, including sums over polarizations. $d\tau_n$ denotes the integration over the phase space of the states $|n\rangle$. To calculate the spectral function $\rho(s) = 1/\pi\,\mathrm{Im}\Pi(s+\varepsilon)$ and to completely understand its behavior is one of the major goals of hadron physics and can not be done easily.

As the sum rule of Eq. (3.2) only provides information on an integral over the spectral function, one can only hope to extract from it some bulk properties of the spectrum, but not all its detailed features. Therefore, traditionally it has been the custom in practical sum rule analyses to make a deliberated guess on the concrete structure

of the spectral function, parametrize this structure with a small number of parameters and finally fit these parameters with the help of the sum rule. In this thesis, we will adopt a different approach, trying to obtain the spectral function directly from the sum rule with the help of the maximum entropy method. The details of this procedure are explained in Chapter 4 of this thesis.

Here, the conventional approach to the sum rules will be explained, which is important for understanding the motivation for the Bayesian approach, proposed in the present work. As a first step, Eq. (3.5) is divided into two energy regions:

$$\mathrm{Im}\Pi(s) \equiv \theta(s_{th} - s)\mathrm{Im}\Pi^{<}(s) + \theta(s - s_{th})\mathrm{Im}\Pi^{>}(s). \tag{3.6}$$

s_{th} is the so-called threshold parameter and marks the boundary of the two domains. The high energy part of the correlator is then approximated using the quark-hadron duality (Poggio et al. 1976), which connects the hadronic spectral function with the perturbative calculations using the OPE technique, for which the degrees of freedom are quarks and gluons:

$$\mathrm{Im}\Pi^{>}(s) \approx \Pi^{OPE}(s). \tag{3.7}$$

This approximation is only valid at sufficiently high energies, where the perturbative terms dominate over the low-energy correlations.

Usually, one is interested in the ground state or lowest resonance coupling to $J(x)$, which lies below the threshold parameter s_{th}. Here, QCD sum rules often use the one pole approximation, completely ignoring the possible width of the state and potentially occurring scattering states. Such scattering states may be present even below the lowest peak, if it is a resonance. The one pole approximation is given as

$$\mathrm{Im}\Pi^{<}(s) \approx \pi|\lambda|^2\delta(s - m^2), \tag{3.8}$$

where m is the mass of the perceived ground state and $|\lambda|^2$ is the coupling strength of the pole to the interpolating field $J(x)$. Equations (3.7) and (3.8) are the basic assumptions of the usual QCD sum rule analysis and especially in the case of Eq. (3.8), one in principle does not know if it is accurate without any additional information. Thus, the validity of these assumptions has to be carefully checked in the course of the calculation.

3.1.3 Practical Versions of the Sum Rules

Even though we have already obtained a sum rule in Eq. (3.2), it is for several reasons not very useful in this form. First of all, the integral over the spectral function most likely does not converge to a finite value. As explained in Appendix A, this problem can be cured by introducing subtraction terms, which are polynomials in q^2 with infinite coefficients. However, at this point, both sides of Eq. (3.2) still diverge and the sum rule is therefore quite useless in practice. Furthermore, experience of actual

OPE calculations show that the convergence of the expansion in the present form is slow and high-order terms can be quite important.

To improve the situation, it was proposed in Shifman et al. (1979a, b) to use the Borel transformation $\widehat{L}_M$, which is defined as

$$\Pi(M^2) \equiv \widehat{L}_M \Pi(q^2) \equiv \lim_{\substack{-q^2,n\to\infty,\\ -q^2/n=M^2}} \frac{(-q^2)^n}{(n-1)!}\left(\frac{d}{dq^2}\right)^n \Pi(q^2). \tag{3.9}$$

After applying this operator to the dispersion relation of Eq. (3.2), one obtains

$$\Pi(M^2) = \frac{1}{M^2}\int_0^\infty ds e^{-s/M^2}\rho^J(s). \tag{3.10}$$

As one can see in Eq. (3.9), the Borel transformation includes an infinite number of derivatives, therefore all subtraction terms automatically vanish. At the same time, the high energy part of the dispersion integral is exponentially suppressed, meaning that the integral now converges to a finite value (as long as the spectral function itself does not grow exponentially, which is not the case in practice).

Some typical and useful examples of the Borel transformation are shown below,

$$\widehat{L}_M (q^2)^k = 0, \tag{3.11}$$

$$\widehat{L}_M (q^2)^k \ln(-q^2) = -k!(M^2)^k, \tag{3.12}$$

$$\widehat{L}_M \left(\frac{1}{q^2}\right)^k = \frac{(-1)^k}{(k-1)!}\left(\frac{1}{M^2}\right)^k, \tag{3.13}$$

$$\widehat{L}_M \left(\frac{1}{s-q^2}\right)^k = \frac{1}{(k-1)!}\left(\frac{1}{M^2}\right)^k e^{-s/M^2}. \tag{3.14}$$

Here, k is a positive integer, and M the so-called Borel mass. As will be discussed later, higher dimensional terms are usually proportional to $(1/q^2)^k$, the power growing with dimension. Hence, it is observed in Eq. (3.13) that higher dimensional terms of the OPE are suppressed by an additional factor of $\frac{1}{(k-1)!}$, which considerably improves the convergence of the OPE.

The Borel transformed sum rule is, however, not the only way to improve the behavior of Eq. (3.2). As we will see in Chap. 6 of this thesis, there are cases in which one can do better. For instance, in Bertlmann et al. (1985), a Gaussian kernel was derived instead of the exponential one in Eq. (3.10). This leads to the "Gaussian sum rules", which can be given as,

$$\Pi(s,\tau) = \frac{1}{\sqrt{4\pi\tau}}\int_0^\infty dt e^{-\frac{(t-s)^2}{4\tau}}\rho^J(t), \tag{3.15}$$

where s and τ are free parameters and correspond to the Borel mass M in the Borel sum rule case. The advantage of this sum rule is that two parameters can be varied,

allowing one to extract more information on the spectral function $\rho^J(t)$ from the sum rule. Furthermore, the kernel of the Gaussian sum rule has a distinct peak at $t = s$, which means that the various peak structures that might be present in the spectral function are more likely to be preserved in $\Pi(s, \tau)$, rather than washed out as it is often the case for the Borel sum rule.

To summarize, Eqs. (3.10) and (3.15) are the final forms of the sum rules that will be repeatedly appear in this work. The goal is now to analyze these equations and to extract as much information as possible from them with the least amount of artificial assumptions. Our proposed approach of how to deal with this task will be discussed in Chap. 4 of this thesis.

3.2 More on the Operator Product Expansion

The operator product expansion (OPE) has already appeared in Eq. (3.3) of the preceding section. As this expansion plays a key role for the formulation of QCD sum rules, we will discuss it in more detail below.

3.2.1 Theoretical Foundations

The original idea of the OPE, proposed by Wilson (1969), can be formulated as follows:

$$\mathrm{T}\big[\widehat{A}(x)\widehat{B}(y)\big] \xrightarrow{x\to y} \sum_n C_n(x-y)\,\widehat{O}_n\Big(\frac{x+y}{2}\Big). \tag{3.16}$$

Spelled out in words, the above statement means that when two local operators are located at nearby space-time points, they can be rewritten as a series of local operators, defined at the average point between the two original operators, with respective numerical coefficients (the "Wilson coefficients"), which can be calculated perturbatively, due to asymptotic freedom of QCD. Fourier transforming the relation of Eq. (3.16) into momentum space and assuming that the local operators $\widehat{O}_n$ do not depend on the space-time coordinates, we get,

$$i\int d^4x e^{iq(x-y)}\mathrm{T}\big[\widehat{A}(x)\widehat{B}(y)\big] \xrightarrow{|q^2|\to\infty} \sum_n C_n(q^2)\,\widehat{O}_n. \tag{3.17}$$

It can be explicitly shown that the correlator at $|q^2| \to \infty$ corresponds to the short distance limit of Eq. (3.16) (Colangelo and Khodjamirian 2001). This Fourier transformed OPE makes it more apparent that one here essentially carries out a division of scales (also referred to as "factorization"), meaning that the perturbative high energy part goes into the Wilson coefficients, while the low energy parts contribute to the local operators.

The OPE is formulated to be a relation between operators, therefore it does not depend on what kind of matrix elements one calculates from Eqs. (3.16) and (3.17). Thus, the OPE provides some universal relations between various different matrix elements, which belong to different processes. These relations are exploited especially in the analysis of deep inelastic scattering, where the OPE has proven to be a very useful tool (Muta 1998). However, one should keep in mind that, even today, the formulation of the OPE is in fact not yet complete. Even though the correctness of the OPE was proven by Zimmermann in a completely perturbative regime (Zimmermann 1970), no proof for it is given for the case when non-perturbative effects are included. This situation is not very satisfactory and a proof for the general non-perturbative case is certainly desirable. However, the success of the OPE in its many places of application does at least provide convincing evidence for its validity.

3.2.1.1 Ordering of Operators According to Their Dimension

To understand why the OPE is a valid expansion and why usually only the first few terms suffice to obtain reliable results, one can refer to an easy argument by dimensional analysis. In the short distance (or large energy) limit $x \to y$, quark masses and other low energy scales can be ignored, meaning that the only dimensionful quantity appearing in the Wilson coefficients is $(x-y)^2$. Thus, in this limit, the functional dependence of $C_n(x-y)$ is fixed by its dimension. Denoting the mass dimensions of $\widehat{A}(x)$, $\widehat{B}(y)$ and $\widehat{O}_n$ as d_A, d_B and d_n, respectively, $C_n(x-y)$ will behave in the following way:

$$C_n(x-y) \xrightarrow{x\to y} \left(\frac{1}{(x-y)^2}\right)^{(d_A+d_B-d_n)/2}. \tag{3.18}$$

As the dimension of the condensates becomes larger, the corresponding singularities are weakened, until at a certain dimension they vanish completely. In momentum space the above relation is rewritten as

$$\begin{aligned} C_n(q^2) &\xrightarrow{|q^2|\to\infty} (q^2)^{(d_A+d_B-d_n-4)/2}\log(-q^2) \\ &\qquad (d_A+d_B-d_n-4 \geq 0), \\ C_n(q^2) &\xrightarrow{|q^2|\to\infty} (q^2)^{(d_A+d_B-d_n-4)/2} \\ &\qquad (d_A+d_B-d_n-4 < 0). \end{aligned} \tag{3.19}$$

Here, it is seen that the in the limit of $|q^2| \to \infty$, the terms corresponding to the operators with the lowest dimension d_n dominate the expansion.

It is, however, important to note that in reality, the behavior of the OPE is not as simple as the above argument suggests. Most importantly, the numerical factors appearing in front of the expressions of Eq. (3.19) can be large and thus distort the ordering of the various terms according to their dimension. As was pointed out in

Ioffe (1981a, b), the lowest order terms usually contain a number of loops, which are numerically suppressed due to the momentum integral related to the loops. At higher orders of the OPE, these loops are cut, which leads to an enhancement of higher order terms compared to the leading order. Therefore, as a simple rule, one should include all orders in the OPE until the loops are cut, to make sure the expansion converges.

3.2.1.2 Accurate Treatment of Factorization

After the OPE was applied to low energy QCD in the framework of QCD sum rules, some criticism was raised by several authors about the validity of this expansion (Quinn and Gupta 1982; David 1984), pointing out that there are problems for the OPE in a theory with spontaneously broken symmetries and that there are ambiguities in the definition of condensates due to IR singularities. Later, these issues were addressed in Novikov et al. (1985), in which an unambiguous definition of the OPE was given. Essentially, the idea is that the simple expression of Eq. (3.3) is in fact not completely correct, but should be written down including a scale μ, specifying the energy at which one separates the high- and low-energy contributions:

$$\Pi(q^2) = \sum_d C_d(q^2, \mu)\langle 0|O_d|0\rangle(\mu). \tag{3.20}$$

Here, both the Wilson coefficients $C_d(q^2, \mu)$ and the condensates $\langle 0|O_d|0\rangle(\mu)$ depend on the factorization scale μ. However, as a whole, this expression should not depend on μ and hence these dependencies of the different parts cancel.

What this means is that the factorization of the OPE should *not* be considered to be a division between perturbative and non-perturbative contributions, but rather between high-energy ("hard") and low-energy ("soft") parts. How this works has been beautifully shown in Shifman (1998), where, as an example, the gluonic contribution to a mesonic correlator was discussed and the division between the hard and soft domain was explicitly carried out. Here, we will only explain the general idea of this procedure. Consider for instance the integral,

$$\int_0^\infty dq^2 G_{\text{full}}(q^2), \tag{3.21}$$

in which $G_{\text{full}}(q^2)$ is a full propagator (or a combination of several of them) contributing to some correlator. One can now easily divide this integral into a hard and a soft part, the division scale being μ:

$$\int_0^\infty dq^2 G_{\text{full}}(q^2) = \int_{\mu^2}^\infty dq^2 G_{\text{full}}(q^2) + \int_0^{\mu^2} dq^2 G_{\text{full}}(q^2). \tag{3.22}$$

The hard part can be replaced by the corresponding perturbative propagator, while the soft part is parametrized by the appropriate condensates $\langle O_G\rangle(\mu)$:

$$\int_0^\infty dq^2 G_{\text{full}}(q^2) \simeq \int_{\mu^2}^\infty dq^2 G_{\text{pert}}(q^2) + \langle O_G \rangle(\mu). \tag{3.23}$$

Formally, this is the correct way to implement the factorization in some OPE calculation. However, in practice such a clear division is not always feasible and one usually follows a much simpler method. Namely, one carries out the full perturbative calculation without the lower boundary μ^2:

$$\begin{aligned}\int_0^\infty dq^2 G_{\text{full}}(q^2) &\simeq \int_0^\infty dq^2 G_{\text{pert}}(q^2) + \left\{ \langle O_G \rangle(\mu) - \int_0^{\mu^2} dq^2 G_{\text{pert}}(q^2) \right\} \\ &= \int_0^\infty dq^2 G_{\text{pert}}(q^2) + \langle O_G^{'} \rangle(\mu).\end{aligned} \tag{3.24}$$

In this way, the perturbative part of the low energy domain is included into the effective condensates $\langle O_G^{'} \rangle(\mu)$. Of course, this treatment is only possible if the perturbative part below μ contains no divergences. If it does, one has to go back to the original form of the factorization, given in Eq. (3.23).

In Novikov et al. (1985), some arguments are given that in QCD, in contrast to other cases such as the $\lambda\phi^4$ theory or the $O(N)$ sigma model at low N, the perturbative effects on the effective condensates $\langle O_G^{'} \rangle(\mu)$ are small and can be essentially neglected. This situation can be related to the existence of a suitable "window" of the factorization scale μ (not to be confused with the "Borel window" appearing in conventional QCD sum rule analyses). Within this window, μ is large enough so that $\alpha_s(\mu)/\pi \ll 1$ and at the same time small enough so that the perturbative contributions to $\langle O_G^{'} \rangle(\mu)$ are small.

3.2.2 Calculation of Wilson Coefficients

Here, we show the fundamental steps of how to obtain the various Wilson coefficients. Derivations of some formulae appearing in this section can be found in the appendices. We here concentrate on the light quark sector, in which the quark masses can be treated as small perturbations. Calculations using heavy quark masses are discussed for instance in Novikov et al. (1984).

3.2.2.1 Fock-Schwinger Gauge

The Fock-Schwinger gauge (Fock 1937; Schwinger 1954) is defined as

$$(x - x_0)^\mu A_\mu^a(x) = 0, \tag{3.25}$$

where x_0 is a space-time point that can be freely chosen. For simplicity, it is usually taken to be 0, which will be done in the following discussion. The merit of using this gauge is that one can derivatively expand the gluon and quark fields in a covariant way, which was first pointed out in Cronstrom (1980) and further discussed in Dubovikov and Smilga (1981), Shifman (1980). Specifically, one obtains,

$$A_\mu^a(x) = \frac{1}{2}x^\nu G_{\nu\mu}^a(0) + \frac{1}{3}x^\nu x^\alpha [D_\alpha G_{\nu\mu}(0)]^a + \frac{1}{8}x^\nu x^\alpha x^\beta [D_\alpha D_\beta G_{\nu\mu}(0)]^a + \cdots \tag{3.26}$$

for the gluon fields and

$$q(x) = q(0) + x^\alpha D_\alpha q(0) + \frac{1}{2}x^\alpha x^\beta D_\alpha D_\beta q(0) + \cdots \tag{3.27}$$

for the quark fields. Note that even though we have used the same notation, the covariant derivatives in Eqs. (3.26) and (3.27) belong to the adjoint and fundamental representation, respectively. A detailed derivation of the above equations is given in Appendix B.

3.2.2.2 The Quark-Propagator in a Non-perturbative Background

As a first step of the actual OPE calculation, it is convenient to derive the quark propagator in the non-perturbative QCD vacuum, which can then be used repeatedly. For this, we will rely on the Fock-Schwinger gauge and treat the quark mass m_q as a small parameter, retaining only terms up to the first power of m_q. Furthermore, we give all the results in the x-representation, which will proof to be useful in later calculations.

Relegating the derivation to Appendix C, we here directly give the final form of the quark propagator in a non-perturbative background:

$$\begin{aligned}
&\langle 0|T[q_i^a(x)\overline{q}_j^b(0)]|0\rangle \\
&\approx \frac{i}{2\pi^2 x^4}(\not{x})_{ij}\delta^{ab} - \frac{m_q}{2^2\pi^2 x^2}\delta_{ij}\delta^{ab} - \frac{ig}{2^5\pi^2 x^2}G_{\mu\nu}^k(\sigma^{\mu\nu}\not{x} + \not{x}\sigma^{\mu\nu})_{ij}\left(\frac{\lambda^k}{2}\right)^{ab} \\
&- \frac{gm_q}{2^5\pi^2}\ln\left(-\frac{x^2\Lambda^2}{4} + 2\gamma_{EM}\right)G_{\mu\nu}^k(\sigma^{\mu\nu})_{ij}\left(\frac{\lambda^k}{2}\right)^{ab} \\
&- \frac{1}{2^2 3}\delta_{ij}\delta^{ab}\langle\overline{q}q\rangle + \frac{im_q}{2^4 3}(\not{x})_{ij}\delta^{ab}\langle\overline{q}q\rangle - \frac{x^2}{2^6 3}\delta_{ij}\delta^{ab}\langle\overline{q}g\sigma Gq\rangle \\
&+ \frac{im_q x^2}{2^7 3^2}(\not{x})_{ij}\delta^{ab}\langle\overline{q}g\sigma Gq\rangle - \frac{\pi^2 x^4}{2^8 3^3}\delta_{ij}\delta^{ab}\langle\overline{q}q\rangle\langle\frac{\alpha_s}{\pi}G^2\rangle \\
&+ \mathcal{O}(m_q^2, g^2)
\end{aligned} \tag{3.28}$$

Here, a, b are color indices, while i, j stand for the Dirac indices. Furthermore, γ_{EM} is the Euler-Mascheroni constant, which has the value $\gamma_{EM} = 0.57721\ldots$. The first term in the third line of Eq. (3.28) with $\ln\left(-\frac{x^2\Lambda^2}{4} + 2\gamma_{EM}\right)$ corresponds to an IR singularity, which arises because of the perturbative treatment of the quark mass m_q. Using the techniques explained in Appendix C, one can in principle systematically calculate higher orders, which is indeed necessary in certain cases.

3.2.2.3 Non-perturbative Coupling of Quarks and Gluons

As one further step, the non-perturbative contractions of a gluon and a quark must be calculated. These will, at least in leading order, be proportional to the mixed condensate $\langle \overline{q} g\sigma Gq \rangle$. For this we have to carry out the OPE of the following operator:

$$\begin{aligned}\langle 0|q_i^a(x) g G_{\mu\nu}^k(0) \overline{q}_j^b(0)|0\rangle &= \langle 0|q_i^a(0) g G_{\mu\nu}^k(0) \overline{q}_j^b(0)|0\rangle \\ &\quad + x^\alpha \langle 0|D_\alpha q_i^a(0) g G_{\mu\nu}^k(0) \overline{q}_j^b(0)|0\rangle + \cdots \end{aligned} \tag{3.29}$$

For these first two terms, the derivation can be found in the Appendix D, where their final form is obtained as

$$\begin{aligned}\langle 0|q_i^a(x) g G_{\mu\nu}^k(0) \overline{q}_j^b(0)|0\rangle \approx &-\frac{(\sigma_{\mu\nu})_{ij}}{2^6 3}\left(\frac{\lambda^k}{2}\right)^{ab} \langle \overline{q} g\sigma Gq\rangle \\ &+ \frac{i m_q}{2^8 3}(\not{x}\sigma^{\mu\nu} - \sigma^{\mu\nu}\not{x})_{ij}\left(\frac{\lambda^k}{2}\right)^{ab} \langle \overline{q} g\sigma Gq\rangle \\ &+ \mathcal{O}(m_q^2, g^2). \end{aligned} \tag{3.30}$$

3.2.2.4 A Simple Example

To show the procedure of an actual calculation, we demonstrate in this subsection the OPE of the vector current, composed of an (anti-) u and a d quark, which strongly couples to the ρ-meson in the low energy domain. The OPE of this current will be analyzed by the MEM technique in Chap. 5 of this thesis.

Our aim here is to carry out the OPE of the following correlator:

$$\Pi_{\mu\nu}(q) = i \int d^4 e^{iqx} \langle 0|\mathrm{T}[j_\mu(x) j_\nu^\dagger(0)]|0\rangle = (q_\mu q_\nu - q^2 g_{\mu\nu})\Pi(q^2). \tag{3.31}$$

Here the operator $j_\mu(x)$ is defined as $j_\mu(x) = \bar{u}(x)\gamma_\mu d(x)$. The last equality in the above equation follows from the fact that $j_\mu(x)$ is a conserved current. Therefore, to simplify the calculations, we can directly work with the contracted correlator,

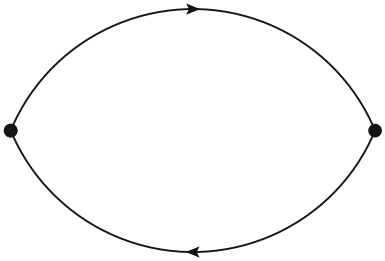

Fig. 3.1 The leading perturbative diagram contributing to the OPE of the vector current correlator of Eq. (3.31)

$$\Pi^{\mu}_{\mu}(q) = -3q^2\Pi(q^2). \tag{3.32}$$

For light quarks, it is most convenient to perform the OPE directly in x-space and to implement the Fourier transform only at the very end of the calculation, which we will do in the following.

The first step is to contract the quark operator by the Wick theorem. This gives

$$\langle 0|\mathrm{T}[j^{\mu}(x)j^{\dagger}_{\mu}(0)]|0\rangle = -\mathrm{Tr}[S(x)\gamma^{\mu}S(-x)\gamma_{\mu}], \tag{3.33}$$

where the trace Tr is taken for both spinor and color indices. $S(x)$ stands for the quark propagator of Eq. (3.28), which contains both perturbative and non-perturbative contributions.

Let us now calculate the leading perturbative term, which corresponds to the Feynman diagram of Fig. 3.1. For this, we simply have to substitute the first term of Eq. (3.28) into Eq. (3.33) and take the traces, which leads to

$$\langle 0|\mathrm{T}[j^{\mu}(x)j^{\dagger}_{\mu}(0)]|0\rangle\Big|_{\text{leading pert.}} = \frac{6}{\pi^4}\frac{1}{x^6}. \tag{3.34}$$

To calculate the first order α_s correction of the perturbative term, one needs to calculate the diagrams shown in Fig. 3.2. This calculation has to be done in momen-

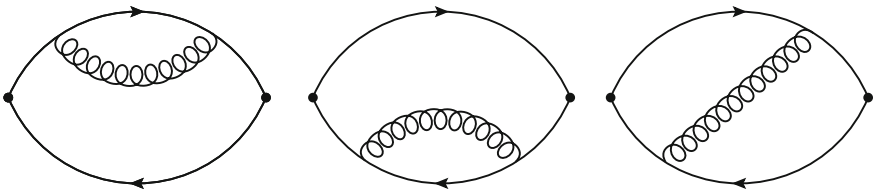

Fig. 3.2 The diagrams contributing to the first order α_s correction of the perturbative OPE term

tum space and is quite involved. Therefore, we do not discuss it here and refer the reader to literature (Colangelo and Khodjamirian 2001; Schwinger 1998) for details.

The next term would in principle be the one with the quark condensate with mass dimension 3. This term, however vanishes due to chiral symmetry. Therefore, we now consider the next order, which involves condensates with mass dimension 4. For demonstration of the method, we calculate the Wilson coefficient for the gluon condensate, for which we have to evaluate the diagrams shown in Fig. 3.3. Here we notice that in the limit of vanishing quark mass, the quark propagator with two attached gluons vanishes (Ioffe et al. 2010) and is therefore strongly suppressed for light quarks. Therefore, we only have to calculate the graph on the left side of Fig. 3.3. (Note, however, that for heavy quarks, all three graphs give contributions of comparable size and thus have to be taken into account (Novikov et al. 1984).)

The concrete evaluation of the relevant diagram of Fig. 3.3 is then quite simple. One substitutes the third term of Eq. (3.28) into Eq.(3.33), takes the traces and contracts the gluon operators according to

$$\langle G^k_{\mu\nu} G^l_{\rho\sigma} \rangle = \frac{\delta^{kl}}{2^5 3} (g_{\mu\rho} g_{\nu\sigma} - g_{\mu\sigma} g_{\nu\rho}) \langle G^2 \rangle. \tag{3.35}$$

At the end of the calculation, we get

$$\langle 0|\mathrm{T}[j^\mu(x)\, j^\dagger_\mu(0)]|0\rangle \Big|_{\text{gluon cond.}} = -\frac{1}{2^4} \frac{1}{x^2} \langle \frac{\alpha_s}{\pi} G^2 \rangle. \tag{3.36}$$

The calculation can be continued to include more higher order terms, up to the point when the expansion is expected to converge. As a last step, one then Fourier transforms the obtained results back to momentum space with the help of the formulae given in Appendix F.

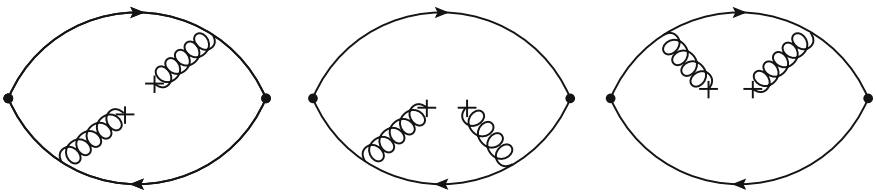

Fig. 3.3 The diagrams contributing to the Wilson coefficient of the gluon condensate. In the light quark case, the diagram on the *left* gives the dominant contribution, while the other two are suppressed

3.3 More on the QCD Vacuum

The vacuum expectation values of Lorentz- and gauge-invariant operators are fundamental parameters of QCD. As they are genuinely non-perturbative objects, their values cannot be calculated analytically and have to be determined by more involved methods. In the following, the values of the most important condensates and their derivation is briefly reviewed.

3.3.1 The Quark Condensate

As already discussed in Sect. 2.3.2, a finite value of the quark condensate signals the spontaneous breaking of the chiral symmetry. $\langle \overline{q}q \rangle$ is therefore probably the most important and most well-known of all the condensates. Its properties have been discussed already many years ago, and its value (for the light quarks) can be obtained via the Gell-Mann-Oakes-Renner relation of Eq. (2.23), approximately giving

$$\langle \overline{q}q \rangle = -(240 \pm 10\,\mathrm{MeV})^3. \tag{3.37}$$

In all the considerations above, the exact realization of the isospin symmetry has been assumed. Moreover, the value of (3.37) is in fact scale dependent due to logarithmical quantum corrections and thus, a renormalization scale has to be given to make it fully determined. The value given above corresponds to the scale of $\mu \sim 1\,\mathrm{GeV}$.

For determining the condensate of the strange quark condensate $\langle \overline{s}s \rangle$, mainly the QCD sum rule analyses of the octet strange baryons have been used (Reinders et al. 1985). These give

$$\frac{\langle \overline{s}s \rangle}{\langle \overline{q}q \rangle} = 0.8 \pm 0.2. \tag{3.38}$$

This is one example, where the value of a condensate actually has been constrained by a QCD sum rule result.

3.3.2 The Gluon Condensate

The gluon condensate $\langle \frac{\alpha_s}{\pi} G^2 \rangle$ is another important quantity of QCD, closely connected to the trace of the energy-momentum tensor. The derivation of its value has first been given in the founding papers of the QCD sum rule approach (Shifman et al. 1979a, b). There, a calculation of the charmonium sum rules was used to extract the value of $\left\langle \frac{\alpha_s}{\pi} G^2 \right\rangle$. As a result

$$\left\langle \frac{\alpha_s}{\pi} G^2 \right\rangle = (0.012 \pm 0.004)\,\mathrm{GeV}^4 \tag{3.39}$$

was obtained. Some later evaluations lead to values quite close to the one above used (Reinders et al. 1985), while others have yielded values about two times larger (Narison 2004), four to five times larger (Marrow et al. 1987), about two times smaller (Ioffe et al. 2010) or even consistent with zero (Brodsky and Shrock 2011). Thus, it has to be said that the uncertainty in the value of $\left\langle \frac{\alpha_s}{\pi} G^2 \right\rangle$ is still very large.

3.3.3 The Mixed Condensate

The mixed condensate $\langle \overline{q} g\sigma Gq \rangle$ is a further quantity, which appears in the OPE of QCD sum rules. It is constructed from quarks and gluons and is therefore called the mixed condensate. Although its dimension is five and it thus appears only at relatively high orders compared to the above two condensates, it actually plays a quite important role in the determination of the masses of various baryons.

For the light u, d quarks, the mixed condensate is usually parametrized relative to the quark condensate:

$$\frac{\langle \overline{q} g\sigma Gq \rangle}{\langle \overline{q} q \rangle} \equiv m_0^2. \tag{3.40}$$

The value of m_0^2 has been estimated in the early days of the QCD sum rule approach (Belyaev and Ioffe 1982) and has since not much changed:

$$m_0^2 = (0.8 \pm 0.1)\,\mathrm{GeV}^2. \tag{3.41}$$

There is also an estimate of $\langle \overline{q} g\sigma Gq \rangle$ by lattice QCD (Doi et al. 2003), which gives a value much larger than the one obtained by QCD sum rules ($\sim$2.5 GeV^2). This disagreement between the results of these two approaches certainly needs further investigation. We will use the value of (3.41) in our calculations.

The mixed condensate for the strange quarks $\langle \overline{q} g\sigma Gq \rangle$ is parametrized similarly to Eq. (3.40):

$$\frac{\langle \overline{s} g\sigma Gs \rangle}{\langle \overline{s} s \rangle} \equiv m_1^2. \tag{3.42}$$

Its value has been estimated to be the same as for the light quarks (Beneke and Dosch 1992):

$$m_1^2 = (0.8 \pm 0.1)\,\mathrm{GeV}^2. \tag{3.43}$$

3.3.4 Higher Order Condensates

There are two sorts of condensates with mass dimension six. Firstly there is the three-gluon condensate $\langle g^3 f_{abc} G_\mu^{av} G_\nu^{b\lambda} G_\lambda^{c\mu} \rangle$, where f_{abc} are the structure constants

of the $SU(3)$ group. The value of this condensate is not very well known and only a crude estimate based on the dilute instanton gas model exists (Novikov et al. 1979). This model gives

$$\langle g^3 f_{abc} G_\mu^{a\nu} G_\nu^{b\lambda} G_\lambda^{c\mu} \rangle \approx 0.045\,\mathrm{GeV}^6. \tag{3.44}$$

There have also been attempts to calculate this quantity on the lattice (Panagopoulos and Vicari 1990), but the results are not yet conclusive and more research is needed for obtaining a reliable estimate.

Secondly, there are the condensates containing four quarks. Generally, these can be expressed as

$$\langle \overline{q}_\alpha^i \overline{q}_\beta^k q_\gamma^l g_\delta^m \rangle, \tag{3.45}$$

where the color and spinor indices have to be contracted to give a singlet, a task for which there exist various combinations. The resulting condensates are, however, only very poorly known and one usually resorts to the vacuum saturation approximation, which reads as follows:

$$\langle \overline{q}_\alpha^i \overline{q}_\beta^k q_\gamma^l q_\delta^m \rangle \simeq \frac{1}{144} (\delta^{im}\delta^{kl}\delta_{\alpha\delta}\delta_{\beta\gamma} - \delta^{il}\delta^{km}\delta_{\alpha\gamma}\delta_{\beta\delta}) \langle \overline{q} q \rangle^2. \tag{3.46}$$

What this equation essentially means is that one inserts a complete set of states between the two pairs of quarks and assumes that the vacuum part gives the dominant contribution. This approximation can be justified in the limit of an infinite number of colors N_c (Novikov et al. 1984), but it is not clear to what degree it holds for QCD with $N_c = 3$. For parametrizing the violation of this approximation, a factor κ is sometimes introduced, which is simply multiplied to Eq. (3.46). Estimates for this parameter range from close to 1 (Ioffe et al. 2010), $2 \sim 3$ (Narison 2004), up to ~ 6 (Leinweber 1997).

For the operators containing even more quarks or gluons, the vacuum saturation approximation discussed above is frequently used, even though the accuracy of this approximation is not properly justified for the various operators that appear at higher orders. Therefore it has to be admitted that the OPE results at higher orders become quite ambiguous. Thus, it is desirable that the OPE converges sufficiently fast, so that these terms do not have to be taken into account. In all, it can be said that the OPE calculation of higher order terms leaves room for further improvement of the QCD sum rule method.

3.4 Parity Projection for Baryonic Sum Rules

The problem of how to accurately project out positive or negative states from baryonic sum rules has been discussed only quite recently and is still not completely resolved. In this section, we introduce an improved method for constructing the parity projected sum rules, with which we hope to clarify this issue.

3.4.1 The Problem of Parity Projection in Baryonic Sum Rules

In contrast to mesons, the operators with quantum numbers of baryons couple to states of both parity, and one can thus not simply fix this quantum number by choosing an appropriate interpolating field. For illustration, let us consider the spatial inversion operation applied to a general Dirac operator $\eta(t, \mathbf{x})$ for a spin $\frac{1}{2}$ state.

$$\eta(t, \mathbf{x}) \to \pm\gamma_0 \eta(t, -\mathbf{x}) \tag{3.47}$$

The parity of this operator can be switched by simply multiplying γ_5:

$$\gamma_5 \eta(t, \mathbf{x}) \to \pm\gamma_5\gamma_0 \eta(t, -\mathbf{x}) = \mp\gamma_0\gamma_5 \eta(t, -\mathbf{x}). \tag{3.48}$$

Therefore, $\eta(t, \mathbf{x})$ couples to both positive and negative parity states. Supposing that the intrinsic parity of $\eta(t, \mathbf{x})$ is positive, we thus get:

$$\begin{aligned} \langle 0|\eta(x)|n^+(q)\rangle &= \lambda_+^n u_+(p) e^{-iqx}, \\ \langle 0|\eta(x)|n^-(q)\rangle &= \lambda_-^n \gamma_5 u_-(p) e^{-iqx}. \end{aligned} \tag{3.49}$$

Here, $u_\pm(q)$ are Dirac spinors with positive and negative parity and $\lambda_\pm^n$ parameters corresponding to the strength of the coupling of η to the state $|n\rangle$. Using these definitions, the correlator of η will have the following form (in which the continuum states are omitted for simplicity):

$$\begin{aligned} \Pi(q) &= i \int d^4x e^{iqx} \langle 0|T[\eta(x)\overline{\eta}(0)]|0\rangle \\ &= \sum_n \left\{ -|\lambda_+^n|^2 \frac{\not{q} + m_+^n}{q^2 - (m_+^n)^2 + i\varepsilon} - |\lambda_-^n|^2 \frac{\not{q} - m_-^n}{q^2 - (m_-^n)^2 + i\varepsilon} \right\} \\ &\equiv \not{q}\Pi_1(q^2) + \Pi_2(q^2). \end{aligned} \tag{3.50}$$

When only the sum rule for $\Pi_1(q^2)$ is used, it is not possible to determine the parity of the state contributing to some specific pole, as both positive and negative parity states couple to its spectral function:

$$\frac{1}{\pi}\mathrm{Im}\Pi_1(q^2) = \sum_n \left[\left|\lambda_+^n\right|^2 \delta(q^2 - (m_+^n)^2) + \left|\lambda_-^n\right|^2 \delta(q^2 - (m_-^n)^2) \right]. \tag{3.51}$$

In contrast, $\Pi_2(q^2)$ contains the positive and negative parity states with different signs:

$$\frac{1}{\pi}\mathrm{Im}\Pi_2(q^2) = \sum_n \left[\left|\lambda_+^n\right|^2 m_+^n \delta(q^2 - (m_+^n)^2) - \left|\lambda_-^n\right|^2 m_-^n \delta(q^2 - (m_-^n)^2) \right]. \tag{3.52}$$

Thus, the problem of parity projection boils down to consistently disentangling the contributions of positive and negative parity to the spectral functions of Eqs. (3.51) and (3.52).

3.4.2 Use of the "Old Fashioned" Correlator

As a solution to the problem discussed above, the use of the "old fashioned" correlator in the rest frame ($\mathbf{q} = 0$) was proposed in Jido et al. (1996),

$$\begin{aligned}\Pi^{\rm of}(q_0) &= i\int d^4x e^{iqx}\theta(x_0)\langle 0|\eta(x)\bar\eta(0)|0\rangle\Big|_{\mathbf{q}=0}\\ &\equiv \gamma_0\Pi_1^{\rm of}(q_0)+\Pi_2^{\rm of}(q_0),\end{aligned}\tag{3.53}$$

where the essential difference to Eq. (3.50) is the insertion of the Heaviside step-function $\theta(x_0)$ before carrying out the Fourier transform. Let us examine the effect of this step-function. First, we note that $\eta(x)\bar\eta(0)$ can be replaced by the corresponding time-ordered product $T[\eta(x)\bar\eta(0)]$. (Alternatively, we could have started directly from Eq. (3.53) with $T[\eta(x)\bar\eta(0)]$ instead of $\eta(x)\bar\eta(0)$.) Next, we use Eq. (3.50) and the Fourier transform of the Heaviside step-function, giving

$$\begin{aligned}&\Pi^{\rm of}(q_0)\\ &= \int d^4x e^{iqx}\frac{1}{2\pi i}\int dk_0\frac{1}{k_0-i\varepsilon}e^{ik_0x_0}\int\frac{d^4p}{(2\pi)^4}e^{-ipx}\\ &\quad\times\sum_n\left\{-|\lambda_+^n|^2\frac{\not p+m_+^n}{p^2-(m_+^n)^2+i\varepsilon}-|\lambda_-^n|^2\frac{\not p-m_-^n}{p^2-(m_-^n)^2+i\varepsilon}\right\}\Big|_{\mathbf{q}=0}\\ &= \frac{1}{2\pi i}\int dk_0\frac{1}{k_0-i\varepsilon}\sum_n\left\{-|\lambda_+^n|^2\frac{(k_0+q_0)\gamma_0-\mathbf{q}\cdot\gamma+m_+^n}{(k_0+q_0)^2-\mathbf{q}^2-(m_+^n)^2+i\varepsilon}\right.\\ &\quad\left.-|\lambda_-^n|^2\frac{(k_0+q_0)\gamma_0-\mathbf{q}\cdot\gamma-m_-^n}{(k_0+q_0)^2-\mathbf{q}^2-(m_-^n)^2+i\varepsilon}\right\}\Big|_{\mathbf{q}=0}.\end{aligned}\tag{3.54}$$

The integrand of the above result contains three poles, two in the upper half of the imaginary plane of k_0 and one in the lower half. Therefore, closing the contour of integration in the lower half and using the Cauchy theorem, we pick up the residue of one pole and get

$$\begin{aligned}&\Pi^{\rm of}(q_0)\\ &= \frac{1}{2}\sum_n\left\{-|\lambda_+^n|^2\frac{1}{q_0-\sqrt{\mathbf{q}^2+(m_+^n)^2}+i\varepsilon}\left[\gamma_0-\frac{\mathbf{q}\cdot\gamma-m_+^n}{\sqrt{\mathbf{q}^2+(m_+^n)^2}-i\varepsilon}\right]\right.\end{aligned}$$

$$- |\lambda_-^n|^2 \frac{1}{q_0 - \sqrt{\mathbf{q}^2 + (m_-^n)^2} + i\varepsilon} \Big[\gamma_0 - \frac{\mathbf{q}\cdot\boldsymbol{\gamma} + m_-^n}{\sqrt{\mathbf{q}^2 + (m_-^n)^2} - i\varepsilon}\Big]\Bigg\}\Bigg|_{\mathbf{q}=0} \tag{3.55}$$

$$= \frac{1}{2}\sum_n \left\{ -|\lambda_+^n|^2 \frac{1}{q_0 - m_+^n + i\varepsilon}(\gamma_0 + 1) - |\lambda_-^n|^2 \frac{1}{q_0 - m_-^n + i\varepsilon}(\gamma_0 - 1) \right\}.$$

From this result it can be understood that the functions $\Pi_1^{\text{of}}(q_0)$ and $\Pi_2^{\text{of}}(q_0)$ contain only poles in the positive q_0 region and that they are analytic for $\text{Im}\, q_0 \geq 0$.

Furthermore, by applying the projection operator $\frac{1}{2}(\gamma_0 \pm 1)$ to the "old fashioned" correlator of Eq. (3.53) and taking the trace over the spinor indices, we are able to construct functions that only contain positive or negative parity states, as

$$\begin{aligned} \frac{1}{2}\text{Tr}\Big[\frac{1}{2}(\gamma_0 \pm 1)\Pi^{\text{of}}(q_0)\Big] &= \Pi_1^{\text{of}}(q_0) \pm \Pi_2^{\text{of}}(q_0) \equiv \Pi^{\pm}(q_0) \\ &= -\sum_n |\lambda_\pm^n|^2 \frac{1}{q_0 - m_\pm^n + i\varepsilon}. \end{aligned} \tag{3.56}$$

The imaginary parts of $\Pi^{\pm}(q_0)$ defined above then give the desired parity projected spectral functions:

$$\frac{1}{\pi}\text{Im}\,\Pi^{\pm}(q_0) = \sum_n |\lambda_n^+|^2 \delta(q_0 - m_n^{\pm}). \tag{3.57}$$

These can, however, not be calculated directly because perturbation theory is not reliable in the low q_0 region, but only at $|q_0| \to \infty$. We thus have to rely on the analytic properties of $\Pi^{\pm}(q_0)$, which allows us to extract information on the spectral functions via certain sum rules, that we will discuss in the next subsection.

3.4.3 Construction of the Sum Rules

We now use the analyticity of the functions $\Pi^{\pm}(q_0)$ to construct the sum rules. To do that, we first have to remember that there are two distinct ways of expressing $\Pi^{\pm}(q_0)$. The first expression uses the OPE and is written down in the language of the elementary degrees of freedom of QCD. In the following discussion we denote it as $\Pi^{\pm}_{\text{OPE}}(q_0)$. This expression is only reliable at high energies. The second one employs the hadronic degrees of freedom, contained in the experimentally observable physical spectrum. We have already given its concrete form in Eq. (3.56) and in the following denote it as $\Pi^{\pm}_{\text{Phys.}}(q_0)$.

To construct the sum rules, we consider the contour integral

$$\oint_C dq_0 \big[\Pi^{\pm}_{\text{OPE}}(q_0) - \Pi^{\pm}_{\text{Phys.}}(q_0)\big] W(q_0) = 0, \tag{3.58}$$

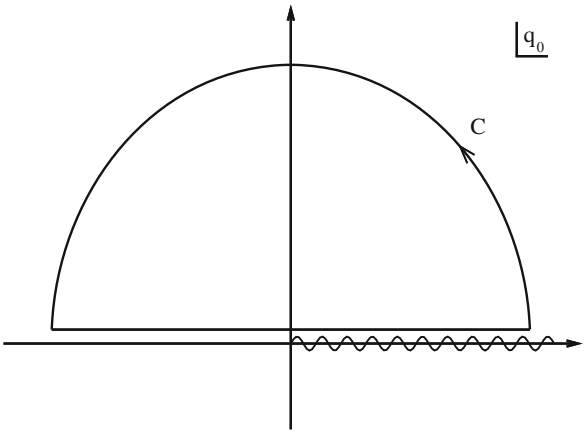

Fig. 3.4 The contour integral C on the complex plane of the variable q_0, used in Eq. (3.58). For the actual calculations, the radius of the half circle of C is taken to infinity. The wavy line denotes the non-analytic cut (or poles) of $\Pi^{\pm}(q_0)$ on the positive side of the real axis. Compared to the discussion in the main text, we have here slightly shifted the contour away from the real axis into the upper half of the imaginary plane for better visuality

where the contour C is given in Fig. 3.4. $W(q_0)$ must be analytic in the upper half of the imaginary plane and real on the real axis. As long as it satisfies these conditions, it can be arbitrarily chosen. That Eq. (3.58) gives 0 follows from the analyticity of both $\Pi^{\pm}_{\text{OPE}}(q_0)$ and $\Pi^{\pm}_{\text{Phys.}}(q_0)$. In other words, there are no poles or cuts inside of the contour C. We know from asymptotic freedom that the perturbative expression $\Pi^{\pm}_{\text{OPE}}(q_0)$ approaches the hadronic one at $|q_0| \to \infty$. Thus, the integrand of Eq. (3.58) vanishes on the half circle of C, whose radius is taken to infinity. Therefore, we are left with the section of C which runs along the real axis and can write down the sum rule as

$$\int_{-\infty}^{\infty} dq_0 \Pi^{\pm}_{\text{Phys.}}(q_0) W(q_0) = \int_{-\infty}^{\infty} dq_0 \Pi^{\pm}_{\text{OPE}}(q_0) W(q_0). \tag{3.59}$$

Finally, taking the imaginary part of the above equation we get

$$\int_{0}^{\infty} dq_0 \rho^{\pm}_{\text{Phys.}}(q_0) W(q_0) = \int_{-\infty}^{\infty} dq_0 \rho^{\pm}_{\text{OPE}}(q_0) W(q_0), \tag{3.60}$$

where the definitions $\rho^{\pm}_{\text{Phys.}}(q_0) \equiv \frac{1}{\pi}\text{Im}\Pi^{\pm}_{\text{Phys.}}(q_0)$ and $\rho^{\pm}_{\text{OPE}}(q_0) \equiv \frac{1}{\pi}\text{Im}\Pi^{\pm}_{\text{OPE}}(q_0)$ were used. Moreover, on the left-hand side, we have exploited the fact that $\Pi^{\pm}_{\text{Phys.}}(q_0)$ only has poles on the positive side of the real axis (see Eq. (3.56)) and have hence restricted the integral to this region.

The authors of Jido et al. (1996) have also restricted to integral on the right-hand side of Eq. (3.60) to positive values. Even though this procedure is roughly correct,

it leads to ambiguities in the higher order OPE terms, which in the chiral limit have poles at $q_0 = 0$. The problem here is that one does not know, whether these poles should be included into the integral of positive values, as they lie just on the border. Furthermore, in Jido et al. (1996) the imaginary part of the time ordered correlator was used instead of $\mathrm{Im}\Pi^{\pm}_{\mathrm{OPE}}(q_0)$, which in principle should be derived from the "old fashioned" correlator of Eq. (3.53). As was pointed out in Kondo et al. (2006), it is not entirely clear whether this prescription is justifiable. Therefore, in this study, we implement two essential improvements as compared to Jido et al. (1996): (1) We do not use the time ordered correlator, but derive all results directly from the "old fashioned" correlator of Eq. (3.53). (2) We do not restrict the region of integration of the OPE side of Eq. (3.60) to positive values and therefore remove the ambiguities that might occur for higher order OPE terms. How this is done will be explained in the next subsection.

3.4.4 General Analysis of the Sum Rules for Three-Quark Baryons

The OPE result of a time ordered correlator with interpolating fields containing three quarks can in coordinate space generally be expressed as

$$\begin{aligned}\Pi(x) = \big[&C_x^{(0)}\frac{1}{(x^2-i\varepsilon)^5} + C_x^{(4)}\frac{1}{(x^2-i\varepsilon)^3}\\ &+ C_x^{(6)}\frac{1}{(x^2-i\varepsilon)^2} + C_x^{(8)}\frac{1}{x^2-i\varepsilon} + \cdots\big] \not{x}\\ &+ C_x^{(3)}\frac{i}{(x^2-i\varepsilon)^4} + C_x^{(5)}\frac{i}{(x^2-i\varepsilon)^3}\\ &+ C_x^{(7)}\frac{i}{(x^2-i\varepsilon)^2} + C_x^{(9)}\frac{i}{x^2-i\varepsilon} + \cdots\end{aligned} \tag{3.61}$$

Here, $C_x^{(n)}$ are constants containing condensates with a total mass dimension n and dimensionless numerical factors. This equation is only correct as long as we work at leading order in α_s for the Wilson coefficients, because higher order corrections may involve additional logarithmic dependencies on x^2. In momentum space, this gives,

$$\begin{aligned}\Pi(q) = \big[&C_q^{(0)}q^4\ln(-q^2-i\varepsilon) + C_q^{(4)}\ln(-q^2-i\varepsilon)\\ &+ C_q^{(6)}\frac{1}{q^2+i\varepsilon} + C_q^{(8)}\frac{1}{(q^2+i\varepsilon)^2} + \cdots\big] \not{q}\\ &+ C_q^{(3)}q^2\ln(-q^2-i\varepsilon) + C_q^{(5)}\ln(-q^2-i\varepsilon)\\ &+ C_q^{(7)}\frac{1}{q^2+i\varepsilon} + C_q^{(9)}\frac{1}{(q^2+i\varepsilon)^2} + \cdots,\end{aligned} \tag{3.62}$$

where, as above, $C_q^{(n)}$ contains condensates and numerical factors. Furthermore, we here have neglected all polynomials in q^2, as they are not relevant for the further discussion. Also, note that we here have used the conventional Fourier transform, not the one with a step function as in Eq. (3.53).

Starting from these results, the question is now how to calculate the right-hand side of Eq. (3.60), from these results, which is all we need to get to the final form of the sum rules. This is a straightforward mathematical exercise: one plugs Eq. (3.61) into Eq. (3.53) and calculates the necessary integrals. One key ingredient for this calculation is the use of the Fourier transformed Heaviside step function of Eq. (F.6). Furthermore, we also mention that for the terms up to dimension 5, it is convenient to work in coordinate space while for terms with mass dimension 6 or larger, it is better to start from the momentum space expression of Eq. (3.62). We leave the detailed derivation to Appendix F and show here only the final result, which is obtained as

$$\begin{aligned}\Pi^{\text{of}}(q_0) = \Big[& C_q^{(0)} q_0^5 \ln(-q_0 - i\varepsilon) + C_q^{(4)} q_0 \ln(-q_0 - i\varepsilon) + \frac{1}{2} C_q^{(6)} \frac{1}{q_0 + i\varepsilon} \\ & + \frac{1}{4} C_q^{(8)} \frac{1}{\sqrt{\mathbf{q}^2}} \frac{1}{(q_0 - \sqrt{\mathbf{q}^2} + i\varepsilon)^2} \Big|_{\mathbf{q}=0} + \cdots \Big] \gamma_0 \\ & + C_q^{(3)} q_0^2 \ln(-q_0 - i\varepsilon) + C_q^{(5)} \ln(-q_0 - i\varepsilon) \\ & + \frac{1}{2} C_q^{(7)} \frac{1}{\sqrt{\mathbf{q}^2}} \frac{1}{q_0 - \sqrt{\mathbf{q}^2} + i\varepsilon} \Big|_{\mathbf{q}=0} \\ & + \frac{1}{4} C_q^{(9)} \Big[\frac{1}{(\sqrt{\mathbf{q}^2})^2} \frac{1}{(q_0 - \sqrt{\mathbf{q}^2} + i\varepsilon)^2} \\ & - \frac{1}{(\sqrt{\mathbf{q}^2})^3} \frac{1}{q_0 - \sqrt{\mathbf{q}^2} + i\varepsilon} \Big] \Big|_{\mathbf{q}=0} + \cdots \end{aligned} \tag{3.63}$$

Here, we have set $\sqrt{\mathbf{q}^2} = 0$ wherever this limit does not lead to divergencies. As one can see in the above equation, we are seemingly running into problems for terms with dimension 7 and higher, as the limit $\sqrt{\mathbf{q}^2} \to 0$ leads to a divergence for these terms. However, as we will see below, after substituting Eq. (3.63) into Eq. (3.60) and evaluating the integral over q_0, these divergences in fact vanish, leaving only finite expressions for the final form of the sum rules.

Let us now carry out this last step and calculate the right hand side of Eq. (3.60). For this, we consider two classes of weight functions $W(q_0)$, one which is an even function of q_0, $W_{\text{e}}(q_0) = F(q_0^2)$, the other being an odd function, $W_{\text{o}}(q_0) = q_0 F(q_0^2)$. Here, $F(q_0^2)$ should be an analytic function on the upper half of the imaginary plane and take only real values on the real axis.

The contributions of the terms of dimension 0, 3, 4, 5, and 6 to the sum rules can be easily obtained. As can be observed from Eq. (3.63), their imaginary part is either proportional to a step function $\theta(q_0)$ (dimension 0–5) or to a δ function $\delta(q_0)$ (dimension 6). The calculation of the higher order terms is somewhat more involved and we thus explicitly show their evaluation here. First, we note that in the limit $\sqrt{\mathbf{q}^2} \to 0$, the imaginary parts of the terms of dimension $7, 9, \ldots$ are

odd functions of q_0, while the ones of dimension 6, 8, 10, ... are even functions of q_0. A proof for this statement is given in Sect. F.2.3 of Appendix F. We thus can immediately conclude that the terms with dimension 7, 9, ... vanish when the even weight function $W_{\rm e}(q_0)$ is used while the terms of dimension 6, 8, 10, ... vanish if the odd weight function $W_{\rm o}(q_0)$ is used. Therefore, all we have to do is to evaluate the remaining non-vanishing parts.

For dimension 7, we have

$$-\frac{1}{2\sqrt{\mathbf{q}^2}}C_q^{(7)}\int dq_0\delta(q_0-\sqrt{\mathbf{q}^2})q_0F(q_0^2)\Big|_{\mathbf{q}=0}=-\frac{1}{2}C_q^{(7)}F(0). \tag{3.64}$$

Next, dimension 8 gives

$$\begin{aligned}&\frac{1}{4\sqrt{\mathbf{q}^2}}C_q^{(8)}\int dq_0\delta'(q_0-\sqrt{\mathbf{q}^2})F(q_0^2)\Big|_{\mathbf{q}=0}\\&\quad=-\frac{1}{2\sqrt{\mathbf{q}^2}}C_q^{(8)}\int dq_0\delta(q_0-\sqrt{\mathbf{q}^2})q_0F'(q_0^2)\Big|_{\mathbf{q}=0}=-\frac{1}{2}C_q^{(8)}F'(0).\end{aligned} \tag{3.65}$$

Finally, dimension 9 is evaluated as

$$\begin{aligned}&\frac{1}{4}C_q^{(9)}\Bigg[\frac{1}{(\sqrt{\mathbf{q}^2})^2}\int dq_0\delta'(q_0-\sqrt{\mathbf{q}^2})q_0F(q_0^2)\\&\qquad+\frac{1}{(\sqrt{\mathbf{q}^2})^3}\int dq_0\delta(q_0-\sqrt{\mathbf{q}^2})q_0F(q_0^2)\Bigg]\Bigg|_{\mathbf{q}=0}\\&\quad=\frac{1}{4}C_q^{(9)}\Bigg[-\frac{1}{(\sqrt{\mathbf{q}^2})^2}\int dq_0\delta(q_0-\sqrt{\mathbf{q}^2})[F(q_0^2)+2q_0^2F'(q_0^2)]\\&\qquad+\frac{1}{(\sqrt{\mathbf{q}^2})^2}F(0)\Bigg]\Bigg|_{\mathbf{q}=0}=-\frac{1}{2}C_q^{(9)}F'(0).\end{aligned} \tag{3.66}$$

We have explicitly checked that for the dimension 10 term, all potential divergences vanish in a similar fashion. Thus, the same procedure could presumably be continued to even higher orders, but this is not of much practical use as the OPE is only available up to terms of dimension 9 at present.

Let us also mention that if α_s corrections to the Wilson coefficients are included, oddness (evenness) of the imaginary parts of the dimension 7, 9, ... (6, 8, 10, ...) terms no longer holds due to the additional logarithmic dependencies on q_0. Thus, the sum rules with an $W_{\rm e}(q_0)$ ($W_{\rm o}(q_0)$) weight function do get contributions from the dimension 7, 9, ... (6, 8, 10, ...) terms, but only at subleading order in α_s. The detailed form of these α_s corrections are given elsewhere (Ohtani et al. 2013), here we for simplicity concentrate only on the results at leading order.

Assembling the findings of the preceding paragraphs, we can now write down the sum rules. For the even weight function $W_{\rm e}(q_0)$, we get

$$\begin{aligned}\int_0^\infty dq_0 \rho^{\pm}_{\rm Phys.}(q_0) W_{\rm e}(q_0) &= \int_0^\infty dq_0 \rho^{\pm}_{\rm Phys.}(q_0) F(q_0^2) \\ &= -C_q^{(0)} \int_0^\infty dq_0 q_0^5 F(q_0^2) \mp C_q^{(3)} \int_0^\infty dq_0 q_0^2 F(q_0^2) \\ &\quad - C_q^{(4)} \int_0^\infty dq_0 q_0 F(q_0^2) \mp C_q^{(5)} \int_0^\infty dq_0 F(q_0^2) \\ &\quad - \frac{1}{2} C_q^{(6)} F(0) - \frac{1}{2} C_q^{(8)} F'(0) + \cdots ,\end{aligned} \tag{3.67}$$

while for the odd weight function $W_{\rm o}(q_0)$, the sum rules reads as follows:

$$\begin{aligned}\int_0^\infty dq_0 \rho^{\pm}_{\rm Phys.}(q_0) W_{\rm o}(q_0) &= \int_0^\infty dq_0 \rho^{\pm}_{\rm Phys.}(q_0) q_0 F(q_0^2) \\ &= -C_q^{(0)} \int_0^\infty dq_0 q_0^6 F(q_0^2) \mp C_q^{(3)} \int_0^\infty dq_0 q_0^3 F(q_0^2) \\ &\quad - C_q^{(4)} \int_0^\infty dq_0 q_0^2 F(q_0^2) \mp C_q^{(5)} \int_0^\infty dq_0 q_0 F(q_0^2) \\ &\quad \mp \frac{1}{2} C_q^{(7)} F(0) \mp \frac{1}{2} C_q^{(9)} F'(0) + \cdots\end{aligned} \tag{3.68}$$

The sum rule of Eq. (3.67) has been considered in Jido et al. (1996), while the one of Eq. (3.68) is to our knowledge formulated here for the first time. Eq. (3.68) has the disadvantage of an additional power of q_0 compared to Eq. (3.67), which will enhance the contributions from the continuum to the sum rules. On the other hand, the contribution of the four-quark condensate of dimension 6 vanishes (or, the be more accurate, is suppressed by α_s) for Eq. (3.68). As this term contains large uncertainties, Eq. (3.68) has the potential to be more reliable. To establish which sum rule is more useful, a detailed study of both Eqs. (3.67) and (3.68) is certainly necessary. Also note that if one employs the maximum entropy method (to be discussed in the next chapter) to analyze the sum rules, it even becomes possible to study both sum rules at the same time.

As a last point of this section, we briefly touch on the issue of what function should be chosen for the weight function $F(q_0^2)$. The traditionally favored choice has been the Borel weight function, given as

$$F(q_0^2, M) = e^{-q_0^2/M^2} . \tag{3.69}$$

However, our recent study (Ohtani et al. 2011) of the (non parity projected) nucleon sum rules suggests that rather the Gaussian weight function,

$$F(q_0^2, s, \tau) = \frac{1}{\sqrt{4\pi\tau}} e^{-(q_0^2 - s)^2/(4\tau)}, \tag{3.70}$$

is in fact to be preferred. Again, a more detailed discussion of this question has to be relegated to another publication (Ohtani et al. 2013).

References

Belyaev VM, Ioffe BL (1982) Sov Phys JETP 56:493
Beneke M, Dosch HG (1992) Phys Lett B 284:116
Bertlmann RA, Launer G, de Rafael E (1985) Nucl Phys B250:61
Brodsky SJ, Shrock R (2011) Proc Nat Acad Sci 108:45
Colangelo P, Khodjamirian A (2001) At the frontier of particle physics/handbook of QCD, vol 3. World Scientific, Singapore, p 1495
Cronstrom C (1980) Phys Lett B 90:267
David F (1984) Nucl Phys B234:237
Doi T, Ishii N, Oka M, Suganuma H (2003) Phys Rev D 67:054504
Dubovikov MS, Smilga AV (1981) Nucl Phys B185:109
Fock VA (1937) Sowj Phys 12:404
Ioffe BL (1981a) Nucl Phys B188:317
Ioffe BL (1981b) Nucl Phys B 191:591(E)
Ioffe BL, Fadin VS, Lipatov LN (2010) Quantum chromodynamics, perturbative and nonperturbative aspects (Cambridge monographs on particle physics, nuclear physics and cosmology). Cambridge University Press, Cambridge
Jido D, Kodama N, Oka M (1996) Phys Rev D 54:4532
Kojo T (2008) PhD thesis, Kyoto University
Kondo Y, Morimatsu O, Nishikawa T (2006) Nucl Phys A764:303
Leinweber DB (1997) Ann Phys 254:328
Marrow J, Parker J, Shaw G (1987) Z Phys C 37:103
Muta T (1998) Foundations of quantum chromodynamics, 2nd edn. World Scientific, Singapore
Narison S (2004) QCD as a theory of hadrons, from partons to confinement. Cambridge University Press, Cambridge
Novikov VA, Shifman MA, Vainshtein AI, Zakharov VI (1979) Phys Lett B 86:347
Novikov VA, Shifman MA, Vainshtein AI, Zakharov VI (1984) Fortsch Phys 32:585
Novikov VA, Shifman MA, Vainshtein AI, Voloshin MB, Zakharov VI (1984) Nucl Phys B237:525
Novikov VA, Shifman MA, Vainshtein AI, Zakharov VI (1985) Nucl Phys B249:445
Ohtani K, Gubler P, Oka M (2011) Eur Phys J A 47:114
Ohtani K, Gubler P, Oka M (2013) Phys Rev D 87:034027
Panagopoulos H, Vicari E (1990) Nucl Phys B332:261
Pascual P, Tarrach R (1984) QCD: renormalization for the practitioner. Springer, Berlin
Poggio EC, Quinn HR, Weinberg S (1976) Phys Rev D 13:1958
Quinn HR, Gupta S (1982) Phys Rev D 26:499
Reinders LJ, Rubinstein H, Yazaki S (1985) Phys Rep 127:1
Schwinger J (1954) Phys Rev 82:664
Schwinger J (1998) Particles, sources, and fields. Perseus Books, Reading
Shifman MA (1980) Nucl Phys B 173:13

Shifman MA (1998) Prog Theor Phys Suppl 131:1
Shifman MA, Vainshtein AI, Zakharov VI (1979a) Nucl Phys B 147:385
Shifman MA, Vainshtein AI, Zakharov VI (1979b) Nucl Phys B 147:448
Wilson KG (1969) Phys Rev D 179:1499
Zimmermann W (1970) Lectures on elementary particles and quantum field theory. Deser S et al (eds) MIT Press, Cambridge

Chapter 4
The Maximum Entropy Method

4.1 Basic Concepts

In this section, the essential steps of the MEM procedure are reviewed in brief. For details on applications in condensed matter physics, see Jarrel and Gubernatis (1996). Furthermore, Nakahara et al. (1999) and Asakawa et al. (2001) discuss the implementation of this method to lattice QCD. For the application of MEM to Dyson-Schwinger studies, see also Nickel (2007).

The basic problem that is solved with the help of MEM is the following. Suppose one wants to calculate some function $\rho(\omega)$, but has only information about an integral of $\rho(\omega)$ multiplied by a kernel $K(x, \omega)$:

$$G_{\mathrm{OPE}}(x) = \int_0^\infty d\omega K(x, \omega)\rho(\omega) \equiv G_\rho(x). \tag{4.1}$$

Let us first explain the above notation, which will be used throughout this chapter. $G_{\mathrm{OPE}}(x)$ stands for the result of the operator product expansion after having completed the Borel transform or other procedures to derive the final form of the sum rules. On the other hand, $G_\rho(x)$ represents the result of the integration of the spectral function $\rho(\omega)$ over ω with the kernel $K(x, \omega)$ as a weight function and hence implicitly depends on $\rho(\omega)$. Equation (4.1) corresponds to Eqs. (3.10), (3.15) or (3.60), where x becomes either the Borel mass M or s, τ, the parameters appearing in the Gaussian sum rules. When $G_{\mathrm{OPE}}(x)$ is known only with limited precision or is only calculable in a limited range of x, the problem of obtaining $\rho(\omega)$ from $G_{\mathrm{OPE}}(x)$ is ill-posed and will not be analytically solvable.

The idea of the MEM approach is now to use Bayes' theorem, by which additional information about $\rho(\omega)$ such as positivity and/or its asymptotic behavior at small or large energies can be added to the analysis in a systematic way and by which one can

P. Gubler, *A Bayesian Analysis of QCD Sum Rules*, Springer Theses,
DOI: 10.1007/978-4-431-54318-3_4, © Springer Japan 2013

finally deduce the most probable form of $\rho(\omega)$. Bayes' theorem can be written as

$$P[\rho|G_{\mathrm{OPE}}H] = \frac{P[G_{\mathrm{OPE}}|\rho H]P[\rho|H]}{P[G_{\mathrm{OPE}}|H]}, \tag{4.2}$$

where prior knowledge about $\rho(\omega)$ is denoted as H, and $P[\rho|G_{\mathrm{OPE}}H]$ stands for the conditional probability of $\rho(\omega)$ given $G_{\mathrm{OPE}}(x)$ and H. The above equation can be easily derived with the help of the formula of the joint probability of $\rho(\omega)$ and $G_{\mathrm{OPE}}(x)$:

$$P[\rho G_{\mathrm{OPE}}|H] = P[\rho|G_{\mathrm{OPE}}H]P[G_{\mathrm{OPE}}|H] = P[G_{\mathrm{OPE}}|\rho H]P[\rho|H]. \tag{4.3}$$

Maximizing the functional $P[\rho|G_{\mathrm{OPE}}H]$ will give the most probable form of $\rho(\omega)$. $P[G_{\mathrm{OPE}}|\rho H]$ is the "likelihood function" and $P[\rho|H]$ the "prior probability". Ignoring the prior probability and maximizing only the likelihood function corresponds to the ordinary χ^2-fitting. The constant term $P[G_{\mathrm{OPE}}|H]$ in the denominator is just a normalization constant and can be dropped as it does not depend on $\rho(\omega)$.

4.1.1 The Likelihood Function and the Prior Probability

We will now discuss the likelihood function and the prior probability one after the other. Considering first the likelihood function, it is assumed that the function $G_{\mathrm{OPE}}(x)$ is distributed according to uncorrelated Gaussian distributions. For our analysis of QCD sum rules, we will numerically generate uncorrelated Gaussianly distributed values for $G_{\mathrm{OPE}}(x)$, which satisfy this assumption. We can therefore write for $P[G_{\mathrm{OPE}}|\rho H]$:

$$\begin{aligned} P[G_{\mathrm{OPE}}|\rho H] &= e^{-L[\rho]}, \\ L[\rho] &= \frac{1}{2(x_{\max} - x_{\min})} \int_{x_{\min}}^{x_{\max}} dx \frac{\left[G_{\mathrm{OPE}}(x) - G_\rho(x)\right]^2}{\sigma^2(x)}. \end{aligned} \tag{4.4}$$

We here note that if $G_{\mathrm{OPE}}(x)$ is evaluated using Monte-Carlo methods, the correlation between the values of $G_{\mathrm{OPE}}(x)$ at different x have to be taken into account by the use of the covariance matrix (Jarrel and Gubernatis 1996; Asakawa et al. 2001). $\sigma(x)$ stands for the uncertainty of $G_{\mathrm{OPE}}(x)$ at the corresponding values of x. In practice, we will discretize both the integrals in Eqs. (4.1) and (4.4) and take N_ω data points for $\rho(\omega)$ in the range $\omega_{\min} \sim \omega_{\max}$ and N_x data points for $G_{\mathrm{OPE}}(x)$ and $G_\rho(x)$ in the range from $x_{\min}$ to $x_{\max}$. This gives

$$G_{\mathrm{OPE}}(x_i) = \sum_{\omega_j=\omega_{\min}}^{\omega_{\max}} d\omega K(x_i, \omega_j)\rho(\omega_j) \equiv G_\rho(x_i), \tag{4.5}$$

with $d\omega = (\omega_{\max} - \omega_{\min})/N_\omega$ for Eq. (4.1), and

$$L(\rho) = \frac{1}{2(x_{\max} - x_{\min})} \sum_{x_i = x_{\min}}^{x_{\max}} dx \frac{[G_{\mathrm{OPE}}(x_i) - G_\rho(x_i)]^2}{\sigma^2(x_i)}, \tag{4.6}$$

with $dx = (x_{\max} - x_{\min})/N_x$ for Eq. (4.4).

The prior probability parametrizes the prior knowledge of $\rho(\omega)$ such as positivity and the values at the limiting points, and is given by the Shannon-Jaynes entropy $S[\rho]$:

$$\begin{aligned} P[\rho|H] &= e^{\alpha S[\rho]}, \\ S[\rho] &= \int_0^\infty d\omega \left[\rho(\omega) - m(\omega) - \rho(\omega) \log \left(\frac{\rho(\omega)}{m(\omega)} \right) \right]. \end{aligned} \tag{4.7}$$

It can be either derived from the law of large numbers or axiomatically constructed from requirements such as locality, coordinate invariance, system independence and scaling, as shown in Appendix G. For our purposes, it is important to note that this functional gives the most unbiased probability for the positive function $\rho(\omega)$. The scaling factor α that is newly introduced in Eq. (4.7) will be integrated out in a later step of the MEM procedure. The function $m(\omega)$, which is also introduced in Eq. (4.7), is the so-called "default model". In the case of no available data $G_{\mathrm{OPE}}(x)$, the MEM procedure will just give $m(\omega)$ for $\rho(\omega)$ because this function maximizes $P[\rho|H]$. The default model is often taken to be a constant, but one can also use it to incorporate known information about $\rho(\omega)$ into the calculation. In our QCD sum rule analysis, we will sometimes use $m(\omega)$ to fix the value of $\rho(\omega)$ at both very low and large energies. As for the other integrals above, we will discretize the integral in Eq. (4.7) in the actual calculation and approximate it as a sum of N_ω data points, giving

$$S(\rho) = \sum_{\omega_j = \omega_{\min}}^{\omega_{\max}} d\omega \left[\rho(\omega_j) - m(\omega_j) - \rho(\omega_j) \log \left(\frac{\rho(\omega_j)}{m(\omega_j)} \right) \right], \tag{4.8}$$

where $d\omega$ is the same as for Eq. (4.5).

4.1.2 The Numerical Analysis

Assembling the results from the last section, we obtain the final form for the probability $P[\rho|G_{\mathrm{OPE}}H]$:

$$\begin{aligned} P[\rho|G_{\mathrm{OPE}}H] &\propto P[G_{\mathrm{OPE}}|\rho H]P[\rho|H] \\ &= e^{Q[\rho]}, \\ Q[\rho] &\equiv \alpha S[\rho] - L[\rho]. \end{aligned} \tag{4.9}$$

It is now merely a numerical problem to obtain the form of $\rho(\omega)$ that maximizes $Q[\rho]$ and, therefore, is the most probable $\rho(\omega)$ given $G_{\mathrm{OPE}}(M)$ and H. As we show in Appendix H, it can be proven that the maximum of $Q[\rho]$ is unique if it exists and, therefore, the problem of local minima does not occur. For the numerical determination of the maximum of $Q[\rho]$, we use the Bryan algorithm (Bryan 1990), which uses the singular-value decomposition to reduce the dimension of the configuration space of $\rho(\omega)$ and, therefore, largely shortens the necessary calculation time. We have also introduced some slight modifications to the algorithm, the most important one being that once a maximum is reached, we add to $\rho(\omega)$ a randomly generated small function and let the search start once again. If the result of the second search agrees with the first one within the requested accuracy, it is taken as the final result. If not, we add to $\rho(\omega)$ another randomly generated function and start the whole process from the beginning. We found that this convergence criterion stabilizes the algorithm considerably.

Once a $\rho_\alpha(\omega)$ that maximizes $Q[\rho]$ for a fixed value of α is found, this parameter is integrated out by averaging $\rho(\omega)$ over a range of values of α and assuming that $P[\rho|G_{\mathrm{OPE}}H]$ is sharply peaked around its maximum $P[\rho_\alpha|G_{\mathrm{OPE}}H]$:

$$\begin{aligned}\rho_{\mathrm{out}}(\omega) &= \int [d\rho] \int d\alpha \rho(\omega) P[\rho|G_{\mathrm{OPE}}H] P[\alpha|G_{\mathrm{OPE}}H] \\ &\simeq \int d\alpha \rho_\alpha(\omega) P[\alpha|G_{\mathrm{OPE}}H]. \end{aligned} \tag{4.10}$$

This $\rho_{\mathrm{out}}(\omega)$ is our final result. To estimate the above integral, we once again make use of Bayes' theorem to obtain $P[\alpha|G_{\mathrm{OPE}}H]$:

$$\begin{aligned} P[\alpha|G_{\mathrm{OPE}}H] &= \int [d\rho] P[G_{\mathrm{OPE}}|\rho H\alpha] P[\rho|H\alpha] P[\alpha|H]/P[G_{\mathrm{OPE}}|H] \\ &\propto P[\alpha|H] \int [d\rho] e^{Q(\rho)} \\ &\propto P[\alpha|H] \exp\Big[\frac{1}{2}\sum_k \log\frac{\alpha}{\alpha+\lambda_k} + Q[\rho_\alpha]\Big]. \end{aligned} \tag{4.11}$$

Here, λ_k represents the eigenvalues of the matrix

$$\Lambda_{ij} = \sqrt{\rho_i}\frac{\partial^2 L}{\partial\rho_i \partial\rho_j}\sqrt{\rho_j}\bigg|_{\rho=\rho_\alpha}, \tag{4.12}$$

where ρ_i stands for the discretized data points of $\rho(\omega)$: $\rho_i \equiv \rho(\omega_i)\Delta\omega$, with $\Delta\omega \equiv \frac{\omega_{\max}-\omega_{\min}}{N_\omega}$ and $\omega_i \equiv \frac{i}{N_\omega}(\omega_{\max}-\omega_{\min})+\omega_{\min}$. To get to the final form of the lowest line of Eq. (4.11), we have expanded $\rho(\omega)$ around $\rho_\alpha(\omega)$ and carried out the Gaussian integrals over $\delta\rho(\omega) \equiv \rho(\omega) - \rho_\alpha(\omega)$. Furthermore, we made use of the fact that the measure $[d\rho]$ is given as

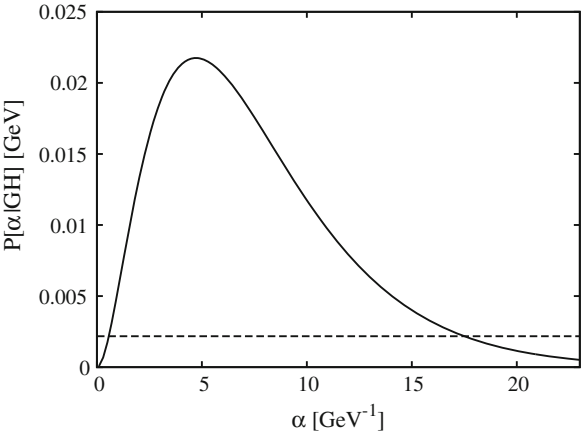

Fig. 4.1 A typical example of the probability $P[\alpha|G_{\rm OPE}H]$, which was obtained from the analysis of mock data of the ρ meson channel. The *dashed line* corresponds to the value of $0.1 \times P[\alpha_{\rm max}|G_{\rm OPE}H]$, from which the boundaries of the integration region, $\alpha_{\rm low}$ and $\alpha_{\rm up}$, are determined

$$[d\rho] \equiv \prod_i \frac{d\rho_i}{\sqrt{\rho_i}}, \tag{4.13}$$

which we have derived in Appendix G. In Eq. (4.11), we still have to specify $P[\alpha|H]$, for which one uses either Laplace's rule with $P[\alpha|H] = $ const. or Jeffreys' rule with $P[\alpha|H] = \frac{1}{\alpha}$. We will employ Laplace's rule throughout this study, but have confirmed that the analysis using Jeffreys' rule gives essentially the same results.

To give an idea of the behavior of the probability $P[\alpha|G_{\rm OPE}H]$, a typical example of this function is shown in Fig. 4.1. Its qualitative structure is the same for all the cases studied in this thesis. To calculate the integral of Eq. (4.10), we first determine the maximum value of $P[\alpha|G_{\rm OPE}H]$, which has a pronounced peak: $P[\alpha_{\rm max}|G_{\rm OPE}H]$. Then, we obtain the lower and upper boundaries of the integration region $(\alpha_{\rm low}, \alpha_{\rm up})$ from the condition $P[\alpha|G_{\rm OPE}H] > 0.1 \times P[\alpha_{\rm max}|G_{\rm OPE}H]$ and normalize $P[\alpha|G_{\rm OPE}H]$ so that its integral within the integration region gives 1. After these preparations, the average of Eq. (4.10) is evaluated numerically.

4.1.3 Error Estimation

As a last step of the MEM analysis, we have to estimate the error of the obtained result $\rho_{\rm out}(\omega)$. The error is calculated for averaged values of $\rho_{\rm out}(\omega)$ over a certain interval (ω_1, ω_2), as shown below.

The variance of $\rho(\omega)$ from its most probable form for fixed α, $\delta\rho(\omega)$, averaged over the interval (ω_1, ω_2), is defined as

$$\langle(\delta\rho)^2\rangle_{\omega_1,\omega_2} \equiv \frac{1}{(\omega_2-\omega_1)^2}\int[d\rho]\int_{\omega_1}^{\omega_2} d\omega d\omega' \delta\rho(\omega)\delta\rho(\omega')P[\rho|G_{\rm OPE}H]$$
$$= -\frac{1}{(\omega_2-\omega_1)^2}\int_{\omega_1}^{\omega_2} d\omega d\omega' \left(\frac{\delta^2 Q}{\delta\rho(\omega)\delta\rho(\omega')}\right)^{-1}\Bigg|_{\rho=\rho_\alpha}, \qquad (4.14)$$

where again the definition of Eq. (4.13) and the Gaussian approximation for the probability $P[\rho|G_{\rm OPE}H]$ were used. Taking the root of this expression and averaging over α in the same way as was done for $\rho_\alpha(\omega)$, we obtain the final result of the error $\delta\rho_{\rm out}(\omega)$, averaged over the interval (ω_1, ω_2):

$$\langle\delta\rho_{\rm out}\rangle_{\omega_1,\omega_2} = \int_{\alpha_{\rm low}}^{\alpha_{\rm up}} d\alpha \sqrt{\langle(\delta\rho)^2\rangle_{\omega_1,\omega_2}}\, P[\alpha|G_{\rm OPE}H]. \qquad (4.15)$$

The interval (ω_1, ω_2) is usually taken to cover a peak or some other structure of interest. The formulas of this subsection will be used to generate the error bars of the various plots of $\rho_{\rm out}(\omega)$, shown in the following parts of this thesis. The use of the error bars is illustrated in Fig. 4.2.

4.2 Sample MEM Analysis of a Toy Model

Before doing the MEM analysis of the actual QCD sum rules, that will be discussed in the second part of this thesis, we here give a brief account of a test analysis of sum rules that can be constructed from a simple non-relativistic potential model. Such sum

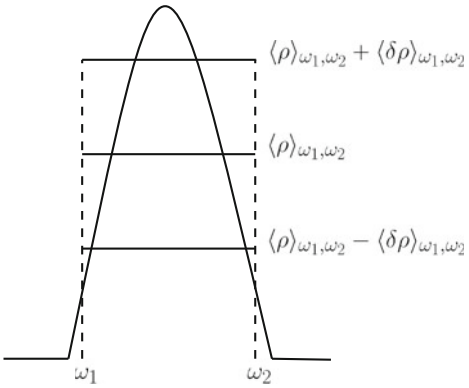

Fig. 4.2 An illustration of the meaning of the error bars used throughout the thesis. Here, the *solid line* which has a *peak* stands for the spectral function $\rho(\omega)$. As discussed in the main text, the *bars* indicate the height of the mean value $\langle\rho\rangle_{\omega_1,\omega_2}$ of the spectral function over the interval (ω_1, ω_2) and the of the corresponding errors added to and subtracted from it, namely $\langle\rho\rangle_{\omega_1,\omega_2} \pm \langle\delta\rho\rangle_{\omega_1,\omega_2}$

rules were discussed already in the early days of QCD sum rule studies (Vainshtein et al. 1980), and have since served as a suitable test case for novel analysis methods in this field (Whitenton et al. 1983; Lucha et al. 2007). We will in this section discuss the MEM analysis of such a potential model. The purpose of this exercise is firstly to investigate what one can hope to achieve with this method and secondly to find out what sort of kernel is most suitable for the MEM analysis of the sum rules.

4.2.1 Construction of the Sum Rules

The general problem to be discussed here can be described by the Schrödinger equation

$$H\psi(\mathbf{x}) \equiv \Big[-\frac{1}{2m}\nabla^2 + V(\mathbf{x})\Big]\psi(\mathbf{x}) = E\psi(\mathbf{x}), \tag{4.16}$$

for which we will specifically treat the case of a three-dimensional oscillator:

$$V(\mathbf{x}) = \frac{m\omega^2}{2}\mathbf{x}^2. \tag{4.17}$$

Let us consider the Greens's function of this system, which we define as

$$G(\mathbf{x}, \mathbf{y}, E) \equiv \langle \mathbf{x}| \frac{1}{E - H} |\mathbf{y}\rangle = \sum_n \frac{\psi_n(\mathbf{x})\psi_n^*(\mathbf{y})}{E - E_n}, \tag{4.18}$$

where we have inserted a complete set of states to obtain the last equation. E_n represents the eigenvalue of the n'th state and ψ_n is the corresponding eigenfunction. Here, we will only use the case, for which $\mathbf{x}$ and $\mathbf{y}$ are 0. This gives,

$$G(E) \equiv G(0, 0, E) = \sum_n \frac{|\psi_n(0)|^2}{E - E_n}. \tag{4.19}$$

The coefficients $|\psi_n(0)|^2$ are known exactly and can be given as

$$|\psi_n(0)|^2 = \frac{(2n+1)!!}{2^n n!}\left(\frac{m\omega}{\pi}\right), \tag{4.20}$$

which scales as $\sqrt{E_n}$ with increasing n. As E_n is certainly real and positive in the present case, we can understand from the above equation that $G(E)$ is analytic in the whole imaginary plane of E except on the positive side of the real axis, where the poles of the eigenstates are located. Therefore, making use of this analyticity and following for instance (Giménez et al. 1991), we can write down a general class of sum rules as

$$\oint_{C(|E|=R)} G(E)W(E)dE = -2i \int_0^R \mathrm{Im}G(E + i\varepsilon)W(E)dE, \tag{4.21}$$

where the weight function $W(E)$ must be analytic in the region where $G(E)$ is analytic and must satisfy $W(E^*)^* = W(E)$. As long as $W(E)$ obeys these constraints, it can be arbitrarily chosen. The contour $C(|E| = R)$ stands for a circle of radius R. We will in the following considerations always take the limit $R \to \infty$.

For the simple harmonic oscillator potential of Eq. (4.17), the Greens's function of Eq. (4.18) and accordingly both sides of Eq. (4.21) are of course already known analytically, the poles appearing in the spectral function (which is proportional to $\mathrm{Im}G(E + i\varepsilon)$) being located at

$$E_n = (2n + \frac{3}{2})\omega \quad (n = 0, 1, 2, \ldots). \tag{4.22}$$

Hence, the sum rules that we have formulated here are of no real practical use. However, in the case of the actual QCD sum rules, we in principle do not know much about the details of the spectral function and can only approximately evaluate the left hand side of Eq. (4.21) with the help of the operator product expansion (OPE). We thus in reality only have the sum rule to extract information on the spectral function. We will simulate this situation here by replacing the exact result of the left hand side by an expansion in ω^2, which has properties analog to the OPE in real QCD.

4.2.1.1 Borel Sum Rule

As a first example, we will formulate and analyze the sum rule with the traditional Borel weight function, given as

$$W(M, E) = e^{-E/M}. \tag{4.23}$$

Using Eq. (4.19), one can see that the left hand side of Eq. (4.21) then can be rewritten as

$$\begin{aligned} \oint_{C(|E|=R)} G(E)W(M, E)dE &= \oint_{C(|E|=R)} G(E)e^{-E/M}dE \\ &= 2\pi i \sum_n |\psi_n(0)|^2 e^{-E_n/M} \\ &= 2\pi i \langle 0|e^{-H/M}|0\rangle. \end{aligned} \tag{4.24}$$

We can observe from the last line of the above equations, that this expression in fact represents the amplitude (times $2\pi i$) for a transition starting at $\mathbf{x} = 0$ and being back at the origin again after an imaginary time of $-i/M$. The value for this amplitude is known analytically (Feynman and Hibbs 1965) and can be given as,

$$\langle 0|e^{-H/M}|0\rangle = \left[\frac{m\omega}{2\pi\,\sinh(\omega/M)}\right]^{3/2}. \tag{4.25}$$

On the other hand, defining the spectral function as

$$\rho(E) = -\frac{1}{\pi}\mathrm{Im}G(E+i\varepsilon) = \sum_n |\psi_n(0)|^2\delta(E-E_n), \tag{4.26}$$

and assembling the of Eqs. (4.21), (4.24) and (4.25), we obtain the following sum rule,

$$\left[\frac{m\omega}{2\pi\,\sinh(\omega/M)}\right]^{3/2} = \int_0^\infty \rho(E)e^{-E/M}dE. \tag{4.27}$$

Here, both sides are still exact, but we now suppose that we do not know $\rho(E)$ and the left hand side only as an expansion in ω^2/M^2. We then get,

$$\left(\frac{mM}{2\pi}\right)^{3/2}\left[1-\frac{1}{4}\left(\frac{\omega}{M}\right)^2+\frac{19}{480}\left(\frac{\omega}{M}\right)^4-\frac{631}{120960}\left(\frac{\omega}{M}\right)^6+\cdots\right] \simeq \int_0^\infty \rho(E)e^{-E/M}dE, \tag{4.28}$$

which is a sum rule that has been discussed many times in the literature (Pascual and Tarrach 1984; Vainshtein et al. 1980; Whitenton et al. 1983; Lucha et al. 2007).

4.2.1.2 Gaussian Sum Rule

One of course does not have to restrict to sum rules to the weight function of Eq. (4.23), but can also consider alternative choices. One such alternative is the Gaussian kernel, given as

$$W(s,\tau,E) = \frac{1}{\sqrt{4\pi\tau}}e^{-(E-s)^2/(4\tau)}. \tag{4.29}$$

To derive the sum rule for this case, it is not necessary to obtain the exact form of the Green's function, as it was done in Eq. (4.27) for the previous example. Instead, we can directly use the results of Eq. (4.28), derive a spectral function corresponding to each term of the expansion and make use of these spectral functions to calculate the "OPE side" of the sum rule.

Specifically, we first find the relation

$$\left(\frac{mM}{2\pi}\right)^{3/2}\left[1-\frac{1}{4}\left(\frac{\omega}{M}\right)^2+\frac{19}{480}\left(\frac{\omega}{M}\right)^4-\frac{631}{120960}\left(\frac{\omega}{M}\right)^6+\cdots\right]$$
$$=-\left(\frac{m^3}{8\pi^4}\right)^{1/2}\int_{C'}\sqrt{E}\left[1+\frac{1}{6}\left(\frac{\omega}{E}\right)^2-\frac{19}{512}\left(\frac{\omega}{E}\right)^4\right. \tag{4.30}$$
$$\left.+\frac{631}{8195}\left(\frac{\omega}{E}\right)^6-\cdots\right]e^{-E/M}dE,$$

for which the contour C′ is shown in Fig. 4.3. One might worry that there could appear a divergence from the part of the path integral around the origin. Such a divergence indeed occurs from the various parts of C′, but cancels once they are added and therefore causes no problem.

To illustrate this cancellation, we here explicitly derive the second term of Eq. (4.33), where the divergence turns up first. We thus have to consider the following contour integral:

$$\int_{C'}\frac{1}{\sqrt{E^3}}e^{-E/M}dE. \tag{4.31}$$

The contour C′ can be split into two parts, one circling around the origin (C'_1) and the other one running above and below the positive real axis (C'_2). For the first part, we take the radius of the contour to be ε (which is infinitesimally small and should be taken to be 0 at the end of the calculation) and obtain,

$$\int_{C'_1}\frac{1}{\sqrt{E^3}}e^{-E/M}dE=\frac{i}{\varepsilon^{1/2}}\int_0^{2\pi}d\theta e^{-i\theta/2}=\frac{4}{\varepsilon^{1/2}}, \tag{4.32}$$

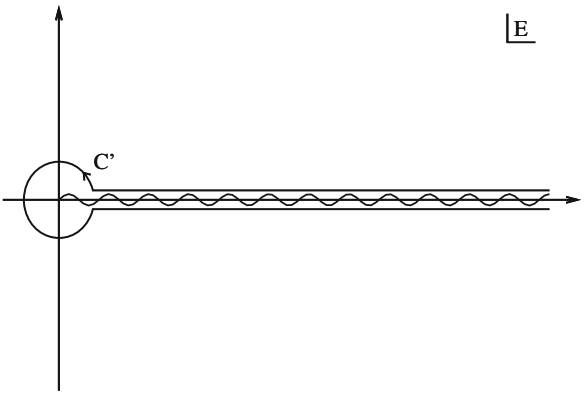

Fig. 4.3 The contour integral C′ on the complex plane of the variable E, used for deriving Eq. (4.33). The *wavy line* denotes the cut of the integrand on the positive side of the real axis

where we have changed the integration variable from E to θ, using $E = \varepsilon e^{i\theta}$. Furthermore, we have dropped all higher orders of ε, which are irrelevant here. Next, the second part gives

$$\begin{aligned}\int_{C_2'} \frac{1}{\sqrt{E^3}} e^{-E/M} dE &= -2\int_{\varepsilon}^{\infty} dE \frac{1}{\sqrt{E^3}} e^{-E/M} \\ &= -\frac{4}{\varepsilon^{1/2}} + \frac{4}{M}\int_{\varepsilon}^{\infty} dE \frac{1}{\sqrt{E}} e^{-E/M} \\ &= -\frac{4}{\varepsilon^{1/2}} + \frac{4\sqrt{\pi}}{M^{1/2}}.\end{aligned} \tag{4.33}$$

Hence, we see that the divergences of both parts indeed cancel and get

$$\int_{C'} \frac{1}{\sqrt{E^3}} e^{-E/M} dE = \frac{4\sqrt{\pi}}{M^{1/2}}, \tag{4.34}$$

as the final result, which was used in Eq. (4.33). The evaluation of the higher order terms are more complicated but the procedure is essentially the same.

Having obtained Eq. (4.30), all we have to do now is to replace $e^{-E/M}$ with the kernel of the Gaussian sum rule, given in Eq. (4.29), and carry out the path integral along C'. After this integration is done, we can write the Gaussian sum rules as follows:

$$\begin{aligned}\left(\frac{2m^3\sqrt{\tau}}{\pi^5}\right)^{1/2}\Bigg[G_0\left(\frac{s}{\sqrt{4\tau}}\right) + \frac{1}{8}\frac{\omega^2}{\tau}G_1\left(\frac{s}{\sqrt{4\tau}}\right) + \frac{19}{640}\frac{\omega^4}{\tau^2}G_2\left(\frac{s}{\sqrt{4\tau}}\right) \\ + \frac{631}{64512}\frac{\omega^6}{\tau^3}G_3\left(\frac{s}{\sqrt{4\tau}}\right) + \cdots\Bigg] \\ \simeq \frac{1}{\sqrt{4\pi\tau}}\int_0^{\infty} \rho(E) e^{-(E-s)^2/(4\tau)} dE.\end{aligned} \tag{4.35}$$

Here, the functions $G_n(x)$ are defined as,

$$\begin{aligned}G_0(x) &= \int_{-x}^{\infty} dy \sqrt{y+x} e^{-y^2}, \\ G_1(x) &= \int_{-x}^{\infty} dy \sqrt{y+x}(1-2y^2) e^{-y^2}, \\ G_2(x) &= \int_{-x}^{\infty} dy \sqrt{y+x}\left(1-4y^2+\frac{4}{3}y^4\right) e^{-y^2}, \\ G_3(x) &= \int_{-x}^{\infty} dy \sqrt{y+x}\left(1-6y^2+4y^4-\frac{8}{15}y^6\right) e^{-y^2}, \\ &\cdots\end{aligned} \tag{4.36}$$

For doing actual calculations, we evaluate the above integrals numerically.

4.2.2 MEM Analysis of the Borel Sum Rules

After having constructed the sum rules, we can now analyze them with the maximum entropy method, according to the procedures described in this chapter. As can be seen for instance in Eq. (4.4), we however first have to determine the error of the "OPE side" of the sum rules as a function of M for obtaining an accurate description of the likelihood function. In the analysis of the real QCD sum rules, we will statistically evaluate this error from the uncertainties of the various condensates that appear in the OPE, as will be explained in Chap. 5. For the present case, we simulate these conditions by assuming that the numerical factors appearing in the expansion of Eq. (4.28) have a certain amount of uncertainty.

First we consider the case, in which all terms (except the leading term, which we assume to be known exactly) have an uncertainty of 1 %. From the point of view of simulating the actual QCD sum rules, this is of course overly optimistic, as most of the condensates have an error of the order of 10 %. A more realistic case will be discussed later. Before giving the results of the MEM analysis, we show the "OPE data" of the left hand side of Eq. (4.28), with the corresponding error in Fig. 4.4. Here and in all what follows, we set $m = \omega$ and give all results in units of ω for simplicity. We clearly see in the figure that the expansion starts to deviate from the exact result around $M/\omega = 0.5$ and explodes below this value. Therefore one should only use the "OPE data" in the regions of M, where the expansion converges. Usually, the lower boundary of the admittable window is determined by the convergence criterion, which demands that the contribution of the highest order term should be less than 10 % of the whole expansion. This criterion gives $M_{\min}/\omega = 0.49$ and we can see from Fig. 4.4 that this choice is reasonable. Furthermore, we can also observe from Fig. 4.4 that the ground state dominates the sum rules for values of M/ω up to around 0.75. One can therefore conclude that there exists a sizable "Borel window" in this particular case, where both the OPE is under control and the ground state dominates the sum rules. As in the present MEM analysis we are also interested in possible excited states, we take the somewhat larger value of $M_{\max}/\omega = 1.25$ for the upper boundary of the used Borel mass region. For the present analysis, we use 100 data points for the OPE ($N_M = 100$) and 800 data points for the spectral function ($N_\omega = 800$), setting $\omega_{\min}$ to 0 and $\omega_{\max}$ to 8 ω.

Let us now turn to the results of the MEM analysis. The obtained spectral function is shown in Fig. 4.5, together with the corresponding error in the region of the lowest peak and the used default model. Note that in this figure, the spectral function is divided by $\sqrt{E/\omega}$ so that its peaks reach a constant strength at high energies. Figure 4.5 clearly shows that the position of the lowest peak is reproduced quite well by the MEM analysis. Numerically, the obtained peak lies at $E/\omega = 1.53$ which is close to the true value of $E/\omega = 1.50$. However, we also observe that the width of the lowest peak is large, opposed to the exact spectral function, where this peak is a

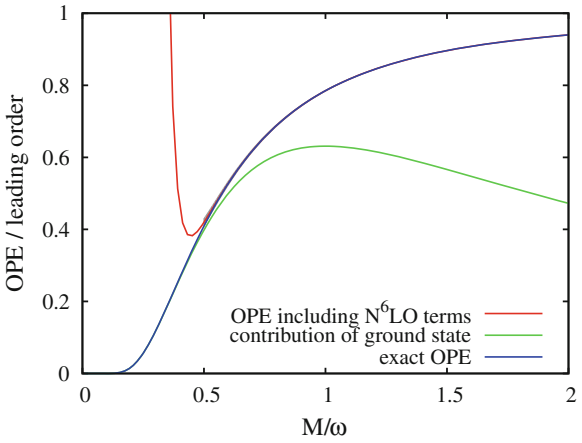

Fig. 4.4 The expansion of Eq. (4.28), normalized by its leading terms is shown by the *red line*. Here, we have included the first 6 terms of the expansion. Our assumed error of this expression is indicated by the *gray band*, which is shown for $0.49 \leq M/\omega \leq 1.25$, the region which is used for the MEM analysis. The *blue line* shows the corresponding exact result of Eq. (4.27). Furthermore, the contribution of the ground state ($E_0 = 3/2\omega$) is shown by the *green line*

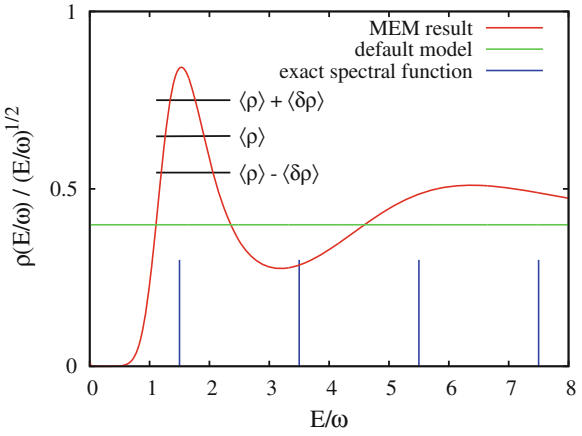

Fig. 4.5 The spectral function obtained from the MEM analysis of Eq. (4.28) with small error (*red line*). The *horizontal bars* give the error of the spectral function, averaged over the region indicated by the horizontal length of the *bars*, as illustrated in Fig. 4.2. For all the following figures of this thesis, we will use the same conventions for the error bars. The employed default model is shown by the *green line*. The *vertical blue lines* indicate the positions of the poles of the true spectral function

delta function. Hence, this width is a pure artifact of the MEM analysis. Furthermore, we see no sign of the excited states in Fig. 4.5 and therefore have to conclude that it is not possible to extract any information on the excited states with the present setting.

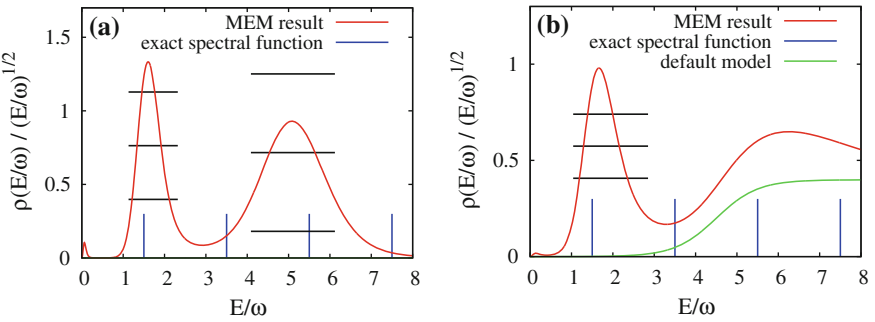

Fig. 4.6 Same as in Fig. 4.5, but with a different default model used for the MEM analysis. **a** A constant default model close to 0 (10^{-3} × high energy limit) is employed (therefore it is not visible in the figure). **b** The default model is a smooth step function, as shown by the *green line*

One might wonder how strong the above results depend on the choice of the default model. To investigate this question, we repeat the same analysis with two other default models, one with a value close to zero (10^{-3} × high energy limit) and another one with a smooth step function like behavior. Both of these default models will play a role in the further course of this thesis. The results of these analyses are shown in Fig. 4.6. One can see in the figures that while the lowest peak is quite stable, the resultant spectral functions significantly differ in the high energy region. Numerically, the lowest peak can be found at $E/\omega = 1.60$ for the default model close to 0, while it is located at $E/\omega = 1.66$ for the smooth step-function like default model. The differences between these results should give an idea of the systematic errors involved with the current MEM analysis. Furthermore, we can see that the high-energy limit of the obtained spectral function is essentially determined by the default model. This is so because the region of the Borel mass M that we use for our analysis strongly depends on the spectral function in the region of the ground state, while it is not very sensitive to the behavior of the spectral function at high energy.

Next, we consider a somewhat more realistic situation. To be more specific, we take the uncertainty of the first condensate appearing in the expansion (which means the second term in Eq. (4.28)) to be 1 %, because it is usually known quite accurately, and for all the remaining condensates to be 10 %. The error for this case is shown by the gray region of Fig. 4.7, which should be compared to the one shown in Fig. 4.4. The result of the corresponding MEM analysis is given in Fig. 4.8, where we have used the same constant default model as in Fig. 4.5. One can see from this result that here even the ground state is almost completely washed out and it therefore seems to be difficult to extract any valuable information from the sum rule by the MEM analysis.

To check if it is possible to improve the reproducibility of the MEM procedure by another choice of the default model, we repeat the above analysis with two other default models, as before. The results are shown in Fig. 4.9. It is observed in these figures, that while the flat default model close to 0 does not give reasonable results,

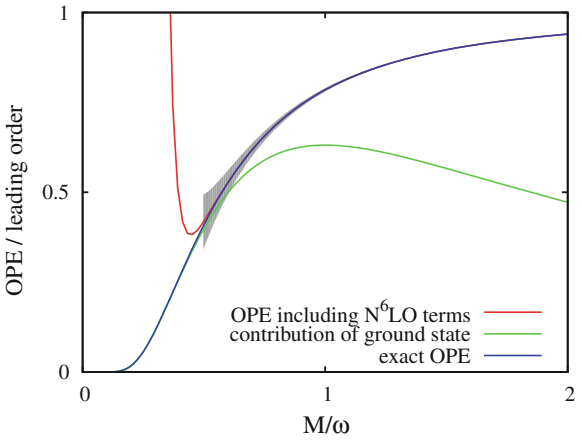

Fig. 4.7 Same as in Fig. 4.4, but with more realistic errors. For details, see the main text

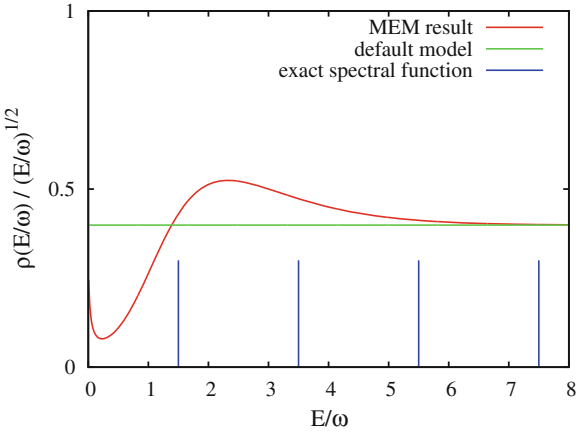

Fig. 4.8 Same as in Fig. 4.5, but with more realistic errors

it is possible to recover the ground state peak at $E/\omega = 1.56$ with the smooth step-function like default model. This finding indicates that, surprisingly, even with uncertainties as large as the ones shown in Fig. 4.7, it is still possible the extract information on the ground state from the sum rule if a suitable default model is chosen. We will encounter a similar case in the analysis of the Borel sum rule of the ρ meson channel. The situation is, however, not entirely satisfactory as the results depend on the default model and the error of the obtained spectral function is large, as can be observed in the right plot of Fig. 4.9. We will therefore consider the Gaussian sum rule in the next section, to examine if it is possible to improve the results obtained so far.

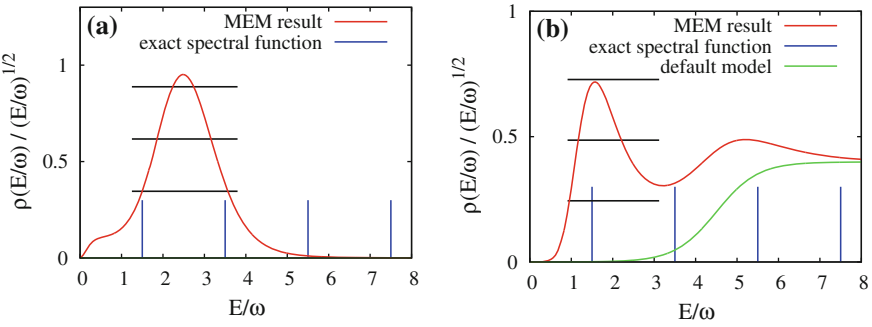

Fig. 4.9 Same as in Fig. 4.6, but with more realistic errors

4.2.3 MEM Analysis of the Gaussian Sum Rules

In this section, we analyze the Gaussian sum rules of the harmonic oscillator potential model, given in Eqs. (4.35) and (4.36). As in the case of the Borel sum rule, we first have to determine the region of the parameters s and τ used for the MEM analysis. Because we now have two parameters instead of one, this task becomes somewhat more complicated. Here, we simply choose four different values of $\tau_i (i = 1 \sim 4)$, and determine the boundaries of s for each value of τ_i separately. The lower boundary of s is as usual fixed by the 10 % convergence criterion of the OPE, while we set the upper boundary to $3.0\,\omega$ for all τ_i. The detailed values of the employed regions are shown below.

$$\tau_1 = 0.75\,\omega^2, \quad s_{1,\mathrm{min}} = -3.39\,\omega, \quad s_{1,\mathrm{max}} = 3.0\,\omega, \tag{4.37}$$

$$\tau_2 = 1.0\,\omega^2, \quad s_{2,\mathrm{min}} = -4.57\,\omega, \quad s_{2,\mathrm{max}} = 3.0\,\omega, \tag{4.38}$$

$$\tau_3 = 1.25\,\omega^2, \quad s_{3,\mathrm{min}} = -6.33\,\omega, \quad s_{3,\mathrm{max}} = 3.0\,\omega, \tag{4.39}$$

$$\tau_4 = 1.5\,\omega^2, \quad s_{4,\mathrm{min}} = -7.41\,\omega, \quad s_{4,\mathrm{max}} = 3.0\,\omega. \tag{4.40}$$

Some comments are in order here. Firstly, let us discuss the method of using the two variables of s and τ at the same time, compared to Sect. 4.1, where only one variable x was considered. Generalizing the method from one variable to two is straightforward, as one only needs to redefine the kernel and the likelihood function. Specifically, we use

$$G(x_i) = \sum_{\omega_j=\omega_{\min}}^{\omega_{\max}} K(x_i, \omega_j)\rho(\omega_j)d\omega$$

$$\downarrow$$

$$G(s_i, \tau_k) = \sum_{\omega_j=\omega_{\min}}^{\omega_{\max}} K(s_i, \tau_k, \omega_j)\rho(\omega_j)d\omega,$$

and

$$L(\rho) = \frac{1}{2(x_{\max} - x_{\min})} \sum_{x_i=x_{\min}}^{x_{\max}} dx \frac{[G_{\mathrm{OPE}}(x_i) - G_\rho(x_i)]^2}{\sigma^2(x_i)}$$

$$\downarrow$$

$$L(\rho) = \sum_{\tau_k=\tau_{\min}}^{\tau_{\max}} \sum_{s_i=s_{k,\min}}^{s_{k,\max}} ds_k \frac{1}{2(s_{k,\max} - s_{k,\min})} \frac{[G_{\mathrm{OPE}}(s_i, \tau_k) - G_\rho(s_i, \tau_k)]^2}{\sigma^2(s_i, \tau_k)}.$$

In the present analysis, we take 4 data points for τ ($N_\tau = 4$), and 50 data points for s ($N_s = 50$) at each τ_i and 800 data points for ω ($N_\omega = 800$). Furthermore, we fix $\omega_{\min}$ and $\omega_{\max}$ to 0 and 8 ω, respectively.

As a second point, it may seem somewhat counterintuitive that we use negative values for the parameter s here, as one would naively expect, that the region where s takes values close to the lowest peak position are most sensitive to its properties. As we will however see below in Fig. 4.10, the OPE at positive s values contain considerable contributions from the excited states and it is only for negative s values that the ground state dominates the sum rules. For a related discussion in the context of the nucleon sum rules, see Sect. 6.4.1 of this thesis.

The results of the OPE, together with the exact result summed up to all orders is shown in Fig. 4.10. In the same figure, the contribution of the ground state is also indicated. We clearly see that, in contrast to the Borel sum rule result given in Fig. 4.4, there is a quite large window, in which the ground state completely dominates the sum rules and where the OPE is convergent. This observation lets one hope that the spectral function can be extracted more accurately than that for the Borel sum rule. As we will see below, this is indeed the case.

The errors shown by the gray bands in Fig. 4.10 are, as before, obtained by assuming that all "condensates" have an uncertainty of 1 %. We now show the extracted spectral functions for this case in Fig. 4.11, which should be compared to the earlier result of Fig. 4.5 of the Borel sum rule. It is obvious from the two figures, that the Gaussian sum rule considerably improves the extraction of the spectral function. We here not only get a sharp peak for the ground state, but also can determine the position of the first excited state with high precision. Quantitatively, the extracted ground state lies at $E/\omega = 1.49$ and the first excited state at $E/\omega = 3.51$, thus both are being obtained with an error of below 1 %. Surprisingly, there even appears a third peak which lies in the region of the second excited state. It is found at $E/\omega = 5.74$, which

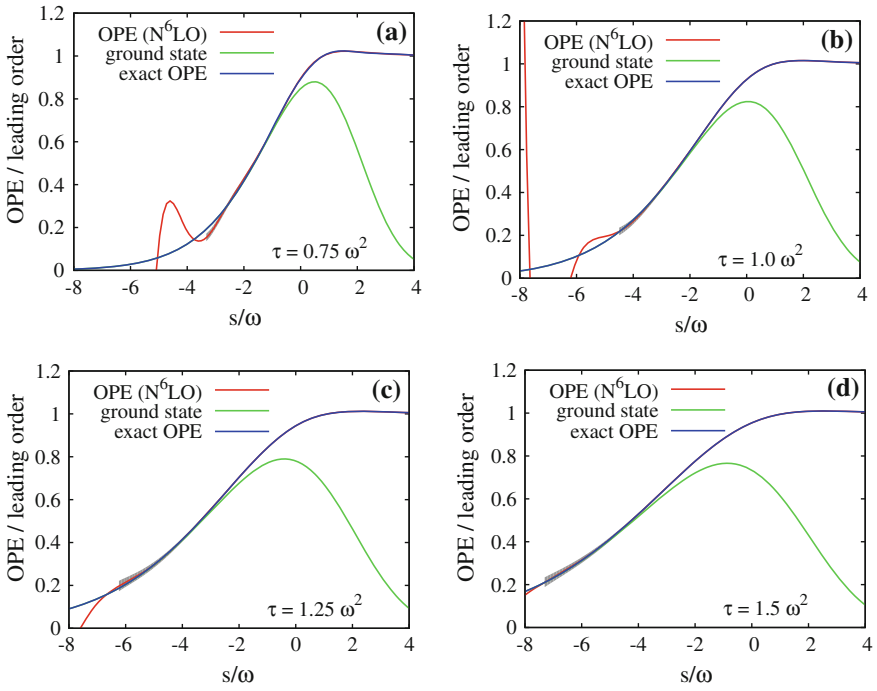

Fig. 4.10 The expansion of Eq. (4.35), normalized by its leading terms is shown by the *red lines* for four values of τ. Here, we have included the first 6 terms of the expansion. Our assumed error of this expression is indicated by the *gray bands*, which are shown for the regions of s which are used for the MEM analysis. The *blue lines* give the corresponding exact result. Furthermore, the contribution of the ground state ($E_0 = 3/2\omega$) is shown by the *green line*

is somewhat larger than the correct value, but still within a precision of 5 %. We have also repeated the same analysis with different default model as in Fig. 4.6 for the Borel sum rules. The results almost completely coincide, and we therefore do not show them here. In all, we can conclude that the MEM analysis of the Gaussian sum rule works very well, shows no dependence on the default model and even makes it possible to obtain reliable information on the excited states of this simple potential model.

The results obtained above are promising, the used errors being, however, unrealistically small. We therefore, repeat the same analysis with an error closer to reality. As for the Borel sum rule, we take the error of the first condensate to be 1 % and the rest 10 %. The uncertainty of the OPE that we obtain in this way is shown by the gray regions of Fig. 4.12. Using these errors, we extract the spectral function given in Fig. 4.13. The figure shows that due to the increased error, the peaks develop an artificial width, which implies a reduced resolution of the MEM. However, we still get clear peaks of the ground state and first excited state at positions $E/\omega = 1.46$

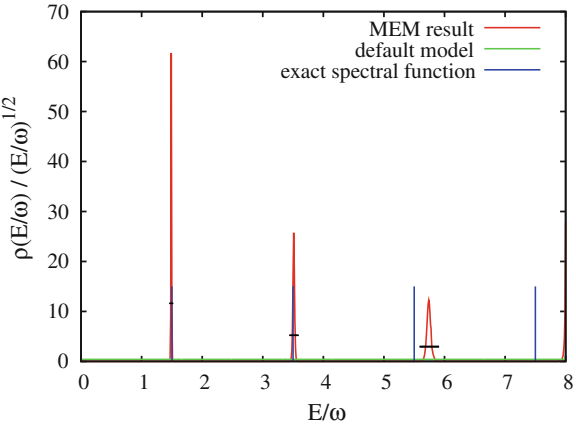

Fig. 4.11 The spectral function obtained from the MEM analysis of Eq. (4.35) with small error (*red line*). The *horizontal bars* give the error of the spectral function, averaged over the region indicated by the horizontal length of the *bars*. The employed default model is shown by the *green line*. Finally, the exact positions of the input spectral function peaks are indicated by the *blue vertical lines*

and $E/\omega = 3.54$, which means that their deviation from the true value is still below 3 %. On the other hand, the third peak is now considerably washed out and its position is moved upwards and presumably overlaps with the next higher peak, so that it becomes difficult to extract useful information on the second excited state.

To see whether these results are stable against changes of the default model, we give the results of the MEM analysis using different default models in Fig. 4.14. We can observe in these figures that the existence of the first two peaks does not depend on the details of the default model. We also see that, while their width and height can change, their position is stable. Quantitatively, they are found at $E/\omega = 1.49$ and $E/\omega = 3.51$ for the default model close to 0, while their position is at $E/\omega = 1.48$ and $E/\omega = 3.48$ for the smooth step-function like default model. This indicates that the width of the peaks is a pure MEM artifact and should not be considered to be property of the real spectral function, while the positions of the (lower) peaks, indeed do reflect the properties of the true spectral function.

4.2.3.1 Up to Which Order Do We Need to Calculate the OPE?

For this simple model, we can calculate the OPE accurately up to any desired order. We can therefore here ask the question up to which order we really need to calculate the OPE for obtaining reliable information on the first few lowest peaks in the spectral function. To investigate this issue, we carry out the MEM analysis for the Gaussian sum rules with various orders of the OPE taken into account. In this calculation, we do not only incorporate the change of the OPE at different orders, but also the shifts of the allowed window, in which the OPE is converging and furthermore the variation

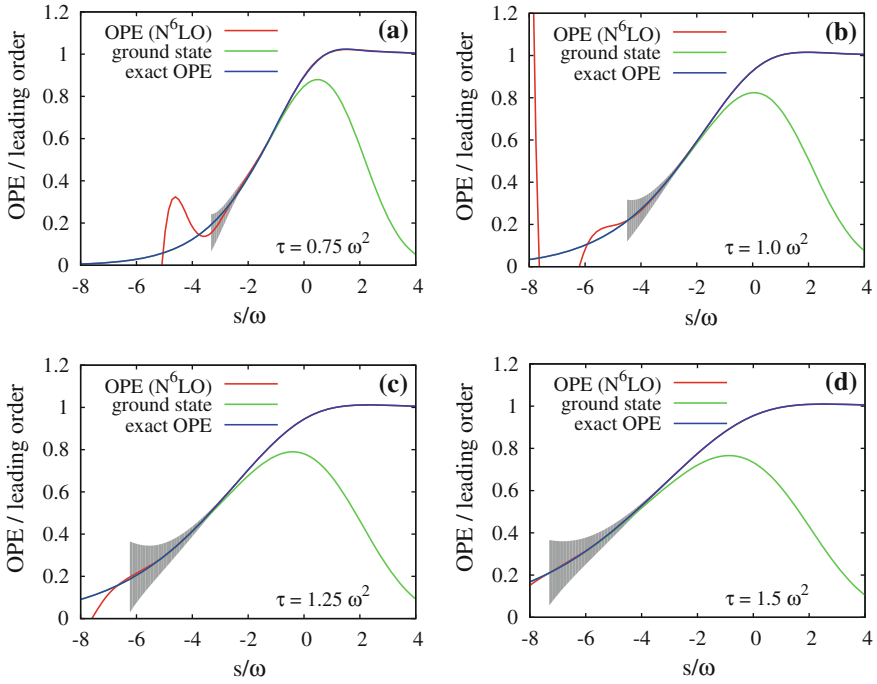

Fig. 4.12 Same as in Fig. 4.10, but with more realistic errors. For details, see the main text

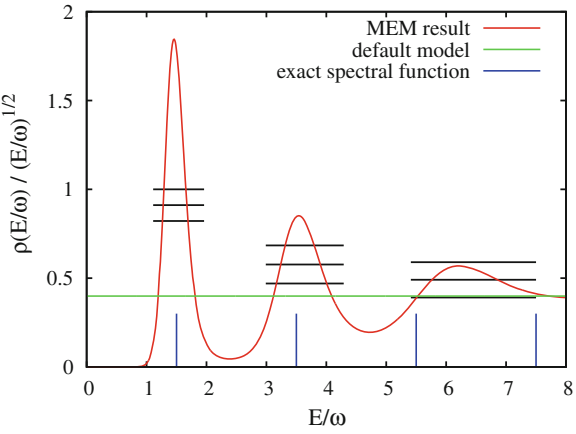

Fig. 4.13 Same as in Fig. 4.11, but with more realistic errors

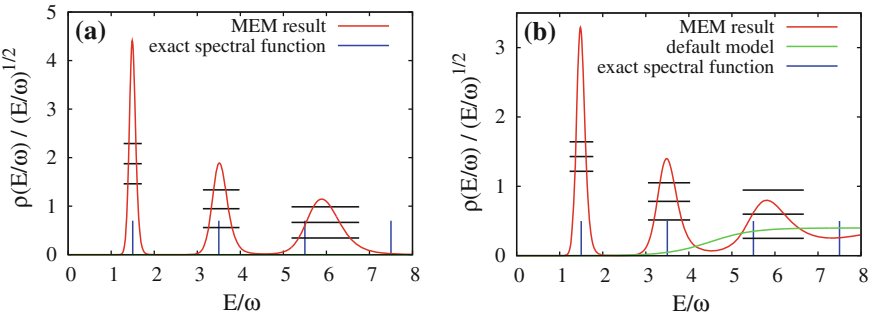

Fig. 4.14 Same as in Fig. 4.13, but with a different default model used for the MEM analysis. **a** A constant default model close to 0 (10^{-3} × high energy limit) is employed (therefore it is not visible in the figure). **b** The default model is a smooth step function, as shown by the *green line*

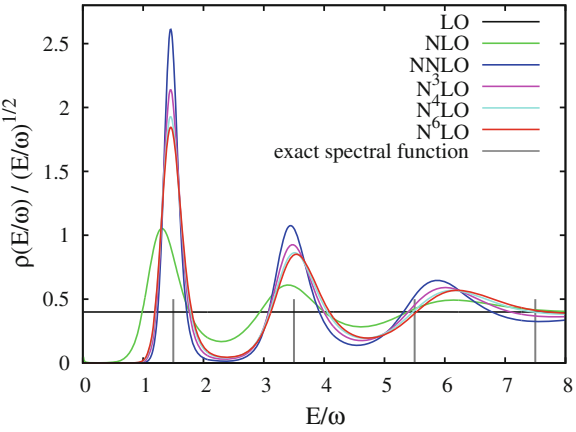

Fig. 4.15 The results of the MEM analysis of the Gaussian sum rules, when only a limited order of the OPE is taken into account. For this figure, the same default model as in Fig. 4.13 was employed

of the error introduced by the higher orders of the OPE. The results are shown in Fig. 4.15.

Examining this plot in detail, we see that for obtaining the ground state at the approximately correct position, we in fact only need the first two OPE terms besides the leading order. This is good news for the actual QCD sum rule studies because for QCD usually only the first two or three OPE terms are known accurately, and further terms suffer from rather large uncertainties due to our lack of knowledge of the values of the higher dimensional condensates. Moreover, we observe that with the first two OPE terms, we also get a peak corresponding to the first excited state, however with a somewhat reduced precision compared to the ground state. This precision is improved if two more terms are taken into account.

Here, we also comment on the reduction (increase) of the height (width) of the ground state peak, which occurs if more than two terms of the OPE are incorporated into the analysis. This happens simply because, by including the higher order terms, we automatically increase the error of the OPE due to the uncertainties of the condensates. A larger error leads to a smaller resolution for the MEM procedure and thus to a smaller and broader peak.

4.2.4 Summary of Toy Model Analysis

Let us briefly summarize what we have learned from the potential model analysis described in the preceding sections.

- Both Borel and Gaussian sum rules contain valuable information on the existence and position of the lowest lying peak of the spectrum, which can in principle be extracted from an MEM analysis. However, for the Borel sum rule with realistic errors, this is only possible for suitable choice of the default model.
- The Gaussian sum rule clearly works better that the Borel sum rule, mainly because it provides a larger parameter window, in which the ground state dominates the sum rule and the OPE converges, as can be seen in Fig. 4.10. In contrast to the Borel sum rule, it makes it even possible to obtain the position of the first excited state with high precision.
- For extracting the position of the ground state, it is sufficient to calculate the first two terms of the OPE, while for the first excited state, the first four terms are necessary if one wants to achieve high precision.

The above observations provide us with a useful conceptual guide for the analysis of the actual QCD sum rules, that will be carried out in the following chapters. One, however, has to keep in mind that the potential model discussed here does not necessarily has exactly the same features as QCD and one thus cannot be sure that the above findings are also true for the real QCD sum rules. For instance, the model considered here does not have a continuum, which usually appears somewhere above (or below) the lowest peak in the various QCD spectra. This continuum will in reality make the extraction of the excited states much more difficult than for the simple case considered here.

References

Asakawa M, Hatsuda T, Nakahara Y (2001) Prog Part Nucl Phys 46:459
Bryan RK (1990) Eur Biophys J 18:165
Feynman R, Hibbs A (1965) Quantum mechanics and path integrals. McGraw-Hill, New York
Giménez V, Bordes J, Peñarrocha J (1991) Nucl Phys B357:3
Jarrel M, Gubernatis JE (1996) Phys Rep 269:133
Lucha W, Melikhov D, Simula S (2007) Phys Lett B 657:148

Nakahara Y, Asakawa M, Hatsuda T (1999) Phys Rev D 60:091503(R)
Nickel D (2007) Ann Phys 322:1949
Pascual P, Tarrach R (1984) QCD: renormalization for the practitioner. Springer, Berlin
Vainshtein AI, Zakharov VI, Novikov VA, Shifman MA (1980) Sov J Nucl Phys 32:840
Whitenton JB, Durand B, Durand L (1983) Phys Rev D 28:597

Part II
Applications

Chapter 5
MEM Analysis of the ρ Meson Sum Rule

5.1 Introduction

The technique of QCD sum rules is well known for its ability to reproduce various properties of hadrons (Shifman et al. 1979; Reinders et al. 1985). Using dispersion relations, this method connects perturbative and non-perturbative sectors of QCD, and therefore, allows one to describe inherently non-perturbative objects such as hadrons by the operator product expansion (OPE), which is essentially a perturbative procedure. The higher-order terms of the OPE contain condensates of various operators, which incorporate information on the QCD vacuum. Hence, QCD sum rules also provide us with nontrivial relations between the properties of hadrons and the QCD vacuum.

Since the early days of the development of QCD sum rules, the range of applications of this method has been constantly expanding, which has helped to explain many aspects of the behavior of hadrons. Nevertheless, QCD sum rules have always been subject to (justified) criticism. One part of this criticism is of mainly technical nature, pointing out that the analysis of QCD sum rules often is not done with the necessary rigor, namely, that the OPE convergence and/or the pole dominance condition are not properly taken into account. Many of the recent works that followed the claimed discovery of the pentaquark $\Theta^+(1540)$ are examples of such a lack of rigor. Nonetheless, these technical problems can be overcome if the analysis is done carefully enough (Gubler et al. 2009a, b).

The second part of the criticism against QCD sum rules is more essential. It is concerned with the ansatz taken to parametrize the spectral function. For instance, it is common to assume the "pole + continuum" functional form, where the pole represents the hadron in question and the continuum stands for the excited and scattering states. While this ansatz may be justified in cases where the low-energy part of the spectral function is dominated by a single pole and the continuum states become important only at higher energies (the ρ-meson channel for instance is such a case), it is not at all clear if it is also valid in other cases. For example, as shown in Kojo and Jido (2008), where the σ-meson channel was investigated using QCD

P. Gubler, *A Bayesian Analysis of QCD Sum Rules*, Springer Theses,
DOI: 10.1007/978-4-431-54318-3_5, © Springer Japan 2013

sum rules, it can be difficult to distinguish the continuum spectrum from a broad resonance, because they lead to a similar behavior of the "pole" mass and residue as functions of the Borel mass. Moreover, using the "pole + continuum" ansatz, the outcome of the analysis usually depends on unphysical parameters such as the Borel mass or the threshold parameter, and it is not always a trivial matter to determine these parameters in a consistent way. After all, the ansatz of parametrizing the spectral function makes a full error estimation impossible in QCD sum rules.

As a possible solution to these problems, we proposed in Gubler and Oka (2010) to analyze QCD sum rules with the help of the Maximum Entropy Method (MEM). This method has already been applied to Monte-Carlo studies in both condensed matter physics (Jarrel and Gubernatis 1996) and lattice QCD (Nakahara et al. 1999; Asakawa et al. 2001), and has applications in many other areas (Wu 1997). It makes use of Bayes' theorem of probability theory, which helps to incorporate known properties of the spectral function such as positivity and asymptotic values into the analysis and finally makes it possible to obtain the most probable spectral function without having to introduce any additional a priori assumptions about its explicit form. It even allows us to estimate the error of the obtained spectral function. Therefore, using this approach, it should in principle be possible to study the spectral function of any channel, including those for which the "pole + continuum" assumption is not appropriate.

However, as a first step it is indispensable to check whether QCD sum rules are a suitable target for MEM and if it is possible to obtain any meaningful information on the spectral function by this method. To provide an answer to these questions is the main object of this chapter. To carry out this check, we have chosen to investigate the sum rule of the ρ-meson. This channel is one of the first subjects that have been studied in QCD sum rules and it is fair to say that it is the channel where this method has so far seen its most impressive success. As mentioned earlier, it is a case where the "pole + continuum" ansatz works well and we thus do not expect to gain anything really new from this analysis. Nevertheless, apart from the aspect of testing the applicability of our new approach, we believe that it is worth examining this channel once more, as MEM also provides a new viewpoint of looking at various aspects of this particular sum rule.

5.2 Analysis Using Mock Data

The uncertainties that are involved in QCD sum rule calculations mainly originate from the ambiguities of the condensates and other parameters such as the strong coupling constant or the quark masses. These uncertainties usually lead to results with relative errors of about 20 %. It is therefore not a trivial question if MEM can be used to analyze the QCD sum rule results, or if the involved uncertainties are too large to allow a sufficiently accurate application of the MEM procedure.

To investigate this question in detail, we carry out the MEM analysis using mock data and realistic errors. Furthermore, we will study the dependence of the results

on various choices of the default model $m(\omega)$ and determine which one is the most suitable for our purposes. This analysis will also provide us with an estimate of the precision of the final results that can be achieved by this method and what kind of general structures of the spectral function can or cannot be reproduced by the MEM procedure.

5.2.1 Generating Mock Data and the Corresponding Errors

Following Asakawa et al. (2001) and Shuryak (1993), we employ a relativistic Breit-Wigner peak and a smooth function describing the transition to the asymptotic value at high energies for our model spectral function of the ρ-meson channel:

$$\rho_{\text{in}}(\omega) = \frac{2F_\rho^2}{\pi} \frac{\Gamma_\rho m_\rho}{(\omega^2 - m_\rho^2)^2 + \Gamma_\rho^2 m_\rho^2} + \frac{1}{4\pi^2}\left(1 + \frac{\alpha_s}{\pi}\right) \frac{1}{1 + e^{(\omega_0 - \omega)/\delta}}, \tag{5.1}$$

$$\Gamma_\rho(\omega) = \frac{g_{\rho\pi\pi}^2}{48\pi} m_\rho \left(1 - \frac{4m_\pi^2}{\omega^2}\right)^{3/2} \theta(\omega - 2m_\pi).$$

The values used for the various parameters are

$$\begin{aligned} m_\rho &= 0.77 \text{ GeV}, \qquad m_\pi = 0.14 \text{ GeV} \\ \omega_0 &= 1.3 \text{ GeV}, \qquad \delta = 0.2 \text{ GeV}, \\ g_{\rho\pi\pi} &= 5.45, \qquad \alpha_s = 0.5, \\ F_\rho &= \frac{m_\rho}{g_{\rho\pi\pi}} = 0.141 \text{ GeV}. \end{aligned} \tag{5.2}$$

The spectral function of Eq. (5.1) is then substituted into Eq. (4.1) and the integration over ω is performed numerically to obtain the central values of the data points of $G_{\text{mock}}(M)$. (In this section, we will use $G_{\text{mock}}(M)$ instead of $G_{OPE}(M)$ to make it clear that we are analyzing mock data.)

We now also have to put some errors $\sigma(M)$ to the function $G_{\text{mock}}(M)$. To make the analysis as realistic as possible, we will use exactly the same errors as in the actual investigation of the OPE results. How these errors are obtained will be discussed later, in the section where the real OPE results are analyzed. We just mention here that when analyzing the OPE results, we will use three different parametrizations for the condensates and other parameters, namely, those given in Colangelo and Khodjamirian (2001), Narison (2004), Ioffe (2006) (see Table 5.1). These parametrizations lead to different estimations of the errors, but for the mock data analysis of this section, these differences are not very important. Here, we will therefore mainly use the errors obtained from the parameters of Ioffe (2006). The resulting function $G_{\text{mock}}(M)$ is given in Fig. 5.1, together with the range $G_{\text{mock}}(M) \pm \sigma(M)$.

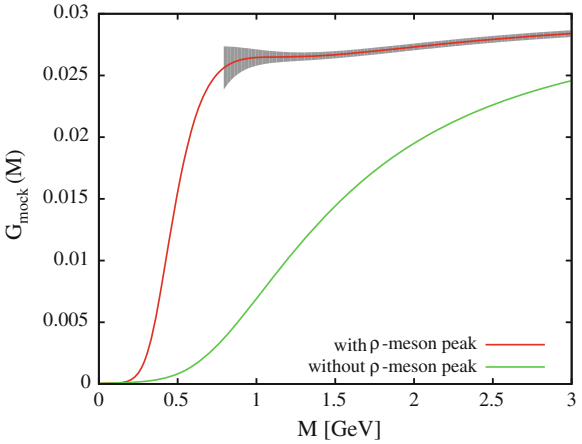

Fig. 5.1 Central values of the mock data (*red line*) obtained by numerically integrating Eq. (4.1) with Eq. (5.1). The errors of the data, extracted from the parameters of Ioffe (2006), are indicated by the *gray region*. The lower boundary of the shown errors corresponds to M_{min}, below which the OPE does not converge. For comparison, the integral of Eq. (4.1) is also shown for the case when only the continuum part, the second term of Eq. (5.1), is taken for the spectral function (*green line*)

5.2.1.1 Determination of the Analyzed Borel Mass Region

Next, we have to decide what range of M to use for the analysis. For the lower boundary, we can employ the usual convergence criterion of the OPE such that the contribution of the highest-dimensional operators is less than 10 % of the sum of all OPE terms. This is a reasonable choice, as the errors originating from the ranges of condensate values lead to uncertainties of up to 20 %, and it would therefore not make much sense to set up a more strict convergence criterion. For the parameters of Ioffe (2006), this gives $M_{\text{min}} = 0.77$ GeV.

Considering the upper boundary of M, the situation is less clear. In the conventional QCD sum rule analysis, it is standard to use the pole dominance condition, which makes sure that the contribution of the continuum states does not become too large. As we do not resort to the "pole + continuum" ansatz in the current approach, the pole dominance criterion does not have to be used and one can, in principle, choose any value for the upper boundary of M. Nevertheless, because we are mainly interested in the lowest resonance peak, we will use a similar pole dominance criterion as in the traditional QCD sum rules. By examining the mock data in Fig. 5.1, one sees that while the resonance pole contributes most strongly to the data around $M \sim 1$ GeV, the contributions from the continuum states grow with increasing M and finally start to dominate the data for values that are larger than 1.5 GeV. We will therefore use $M_{\text{max}} = 1.5$ GeV as the upper boundary of M for the rest of this chapter. The dependence of the final results on this choice is small, as will be shown later in Fig. 5.4.

Finally, the values of the N_M data points of $G_{\mathrm{mock}}(M)$ between M_{min} and M_{max} are randomly generated, using Gaussian distributions with standard deviations $\sigma(M)$, centered at the values obtained from the integration of Eq. (4.1). The ranges of values of $G_{\mathrm{mock}}(M)$ are indicated by the gray region in Fig. 5.1. We take 100 data points for functions of the Borel mass M ($N_M = 100$) and have checked that the results of the analysis do not change when this value is altered. For functions of the energy ω, 600 data points are taken ($N_\omega = 600$).

5.2.2 Choice of an Appropriate Default Model

It is important to understand the meaning of the default model $m(\omega)$ in the present calculation. It is in fact used to fix the value of the spectral function at high and low energies, because the function $G_{OPE}(M)$ contains only little information on these regions. This can be understood firstly by considering the property of the kernel

$$K(M, \omega) = \frac{2\omega}{M^2} e^{-\omega^2/M^2}, \tag{5.3}$$

which is zero at $\omega = 0$. $G_{OPE}(M)$ therefore contains no information on $\rho_{\mathrm{in}}(\omega = 0)$ and the corresponding result of the MEM analysis $\rho_{\mathrm{out}}(\omega = 0)$ will thus always approach the default model, as $G_{OPE}(M)$ does not constrain its value. Secondly, we use $G_{OPE}(M)$ only in a limited range of M, because the operator product expansion diverges for small M and the region of very high $G_{OPE}(M)$ is dominated by the high-energy continuum states, which we are not interested in. The region of the spectral function, which contributes most strongly to $G_{OPE}(M)$ between M_{min} and M_{max}, lies roughly in the range between ω_{min} ($\simeq M_{\mathrm{min}}$) and ω_{max} ($\simeq M_{\mathrm{max}}$), as can be for instance inferred from Fig. 5.1. $\rho_{\mathrm{out}}(\omega)$ will then approach the default model quite quickly outside of this region, because there is no strong constraint from $G_{OPE}(M)$. This implies that the values of $\rho_{\mathrm{out}}(\omega)$ at the boundaries are fixed by the choice of the default model and one should therefore consider these boundary conditions as inputs of the present analysis. Once these limiting values of $\rho(\omega)$ are chosen, the MEM procedure then extracts the most probable spectral function $\rho_{\mathrm{out}}(\omega)$ given $G_{OPE}(M)$ and the boundary conditions supplied by $m(\omega)$.

To illustrate the importance of choosing appropriate boundary conditions, we show the results of the MEM analysis for a constant default model, with a value fixed to the perturbative result at high energy. Here, the boundary condition for the low energy is not correctly chosen, because the spectral function is expected to vanish at very low energy. The result is given in Fig. 5.2 and clearly shows that $\rho_{\mathrm{out}}(\omega)$ does not reproduce $\rho_{\mathrm{in}}(\omega)$.

This is in contrast to the corresponding behavior in lattice QCD, where it suffices to take a constant value of the spectral function, chosen to be consistent with the high-energy behavior of the spectral function to obtain correct results. The reason for this difference is mainly that the OPE in QCD sum rules is not sensitive to the

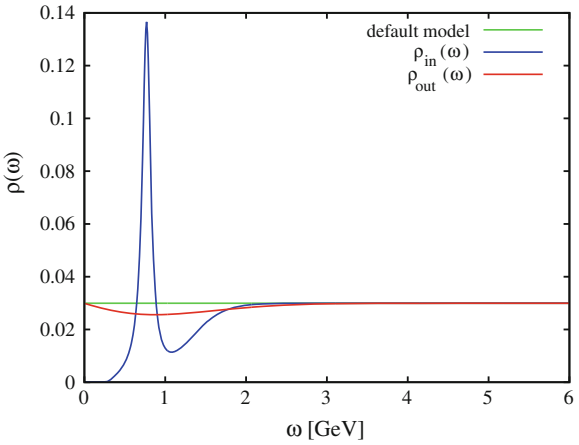

Fig. 5.2 Outcome of the MEM analysis using a constant default model with its value fixed to the perturbative result (*green line*). $\rho_{\text{in}}(\omega)$ is the function that was used to produce the mock data (*blue line*), given in Eq. (5.1), and $\rho_{\text{out}}(\omega)$ shows the result of the MEM procedure (*red line*)

low-energy part of the spectral function, owing to the properties of the kernel and our limitations in the applicability of the OPE. Most importantly, the information that there is essentially no strength in the spectral function below the rho-meson peak, is stored in the region of the Borel mass M around and below $M = 0.5$ GeV. However, as the OPE does not converge in this region, it is not available for our analysis and we therefore have to use the default model to adjust the spectral function to the correct behavior. On the other hand, in lattice QCD, it is possible to calculate the correlator reliably at sufficiently high euclidean time, where the effective mass plot reaches a plateau and thus only the ground state contributes. Therefore, it seems that from lattice QCD one can gain sufficient information on the low-energy part of the spectral function, and one does not need to adjust the default model to obtain physically reasonable results.

For the present analysis, we introduce the following functional form, to smoothly connect low- and high-energy parts of the default model,

$$m(\omega) = \frac{1}{4\pi^2}\left(1 + \frac{\alpha_s}{\pi}\right)\frac{1}{1 + e^{(\omega_0 - \omega)/\delta}}, \tag{5.4}$$

which is close to 0 at low energy and approaches the perturbative value $1 + \frac{\alpha_s}{\pi}$ at high energy, changing most significantly in the region between $\omega_0 - \delta$ and $\omega_0 + \delta$. This function can be considered to be the counterpart of the "continuum" in the "pole + continuum" assumption of

$$\text{Im}\Pi(s) = \pi|\lambda|^2\delta(s - m^2) + \theta(s - s_{th})\,\text{Im}\Pi^{OPE}(s), \tag{5.5}$$

where δ is essentially taken to be 0 and the threshold parameter s_{th} corresponds to ω_0. Equation (5.4) nevertheless enters into the calculation in a very different way than the second term of Eq. (5.5) in the conventional sum rules, and one should therefore not regard these two approaches to the continuum as completely equivalent.

We have tested the MEM analysis of the mock data for several values of ω_0 and δ and the results are shown in Fig. 5.3. One sees that in the cases of **d** and **e**, the sharply rising default model induces an artificial peak in the region of ω_0. Even though these peaks are statistically not significant, they may lead to erroneous conclusions. We will therefore adopt only default models, for which only small artificial structures appear, such as in the case of **a–c**. Comparing these three figures, it is observed that the error of the spectral function relative to the height of the ρ-meson peak is smallest for the parameters of **c**. We therefore adopt this default model with $\omega_0 = 2.0$ and $\delta = 0.1\,\mathrm{GeV}$ in the following investigations.

It is worth considering the results of Fig. 5.3 also from the viewpoint of the dependence of the peak position on the default model. It is observed that even though the height of the ρ-meson peak varies quite strongly depending on which default model is chosen, its position only varies in a range of ± 40 MeV, which shows that the present

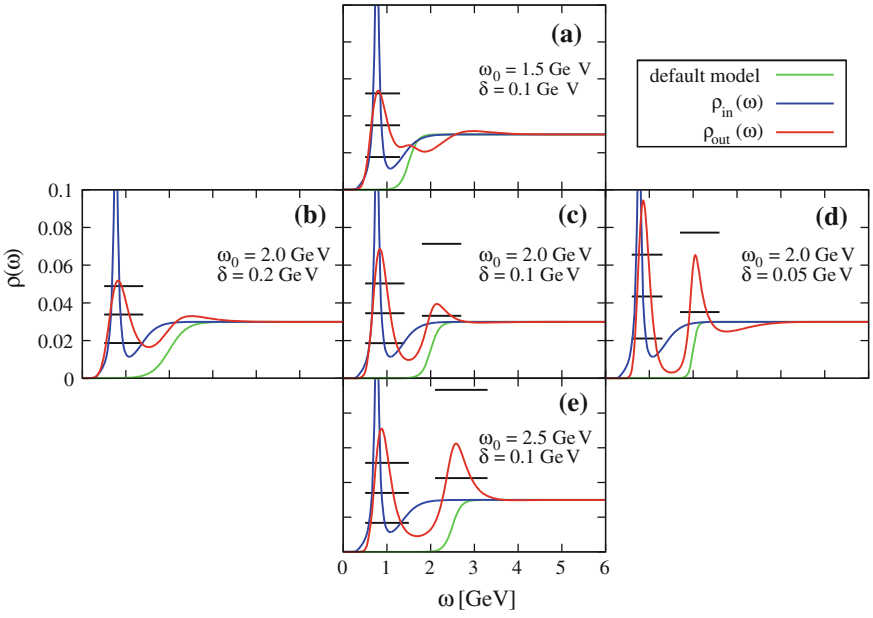

Fig. 5.3 Results of the MEM investigation of mock data with various default models. As in Fig. 5.2, the *red lines* stand for the output of the analysis $\rho_{\mathrm{out}}(\omega)$, the *green lines* for the default model with the parameters shown in the figure, and the *blue lines* for the input function $\rho_{\mathrm{in}}(\omega)$ of Eq. (5.1). The *horizontal bars* show the values of the spectral function, averaged over the peaks $\langle\rho_{\mathrm{out}}\rangle_{\omega_1,\omega_2}$ and the corresponding ranges $\langle\rho_{\mathrm{out}}\rangle_{\omega_1,\omega_2} \pm \langle\delta\rho_{\mathrm{out}}\rangle_{\omega_1,\omega_2}$. Their extent shows the averaged interval (ω_1, ω_2). For figures **c**, **d** and **e**, the lower error bars of the second peak are not shown because they lie below $\rho(\omega) = 0$

evaluation of the lowest pole position does not strongly depend on the detailed values of ω_0 and δ. This behavior should be compared with the results of the usual sum rules, where the dependence of the pole mass on the threshold parameter s_{th} is stronger. On the other hand, we have to mention that ω_0 should not be chosen to have a value much below $\omega_0 = 1.5$ GeV, because, in this region, the artificial structures discussed above start to interfere with the ρ-meson peak. Moreover, if the default model approaches the limit shown in Fig. 5.2, where $m(\omega)$ is just a constant fixed to the asymptotic value at high energy, the ρ-meson peak will gradually disappear.

5.2.3 *Investigation of the Stability of the Obtained Spectral Function*

Next, we will briefly discuss the dependence of our results on the upper boundary of the employed Borel mass region M_{max}. In Fig. 5.4, we show the results for the values $M_{\mathrm{max}} = 1.5$, 2.0 and 2.5 GeV. Here, the default model with parameters $\omega_0 = 2.0$ and $\delta = 0.1$ GeV was used. The spectral functions of these three cases almost coincide and have the same qualitative features. Quantitatively, the peak position of the ρ-meson is shifted only 20 MeV when M_{max} is raised from 1.5 to 2.5 GeV.

Let us also check how the results of the analysis are affected by a different choice of parameters, leading to altered magnitudes of error and also differing lower boundaries of the Borel mass M_{min}. Of the three parameter sets used in this study, given in Table 5.1, the errors of Colangelo and Khodjamirian (2001) are rather small, while the errors of Narison (2004) and Ioffe (2006) are larger and have about the same

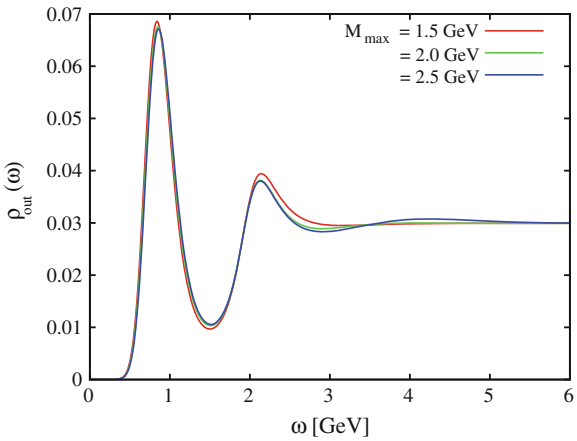

Fig. 5.4 MEM results for three different values of the upper boundary of the Borel mass M_{max}. For the default model, the same version as in Fig. 5.3, **c** was used. Thus, the *red line* of this figure is the same as the one of Fig. 5.3, **c**

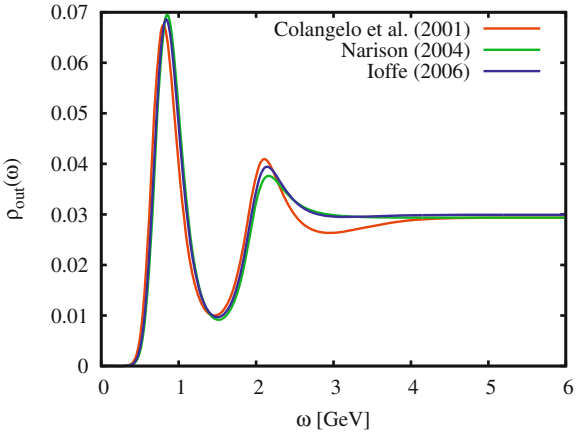

Fig. 5.5 Findings of the MEM analysis with different error estimations and lower boundaries of the Borel mass $M_{\min}$. For the default model, the same version as in Fig. 5.3**c** was used. The *red line* uses parameters of Colangelo and Khodjamirian (2001), the *green line* those of Narison (2004) and the *blue line* those of Ioffe (2006). Note that we use for this plot the same mock data for all three cases and vary only the errors and $M_{\min}$

overall magnitude. Moreover, $M_{\min}$, which is determined from the OPE convergence criterion mentioned earlier, takes values $M_{\min} = 0.72$ GeV for Colangelo and Khodjamirian (2001), 0.83 GeV for Narison (2004) and 0.77 GeV for Ioffe (2006). To understand how these parameters affect the MEM analysis, the results of the calculation using the same central values for the mock data, but different errors and $M_{\min}$, are shown in Fig. 5.5. It is observed that the spectral functions are very similar and depend only weakly on the choice of errors and $M_{\min}$.

Finally, it is important to confirm whether the lowest peak that we have observed in all the results so far really originates from the ρ-meson peak of the input spectral function. In other words, we have to verify that the lowest peak obtained from the MEM analysis really disappears when the ρ-meson peak is removed from the input spectral function. The result for this case is given in Fig. 5.6. We see that while we get the same (non-significant) peak around 2.0 GeV as before, which is induced by the sharply rising default model in this region, the lower peak has completely disappeared. This confirms that the lower peak is directly related to the ρ-meson peak and is not generated by any other effects of the calculation.

5.2.4 Estimation of the Precision of the Final Results

To obtain an estimate of the precision of the current approach, let us now turn to the numerical results of our analysis of mock data. We regard part c of Fig. 5.3 as our main result, the central value of the peak being $m_{\rho,\text{out}} = 0.84$ GeV. The shift from

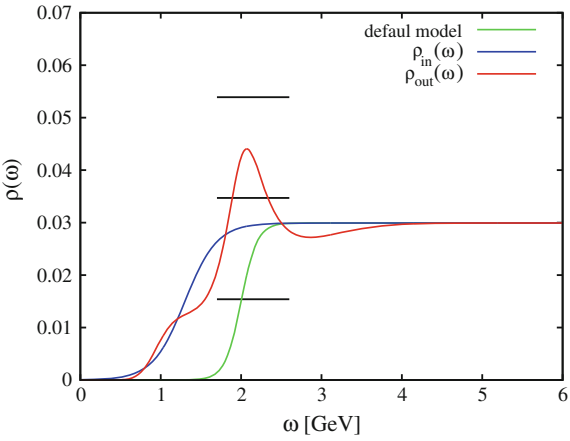

Fig. 5.6 Outcome of the MEM analysis using mock data without the ρ-meson peak. $\rho_{\mathrm{in}}(\omega)$ (*blue line*) is the function that was used to produce the mock data and $\rho_{\mathrm{out}}(\omega)$ (*red line*) shows the result of the MEM procedure. For the default model, the same version as in Fig. 5.3**c** was used (*green line*)

the true value of $m_{\rho,\mathrm{in}} = 0.77$ GeV is caused by the errors of the involved parameters and by the fact that we cannot use all the data points of $G_{\mathrm{mock}}(M)$, but only the ones for which we have some confidence that the OPE converges.

Furthermore, as discussed above, there are some additional uncertainties of ± 40 MeV coming from the choice of the default model and ± 20 MeV from the value of M_{max}. The overall error is then obtained by taking the root of the sum of all squared errors and rounding it up. This gives

$$\Delta m_\rho \simeq \pm 90\,\mathrm{MeV}, \tag{5.6}$$

which is quite large but seems to be realistic when one considers the large errors of the condensates that are involved in the calculation. For the other parametrizations, we get similar results, specifically, $\Delta m_\rho \simeq \pm 60$ MeV for Colangelo and Khodjamirian (2001) and $\Delta m_\rho \simeq \pm 100$ MeV for Narison (2004).

Having the spectral function at our disposal, it becomes possible to extract the coupling strength of the interpolating field to the ρ-meson state, the quantity corresponding to F_ρ in Eq. (5.1). We obtain this coupling strength by fitting the spectral function in the region of the ρ-meson resonance with a relativistic Breit-Wigner peak of the same functional form as the first term of Eq. (5.1) plus a second-order polynomial to describe the continuum background. In order that the background does not become negative and does not contribute to the peak, we have restricted the coefficients of the polynomial to positive values. An example of the resulting curves is given in Fig. 5.7. For part **c** of Fig. 5.3, we get a value of $F_{\rho,\mathrm{out}} = 0.178$ GeV, which is somewhat larger than the true value of $F_{\rho,\mathrm{in}} = 0.141$ GeV. It is not a surprise that the precision of this quantity is worse than that of the peak position, because the shape of the peak is deformed quite strongly owing to the MEM analysis. As can be suspected when looking at Fig. 5.3, there is also a fairly large uncertainty coming from the choice of the default model, which is about ± 0.031 GeV. On the other hand,

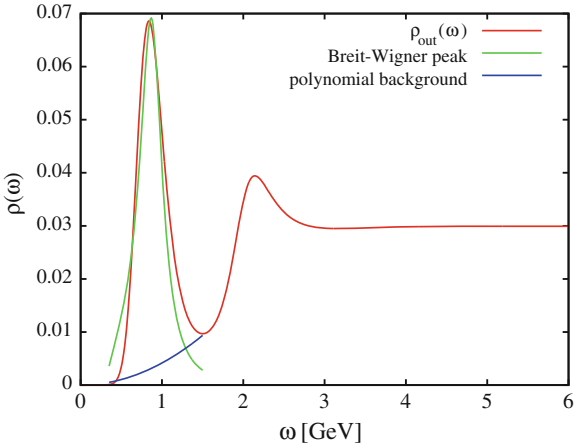

Fig. 5.7 An example of a result of the fitting procedure described in the text. For the spectral function (*red line*) $\rho_{\rm out}(\omega)$, we use here the same as in Fig. 5.3**c**. For the peak, a relativistic Breit-Wigner form is employed (*green line*), while the background is parametrized by a second-order polynomial with positive valued coefficients (*blue line*)

we found that the dependence on the boundaries of the fitting region and on $M_{\rm max}$ is very small, being ± 0.003 and ± 0.001 GeV, respectively. Altogether, this gives the following error for F_ρ:

$$\Delta F_\rho \simeq \pm 0.049\,\text{GeV}. \tag{5.7}$$

A similar analysis for the parameters of Colangelo and Khodjamirian (2001) gives $\Delta F_\rho \simeq \pm 0.038$ GeV and $\Delta F_\rho \simeq \pm 0.049$ GeV for Narison (2004).

As one further point, it is important to investigate if and how the precision of the MEM analysis improves once the OPE data will be available with better precision. To answer this question, we have repeated the analysis using an error with a smaller magnitude and have found that, concerning the pole position, the reproducibility indeed improves with a smaller error. For instance, using the errors obtained from the Ioffe parameters of Ioffe (2006), we get $m_{\rho,\rm out} = 0.84$ GeV, but when we reduce the errors by hand to 20 % of their real value, the result shifts to $m_{\rho,\rm out} = 0.78$ GeV, which almost coincides with the correct value of $m_{\rho,\rm in} = 0.77$ GeV. To a lesser degree, the same is true for the residue. Its value changes from $F_{\rho,\rm out} = 0.178$ GeV to $F_{\rho,\rm out} = 0.167$ GeV with reduced error values, compared with the input value of $F_{\rho,\rm in} = 0.141$ GeV, which is also an improvement. On the other hand, we could not observe a significantly better reproducibility of the width with the reduced error values.

These results show that the MEM analysis of QCD sum rules has the potential to become more accurate in obtaining the position and the residue of the ρ-meson pole, once the values of the condensates are known with better precision. At the same time, it has to be noted that an accurate determination of the width seems to be difficult to achieve in the current approach even with smaller errors. The reason for this is discussed in the following section.

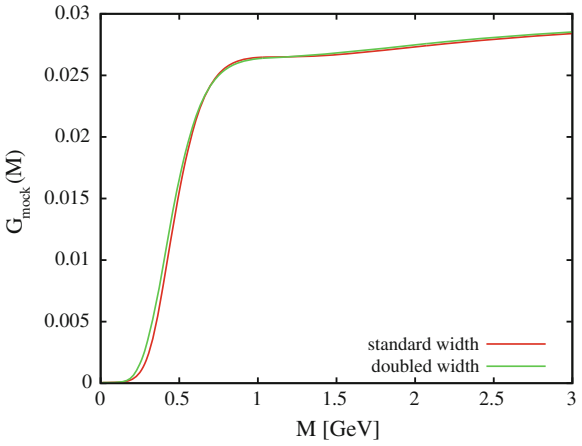

Fig. 5.8 Two versions of $G_{\mathrm{mock}}(M)$, the *red line* showing the one with the standard value for the width as given in Eqs. (5.1) and (5.2), and the *green line* giving a version where $\Gamma_\rho(\omega)$ of Eq. (5.1) is doubled and $F_\rho = 0.149$ GeV is used

5.2.5 Why it is Difficult to Accurately Determine the Width of the ρ Meson

We have so far focused our discussion on the reproduction of the parameters m_ρ and F_ρ of the spectral function and have shown that, by the MEM analysis, they could be reproduced with a precision of approximately 20 %. Considering now the width Γ_ρ, the situation turns out to be quite different, as can be observed for instance in Fig. 5.3. We see there that the values of the width come out about twice as large as in the input spectral function of Eq. (5.1).

The reason for this difficulty of reproducing the width lies in the small dependence of $G_\rho(M)$ on Γ_ρ. This is illustrated in Fig. 5.8. We see there that the curve obtained from our model spectral function of Eqs. (5.1) and (5.2) and the one generated from a spectral function for which the width of the ρ-meson peak has been doubled and a slightly larger value for F_ρ has been used ($F_\rho = 0.149$ GeV compared with the standard value of 0.141 GeV), almost coincide. This means that with the precision available for our QCD sum rule analysis, it is practically impossible to distinguish between these two cases. Examining the curves of Fig. 5.8 a bit more carefully, it is found that the most prominent difference between them lies in the region of 0.5 GeV or below. However, this region cannot be accessed by the current calculation, as the OPE does not converge well for such small values of M. Furthermore, even if we could have calculated the OPE to higher orders and would thus have some knowledge about $G_{OPE}(M)$ in the region below 0.5 GeV, this would most likely not help much, as the uncertainty here will be large owing to the large number of unknown condensates that will appear at higher orders of the OPE. We therefore have to conclude that it is not possible to say anything meaningful about the width

of the ρ-meson peak at the current stage. To predict the width, the OPE has to be computed to higher orders and the various condensates have to be known with much better precision than they are today.

5.3 Analysis Using the OPE Results

5.3.1 The ρ Meson Sum Rule

Carrying out the OPE and applying the Borel transformation, we obtain the left-hand side of Eq. (4.1), $G_{\mathrm{OPE}}(M)$. In the case of the vector meson channel, we use the operator

$$j_\mu = \frac{1}{2}(\bar{u}\gamma_\mu u - \bar{d}\gamma_\mu d), \tag{5.8}$$

which stands for J in Eq. (3.1) and take the terms proportional to the structure $q_\mu q_\nu - q^2 g_{\mu\nu}$. We then arrive at the following expression for $G_{OPE}(M)$, where the OPE has been calculated up to dimension 6:

$$\begin{aligned} G_{OPE}(M) =& \frac{1}{4\pi^2}\Big(1+\eta(\alpha_s)\Big) + \Big(2m\langle\bar{q}q\rangle + \frac{1}{12}\Big\langle\frac{\alpha_s}{\pi}G^2\Big\rangle\Big)\frac{1}{M^4} \\ & - \frac{112\pi}{81}\alpha_s\kappa\langle\bar{q}q\rangle^2\frac{1}{M^6} + \cdots, \\ \eta(\alpha_s) =& \frac{\alpha_s}{\pi} + 0.154\alpha_s^2 - 0.372\alpha_s^3 + \cdots. \end{aligned} \tag{5.9}$$

Here, α_s is the usual strong coupling constant, given as $\frac{g^2}{4\pi}$, m stands for the (averaged) quark mass of the u- and d-quarks, and $\langle\bar{q}q\rangle$ is the corresponding quark condensate. Meanwhile, the gluon condensate $\left\langle\frac{\alpha_s}{\pi}G^2\right\rangle$ is an abbreviated expression for $\left\langle\frac{\alpha_s}{\pi}G^a_{\mu\nu}G^{a\mu\nu}\right\rangle$ and κ parametrizes the breaking of the vacuum saturation approximation, which has been used to obtain the above result for $\kappa = 1$.

A few comments about this result are in order here. For the perturbative term, which is known up to the third order in α_s, we have taken the number of flavors to be $N_f = 4$. Note that only the second and third terms of $\eta(\alpha_s)$ depend (weakly) on N_f (Surguladze and Samuel 1996) and that the final results of the analysis are thus not affected by this choice. We have not considered the running of α_s in deriving Eq. (5.9) for simplicity. If the running is taken into account, the coefficient of the third term of $\eta(\alpha_s)$ changes due to the Borel transformation, as was shown in Shifman (1998). Nevertheless, this again leads only to a minor change in the whole expression of Eq. (5.9) and does not alter any of the results shown in the following. Considering the terms proportional to $1/M^4$, the first-order corrections of the Wilson coefficients are in fact known (Surguladze and Samuel 1996), but we have not included them here as the values of the condensates themselves have large uncertainties, and it is

Table 5.1 Values of the parameters used in the calculation. These have been adjusted to the renormalization scale of 1 GeV

	Colangelo and Khodjamirian (2001)	Narison (2004)	Ioffe (2006)
$\langle\bar{q}q\rangle$ [GeV^3]	$-(0.24 \pm 0.01)^3$	$-(0.238 \pm 0.014)^3$	$-(0.248 \pm 0.013)^3$
$\left\langle\frac{\alpha_s}{\pi}G^2\right\rangle$ [GeV^4]	0.012 ± 0.0036	0.0226 ± 0.0029	0.009 ± 0.007
κ	1	2.5 ± 0.5	1.0 ± 0.1
α_s	0.5	0.50 ± 0.07	0.57 ± 0.08

therefore not necessary to determine the corresponding coefficients with such a high precision. It is nonetheless important to note that these corrections are small (namely of the order of 10 % or smaller, compared with the leading terms) and thus do not introduce any drastic changes into the sum rules.

5.3.1.1 Values of the Parameters and Their Uncertainties

There are various estimates of the values of the condensates and their ranges. We will employ the ones given in three recent publications: Colangelo and Khodjamirian (2001), Narison (2004), Ioffe (2006). The explicit values are given in Table 5.1, where they have been adjusted to the renormalization scale of 1 GeV. For the central value of $m\langle\bar{q}q\rangle$, we make use of the Gell-Mann-Oakes-Renner relation, which gives $m\langle\bar{q}q\rangle = -\frac{1}{2}m_\pi^2 f_\pi^2$ and take the experimental values of m_π and f_π for all three cases, leading to

$$m\langle\bar{q}q\rangle = -8.5 \times 10^{-5}\,\text{GeV}^4. \tag{5.10}$$

Note that due to its smallness, the term containing $m\langle\bar{q}q\rangle$ does not play an important role in the sum rules of Eq. (5.9). The values of Table 5.1 agree well for $\langle\bar{q}q\rangle$ and α_s, while they differ considerably for $\left\langle\frac{\alpha_s}{\pi}G^2\right\rangle$ and κ. Namely, Narison (2004) employs values for $\left\langle\frac{\alpha_s}{\pi}G^2\right\rangle$ and κ that are about two times larger than those of Colangelo and Khodjamirian (2001) and Ioffe (2006). Considering the error estimates of the parameters, Colangelo and Khodjamirian (2001) uses altogether the smallest errors as the breaking of the vacuum saturation approximation is not considered and only a fixed value for α_s is employed. Comparing the results obtained from these three parameter sets will provide us with an estimate of the order of the systematic error inherent in the current calculation.

5.3.1.2 Determination of the Errors of $G_{OPE}(M)$

As can be inferred from Eq. (5.9) and Table 5.1, the uncertainty of $G_{OPE}(M)$ will vary as a function of M and will be larger for small values of M because the contributions of the higher-order terms with large uncertainties of the condensates become more

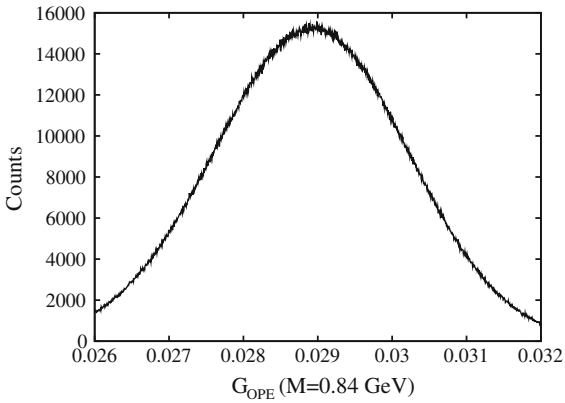

Fig. 5.9 An example of the distributed values of $G_{OPE}(M)$ for $M = 0.84$ GeV. Here, the parameters of Ioffe (2006) have been used

important in that region. To accurately estimate this error, we follow Leinweber (1997) and numerically generate Gaussianly distributed values for the condensates, and then examine the distribution of the resulting values of $G_{OPE}(M)$. For illustration, we show the values of $G_{OPE}(M = 0.84\,\text{GeV})$ in Fig. 5.9, where the parameter estimates of Ioffe (2006) have been used. $\sigma(M)$, the error of $G_{OPE}(M)$, can then be easily extracted from this distribution by using the formula of the standard deviation of a given set.

As for the analysis of the mock data, the data points of $G_{OPE}(M)$ are randomly generated, using Gaussian distributions with standard deviations $\sigma(M)$, centered at $G_{OPE}(M)$ of Eq. (5.9). We here again take $N_M = 100$ and $N_\omega = 600$. M_{min} is determined from the 10 % convergence criterion and M_{max} is fixed to 1.5 GeV.

5.3.2 *Results of the MEM Analysis*

Having finished all the necessary preparations, we now proceed to the actual MEM analysis of the real OPE data. First, we show the central values of the right-hand side of Eq. (5.9) and the corresponding errors for the three parameter sets of Colangelo and Khodjamirian (2001), Narison (2004), Ioffe (2006) on the left side of Fig. 5.10. Comparing these figures with Fig. 5.1, we see that the OPE results and the mock data obtained from Eq. (5.1) exhibit a very similar behavior, even in the region smaller than M_{min}, below which we have no control over the convergence of the OPE.

Using these data, we have carried out the MEM analysis. For the default model, we have adapted the function

$$m(\omega) = \frac{1}{4\pi^2}\Big(1 + \eta(\alpha_s)\Big)\frac{1}{1 + e^{(\omega_0 - \omega)/\delta}} \tag{5.11}$$

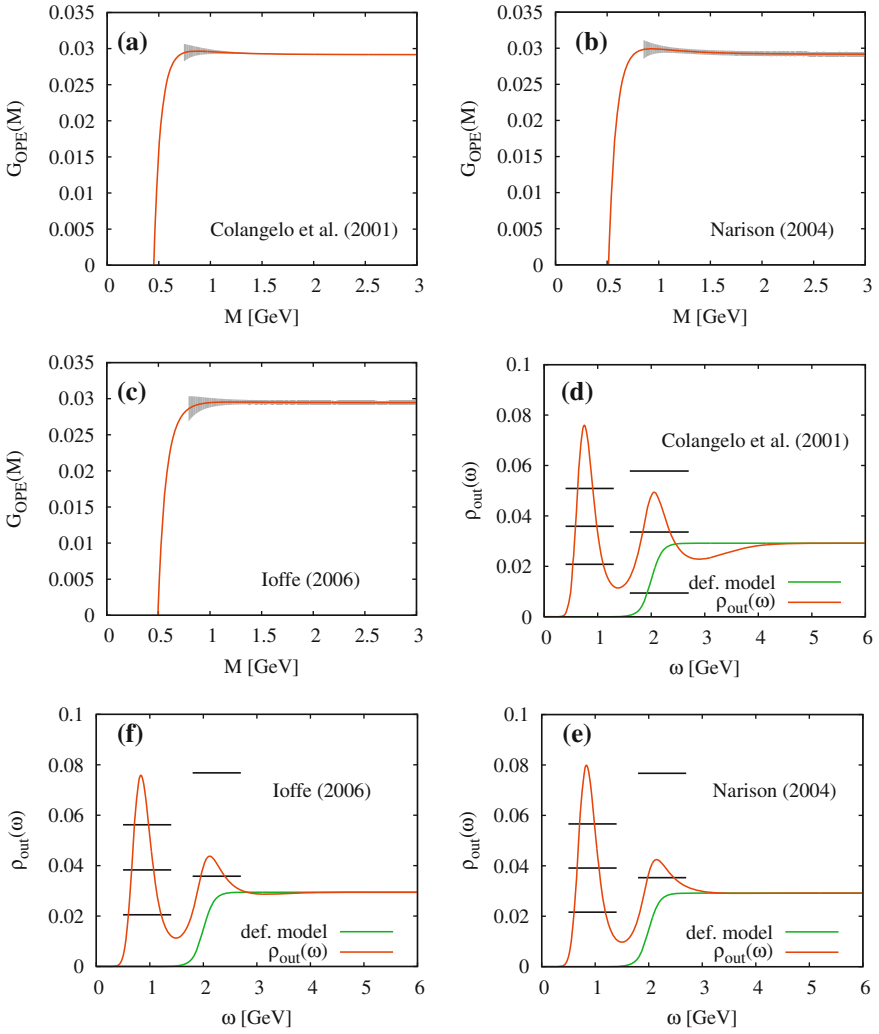

Fig. 5.10 On the *left*, the central values of $G_{OPE}(M)$ (*red lines*) with the errors $\sigma(M)$ (*gray region*) are given. The lower boundary of the shown errors corresponds to $M_{\min}$, below which the OPE does not converge. These plots should be compared with Fig. 5.1. On the *right*, the results of the MEM analysis using these OPE data are displayed (*red lines*). The *green lines* show the default model, and the *horizontal bars* stand for the values of the spectral function, averaged over the peaks $\langle\rho_{\text{out}}\rangle_{\omega_1,\omega_2}$ and the corresponding ranges $\langle\rho_{\text{out}}\rangle_{\omega_1,\omega_2} \pm \langle\delta\rho_{\text{out}}\rangle_{\omega_1,\omega_2}$. Their extent shows the averaged interval (ω_1, ω_2). For the lower two figures the lower error bars of the second peak are not shown because they lie below $\rho(\omega) = 0$

Table 5.2 Final results for the three parameter sets. The respective errors are determined from our analysis of mock data in the previous section

	Colangelo and Khodjamirian (2001)	Narison (2004)	Ioffe (2006)	Experiment
m_ρ (GeV)	0.75 ± 0.06	0.84 ± 0.10	0.83 ± 0.09	0.77
F_ρ (GeV)	0.172 ± 0.038	0.190 ± 0.049	0.186 ± 0.049	0.141

with parameters $\omega_0 = 2.0$ and $\delta = 0.1$ GeV, which we have found in our investigation of the mock data to give the best reproduction of the ρ-meson peak with the smallest relative error and no large artificial peaks. The results are shown on the right side of Fig. 5.10. We clearly see that all three data sets give a significant lowest peak, which corresponds to the ρ-meson resonance. To determine the peak position, we simply adopt the value, where the peak reaches its highest point. The uncertainty of this quantity has already been estimated in our mock-data analysis and we employ the number that has been obtained there to specify the error of our final results that are given in the first line of Table 5.2.

Next, fitting the spectral functions of Fig. 5.10 to a relativistic Breit-Wigner peak plus a second-order polynomial background, as was done in Fig. 5.7, we have determined the coupling strength F_ρ from our obtained spectral function, leading to the results given in the second line of Table 5.2. As could be expected from our experience with the mock data, we get results that are all somewhat larger than the experimental value.

5.3.2.1 Dependence of the ρ-Meson Peak on the Values of the Condensates

Looking at Tables 5.1 and 5.2, it is interesting to observe that even though the parameter sets of Narison (2004) and Ioffe (2006) are quite different, they lead to almost identical results. This fact can be explained from the dependences of the properties of the ρ-meson resonance on $\langle\bar{q}q\rangle$, $\left\langle\frac{\alpha_s}{\pi}G^2\right\rangle$ and κ, as will be shown in this subsection.

Investigating the relation between the ρ-meson resonance and the condensates is also interesting in view of the behavior of this hadron at finite temperature or density, as the values of the condensates will change in such environments. This will in turn alter the QCD sum rule predictions for the various hadron properties. A detailed study of this kind of behavior of the ρ-meson is nevertheless beyond the scope of the present paper and is left for future investigations. Here, we only discuss the change in the mass of the ρ-meson when the values of the condensates are modified by hand. The behavior of the peak position m_ρ is shown in Fig. 5.11.

For obtaining these results, we have used the errors of Colangelo and Khodjamirian (2001) and the corresponding values of κ and α_s, but have confirmed that the qualitative features of Fig. 5.11 do not depend on the explicit values of these parameters. It is seen that while m_ρ decreases somewhat when $\langle\frac{\alpha_s}{\pi}G^2\rangle$ increases, its value grows quite strongly when $\langle\bar{q}q\rangle$ increases, irrespective of the value of the

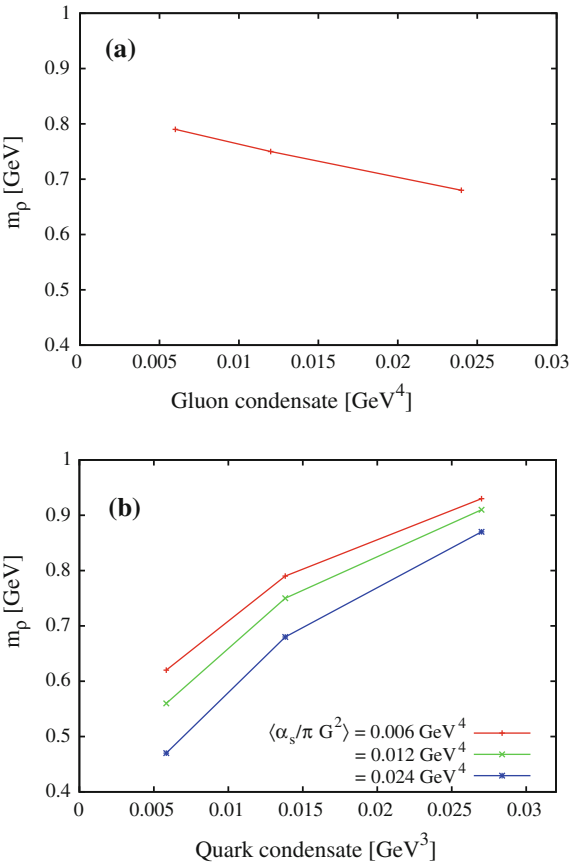

Fig. 5.11 **a** The mass m_ρ shown as a function of the gluon condensate $\langle \frac{\alpha_s}{\pi} G^2 \rangle$. Here, the value of $\langle \bar{q}q \rangle = -(0.24)^3\,\mathrm{GeV}^3$ has been used for the quark condensate. **b** m_ρ shown as a function of the quark condensate $\langle \bar{q}q \rangle$ for three different values of the gluon condensate. To obtain these plots, the errors of Colangelo and Khodjamirian (2001) have been used

gluon condensate. We found that the coupling strength F_ρ shows the same qualitative behavior, slightly decreasing with increasing $\langle \frac{\alpha_s}{\pi} G^2 \rangle$, and strongly increasing with increasing $\langle \bar{q}q \rangle$. A similar tendency for m_ρ and F_ρ has also been observed in Leinweber (1997). This result shows that the quark condensate plays an essential role in determining the properties of the ρ-meson.

It is important to note here that the correlation between m_ρ and $\langle \bar{q}q \rangle$ to a large part occurs due to the last term in $G_{OPE}(M)$ of Eq. (5.9), which contains the squared quark condensate. This means that a similar (but weaker) correlation exists between m_ρ and κ, which is also present in the last term of $G_{OPE}(M)$. Hence, we can now understand why the parameters of Narison (2004) and Ioffe (2006) give such similar results: while the large value of the gluon condensate of Narison (2004) should lead to a smaller m_ρ, this effect is compensated by the large value of κ, which shifts the mass upwards. Therefore, the sum of these changes cancel each other out to a large degree, the net effect being almost no change in the spectral function for both cases.

5.4 Summary and Conclusion

We have applied the MEM technique to the problem of analyzing QCD sum rules. Using MEM has the advantage that we are not forced to introduce any explicit functional form of the spectral function, such as the "pole + continuum" ansatz that has often been employed in QCD sum rule studies. This therefore allows us to investigate any spectral function without prejudice to its actual form. Furthermore, with this technique, we have direct access to the spectral function without the need for interpreting quantities that depend on unphysical parameters such as the Borel mass M and the threshold parameter s_{th}.

To check whether it is really possible to apply MEM to QCD sum rules, we have investigated the vector meson channel in detail, first with mock data obtained from a realistic model spectral function, and then with the actual Borel-transformed results of the operator product expansion. The main results of this investigation are summarized as follows:

- Most importantly, demonstrating that it is possible to extract a significant peak in the spectral function, which corresponds to the ρ-meson resonance, we could show that the MEM technique is quite useful for analyzing QCD sum rules. For both mock or OPE data, we were able to reproduce the experimental ρ-meson mass m_ρ with a precision of about 10 % and the coupling strength F_ρ with a precision of about 30 %.
- We have found that owing to the properties of the kernel of Eq. (5.3), the default model $m(\omega)$ has to be chosen according to the correct behavior of the spectral function at low energy. We therefore have taken a default model that tends to zero at $\omega = 0$. On the other hand, to give the correct behavior at large energies, the same default model is constructed to approach the perturbative value in the high-energy region.
- The position of the ρ-meson peak in the spectral function has turned out to be quite stable against changes in various parameters of the calculation, such as the details of the default model or the range of the analyzed Borel mass region. We have shown that changing these parameters leads to a fluctuation of the peak position of only 20–50 MeV.
- Concerning the width of the lowest lying peak, we are unable to reproduce the value of the input spectral function of the mock data with any reasonable precision and have shown that the reason for this difficulty comes from the insufficient sensitivity of the data $G_\rho(M)$ on the detailed form of the ρ-meson peak. To accurately estimate the width, one needs not only to go to higher orders in the OPE, but also has to have much more precise information on the values of the condensates than is available today.
- Accompanied by a steep rise in the default model $m(\omega)$, we have observed the appearance of artificial peaks in the output spectral function of the MEM analysis. These peaks are MEM artifacts and one has to be careful not to confuse them with the actual peaks that are predicted using the OPE data.

These results are promising and encourage us to apply this approach to other channels, including baryonic ones. It would also be interesting to apply this method to the behavior of hadrons at finite density or temperature, as it would become possible to directly observe the change in the spectral function in hot or dense environments by this approach. Furthermore, it would be of interest to study the various exotic channels, containing more than three quarks. In these channels, the scattering states presumably play an important role and the Bayesian approach of this paper could help clarify the situation and provide a natural way of distinguishing genuine resonances from mere scattering states.

Even though these are interesting subjects for future studies, we want to emphasize here that the uncertainties involved for each channel can differ considerably and one thus should always carry out a detailed analysis with mock data for each case, before investigating the actual sum rules. This procedure is necessary to check whether it is possible to obtain meaningful results from the MEM analysis of QCD sum rules.

References

Asakawa M, Hatsuda T, Nakahara Y (2001) Prog Part Nucl Phys 46:459
Colangelo P, Khodjamirian A (2001) At the frontier of particle physics/ handbook of QCD, vol 3. World Scientific, Singapore, p 1495
Gubler P, Oka M (2010) Prog Theor Phys 125:995
Gubler P, Jido D, Kojo T, Nishikawa T, Oka M (2009a) Phys Rev D79:114011
Gubler P, Jido D, Kojo T, Nishikawa T, Oka M (2009b) Phys Rev D80:114030
Ioffe BL (2006) Prog Part Nucl Phys 56:232
Jarrel M, Gubernatis JE (1996) Phys Rep 269:133
Kojo T, Jido D (2008) Phys Rev D78:114005
Leinweber DB (1997) Ann Phys 254:328
Nakahara Y, Asakawa M, Hatsuda T (1999) Phys Rev D60:091503(R).
Narison S (2004) QCD as a theory of hadrons, from partons to confinement. Cambridge University Press, Cambridge
Reinders LJ, Rubinstein H, Yazaki S (1985) Phys Rep 127:1
Shifman MA, Vainshtein AI, Zakharov VI (1979) Nucl Phys B147(385):448
Shifman M (1998) Prog Theor Phys Suppl 131:1
Shuryak EV (1993) Rev Mod Phys 65:1
Surguladze L, Samuel M (1996) Rev Mod Phys 68:259
Wu N (1997) The maximum entropy method. Springer, Berlin

Chapter 6
MEM Analysis of the Nucleon Sum Rule

6.1 Introduction

The use of QCD sum rules in studies investigating the properties of baryons has already a long history. Since the seminal papers of Shifman et al. (1979a, b), in which the QCD sum rule method was developed and the subsequent first application to baryonic channels by Ioffe (1981a, b), the sum rules of the nucleon have been continuously improved by including higher orders in the perturbative Wilson coefficients (Krasnikov et al. 1983; Chung et al. 1984; Jamin 1988; Ovchinnikov et al. 1991; Shiomi and Hatsuda 1995; Sadovnikova et al. 2005) or non-perturbative power corrections (Chung et al. 1984; Belyaev and Ioffe 1982; Leinweber 1990). This even led to attempts to determine the mass difference between the neutron and the proton (Yang et al. 1993), which certainly is a very difficult task because of the smallness of this difference compared to hadronic scales. Another important development was initiated in a paper by Leinweber (1997), in which among other technical points, the choice of the interpolating field made by Ioffe was criticized and a new statistical method for the analysis of QCD sum rules was introduced. Furthermore, QCD sum rules also have been used to investigate the nucleon properties in nuclear matter (Hatsuda et al. 1991; Cohen et al. 1995; Drukarev et al. 2010) or at finite temperature (Adami and Zahed 1992).

However, in all these studies it was necessary to model the spectral function according to some specific functional form, the "pole + continuum"-ansatz being the most popular one. Such a procedure inevitably incorporates strong assumptions on the spectral function into the analysis. This strategy works well when the actual spectral function has some resemblance with the chosen model, but will of course fail if this is not the case. For instance, as is known from studies using both QCD sum rules and lattice QCD, certain linear combinations of interpolating fields, which in principle carry the quantum numbers of the nucleon, couple to the nucleon ground state only very weakly. A QCD sum rule analysis of such interpolating fields, which uses the "pole + continuum" ansatz, can only lead to ambiguous results. The problems become even more severe in studies of the spectral function at finite temperature

P. Gubler, *A Bayesian Analysis of QCD Sum Rules*, Springer Theses,
DOI: 10.1007/978-4-431-54318-3_6, © Springer Japan 2013

or density, as the validity of the phenomenological "pole + continuum" ansatz for the spectral function in such an environment becomes less certain, rendering it more difficult to make educated guesses about its actual form.

In this chapter, it is our main purpose to examine if and to what extent the MEM analysis can be applied to the sum rule of the nucleon, in order to solve the above mentioned problems (Ohtani et al. 2011). Throughout our investigations, we found that the MEM analysis of the Borel sum rule

$$G_{\mathrm{OPE}}(M) = \frac{1}{M^2}\int_0^\infty dt e^{-\frac{t}{M^2}}\rho(t), \tag{6.1}$$

in fact fails to satisfactorily extract the nucleon spectral function in the ground state region. As we will discuss later in detail, this failure is mainly caused by the large contribution of the continuum to the OPE side of Eq. (6.12), which strongly deteriorates the contribution of the nucleon pole.

On the other hand, the Gaussian sum rule (Bertlmann et al. 1985; Orlandini et al. 2001)

$$G_{\mathrm{OPE}}(s,\tau) = \frac{1}{\sqrt{4\pi\tau}}\int_0^\infty dt e^{-\frac{(t-s)^2}{4\tau}}\rho(t), \tag{6.2}$$

which, for the nucleon, is formulated for the first time in this paper, turns out to give better results and allows us to resolve the nucleon pole from the continuum. There are essentially two reasons for the superiority of the sum rule of Eq. (6.2). First of all, the kernel of Eq. (6.2), a Gaussian centered at s with a width of $\sqrt{2\tau}$, collects more information on the spectral function $\rho(t)$ than the one in Eq. (6.12) when the integration over t is carried out. This is especially true for small values of the width τ, which we, however, cannot take arbitrarily small because below a certain threshold, the convergence of the operator product expansion (OPE) becomes poor. A similar situation occurs in the Borel sum rules, where the Borel mass is restricted from below due to the OPE convergence. The second reason also can be related to the kernel of Eq. (6.2), containing two parameters s and τ, which can be freely chosen as long as the OPE converges. This freedom allows us to vary two parameters at the same time in the MEM analysis, leading to reasonable results for the extracted spectral function. Similar experiences also have been made in nuclear structure studies, where the Lorentz kernel has proven to be useful (Efros et al. 1994, 2007).

Furthermore, using the MEM analysis of the Gaussian sum rules, we are able to extract the spectral function not only in cases where the interpolating field strongly couples to the nucleon ground state and thus the "pole + continuum" ansatz should be valid, but also in cases where only higher energy states contribute to the sum rules and hence the conventional analysis most likely fails to give meaningful results. This advantage is especially useful for examining which kind of interpolating field couples strongly to the nucleon ground state and is thus a suitable interpolating field for the analysis, a question with a long and controversial history in QCD sum rules studies (Ioffe 1981a, b; Leinweber 1997; Chung et al. 1982a, b; Ioffe 1983; Dosch et al. 1989). Our MEM analysis of the general operator given in Eq. (6.5) strongly

suggests that the nucleon ground state only couples to $\eta_1(x)$ ($\beta = 0$) and not to $\eta_2(x)$ ($\beta = \infty$) (see Eqs. (6.3) and (6.4) in Sect. 6.2). In addition, we have obtained some hints of excited states coupling to $\eta_2(x)$. These issues will be discussed in detail in Sect. 6.4.3.

This chapter is organized as follows. In Sect. 6.2, the details of the Borel and Gaussian sum rule for the nucleon are explained. Next, the results of the analysis using the Borel sum rule in combination with MEM are presented in Sect. 6.3. In Sect. 6.4, the results of the analysis for the Gaussian sum rule are outlined, and in Sect. 6.4.3, the differences of the obtained spectral functions depending on the choice of the interpolating fields are discussed. Finally, the summary and conclusions are given in Sect. 6.5.

6.2 QCD Sum Rules for the Nucleon

The method of QCD sum rules is used to carry out the analysis of the spectral function as follows. First, we choose an interpolating field which has the quantum numbers of the nucleon, then its correlation function is calculated in the deep Euclidean 4-momentum region. Alternatively, the same correlation function at the physical 4-momentum region is expressed by the spectral function of the nucleon channel. The sum rules can then be constructed by equating the two expressions using a dispersion relation.

For the nucleon, there are two independent local interpolating operators,

$$\eta_1(x) = \varepsilon^{abc}(u^{Ta}(x)C\gamma_5 d^b(x))u^c(x), \tag{6.3}$$

$$\eta_2(x) = \varepsilon^{abc}(u^{Ta}(x)Cd^b(x))\gamma_5 u^c(x). \tag{6.4}$$

Here, abc are color indices, C is the charge conjugation matrix and T stands for the transposition operation. The spinor indices are omitted for simplicity. A general interpolating operator can thus be expressed as

$$\eta(x) = \eta_1(x) + \beta\eta_2(x), \tag{6.5}$$

where β is a real parameter. Here, the case of $\beta = -1$ is identified as the so-called "Ioffe current" (Ioffe 1981a, b), which is often used in QCD sum rule studies of the nucleon.

Using this interpolating operator, we define the correlation function as

$$\begin{aligned} \Pi(q) &= i\int e^{iqx}\langle 0|T[\eta(x)\overline{\eta}(0)]|0\rangle d^4x \\ &= \not{q}\Pi_1(q^2) + \Pi_2(q^2). \end{aligned} \tag{6.6}$$

The imaginary part of $\Pi_1(q^2)$ satisfies the positivity condition, $\mathrm{Im}\Pi_1(q^2) \geq 0$, while $\mathrm{Im}\Pi_2(q^2)$ is not necessarily positive due to contributions of negative parity states. The positivity condition is, however, essential for the application of the MEM method and we thus consider only $\Pi_1(q^2)$ in the following. To make use of the information contained in $\mathrm{Im}\Pi_2(q^2)$, one would have to analyze the parity projected sum rules (Jido et al. 1996), which we plan to investigate in the future. In the deep Euclidean region ($-q^2 \to \infty$), $\Pi_1(q^2)$ can be calculated by using the operator product expansion (OPE). Including operators up to dimension 8, we get

$$\begin{aligned}\Pi_1(q^2) = &-\frac{5+2\beta+5\beta^2}{128(2\pi)^4}q^4\ln(-q^2)\\ &-\frac{5+2\beta+5\beta^2}{256(2\pi)^2}\langle\frac{\alpha_s}{\pi}G^2\rangle\ln(-q^2)\\ &-\frac{7-2\beta-5\beta^2}{24}\langle\bar{q}q\rangle^2\frac{1}{q^2}\\ &-\frac{13-2\beta-11\beta^2}{96}\langle\bar{q}q\rangle\langle\bar{q}g\sigma\cdot Gq\rangle\frac{1}{q^4}\\ \equiv\ &\Pi_{\mathrm{OPE}}(q^2)\end{aligned} \tag{6.7}$$

To obtain Eq. (6.7), several approximations have been implemented. Firstly, only the lowest order in α_s is taken into account. The validity of this approximation is not obvious, because it is known that the first order α_s corrections are significant and lead to a considerable increase of the continuum contribution (Leinweber 1997). Nevertheless, our main goal of this paper is to examine whether the MEM analysis can be applied to the nucleon sum rule or not, and we thus ignore the α_s corrections here. For a more quantitative future analysis, the higher order corrections should certainly be taken into account. The second approximation arises from the use of the vacuum saturation, by which $\langle\bar{q}q\bar{q}q\rangle$ and $\langle\bar{q}q\bar{q}g\sigma\cdot Gq\rangle$ can be formally reduced to $\langle\bar{q}q\rangle^2$ and $\langle\bar{q}q\rangle\langle\bar{q}g\sigma\cdot Gq\rangle$, respectively. Although this approximation can be justified in the large N_c limit (Shifman et al. 1979a, b), it is not clear to what extent it is trustable at $N_c = 3$. Nonetheless, for the present qualitative analysis, we will assume this approximation to be valid.

As already mentioned, $\Pi_1(q^2)$ can also be expressed in terms of the physical spectral function using the dispersion relation:

$$\begin{aligned}\Pi_1(q^2) &= \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}\Pi_1(t)}{t-q^2}dt = \int_0^\infty \frac{\rho(t)}{t-q^2}dt\\ &\equiv \Pi_\rho(q^2),\end{aligned} \tag{6.8}$$

where the definition $\mathrm{Im}\Pi_1(t) = \pi\rho(t)$ is used for the spectral function. Our goal is now to extract $\rho(t)$ from the sum rule obtained by equating Eqs. (6.7) and (6.8). It should be noted here that subtractions are necessary in order to make the integral of Eq. (6.8) convergent. In the case of the nucleon, the subtraction terms are

$\Pi_1(0) + \Pi_1'(0)q^2 + \frac{\Pi_1''(0)}{2}q^4$, which will disappear after transforming Eq. (6.8) into the Borel or Gaussian sum rules. How this is done will be explained in the following subsections.

6.2.1 Borel Sum Rule

In the case of the Borel sum rule, we transform $\Pi_1(q^2)$ using the Borel transformation $\hat{B}_M$, defined below:

$$\hat{B}_M = \lim_{\substack{-q^2,n\to\infty \\ -q^2/n=M^2}} \frac{(-q^2)^n}{(n-1)!}\left(\frac{d}{dq^2}\right)^n. \tag{6.9}$$

Applying $\hat{B}_M$ to Eq. (6.7), we get for $G_{\mathrm{OPE}}(M) \equiv \hat{B}_M \Pi_{\mathrm{OPE}}(q^2)$:

$$\begin{aligned} G_{\mathrm{OPE}}(M) = {} & \frac{5+2\beta+5\beta^2}{64(2\pi)^4}M^4 \\ & + \frac{5+2\beta+5\beta^2}{256(2\pi)^2}\left\langle\frac{\alpha_s}{\pi}G^2\right\rangle \\ & + \frac{7-2\beta-5\beta^2}{24}\langle\bar{q}q\rangle^2\frac{1}{M^2} \\ & - \frac{13-2\beta-11\beta^2}{96}\langle\bar{q}q\rangle\langle\bar{q}g\sigma\cdot Gq\rangle\frac{1}{M^4}. \end{aligned} \tag{6.10}$$

Meanwhile, applying the Borel transformation to Eq. (6.8), $G_\rho(M) \equiv \hat{B}_M \Pi_\rho(q^2)$ is obtained as:

$$G_\rho(M) = \frac{2}{M^2}\int_0^\infty e^{-\frac{\omega^2}{M^2}}\omega\rho(\omega)d\omega. \tag{6.11}$$

Here, $t = \omega^2$ was used. Note that at this point, the subtraction terms have been eliminated and the integral of the right hand side converges. This leads us to the final form of the Borel sum rule,

$$G_{\mathrm{OPE}}(M) = G_\rho(M) = \frac{2}{M^2}\int_0^\infty e^{-\frac{\omega^2}{M^2}}\omega\rho(\omega)d\omega. \tag{6.12}$$

6.2.2 Gaussian Sum Rule

The Gaussian sum rule, first introduced in Bertlmann et al. (1985), exhibits another way of improving Eq. (6.8). Based on the idea of local duality, it provides a formulation for the convolution of the spectral function with a Gaussian kernel. As this

sort of sum rule is not often discussed in the literature and the specific case of the nucleon has to our knowledge not even been formulated, we will explain each step in some detail, following closely the formulation given in Orlandini et al. (2001).

Going to the complex plane of q^2 and taking the difference between $\Pi_1(q^2 = s + i\Delta)$ and $\Pi_1(s - i\Delta)$ of Eq. (6.8), where s and Δ are real, we obtain

$$\frac{\Pi_1(s + i\Delta) - \Pi_1(s - i\Delta)}{2i\Delta} = \int_0^\infty \frac{\rho(t)}{(t - s)^2 + \Delta^2} dt. \tag{6.13}$$

At this stage, the integral above is not convergent and the subtraction terms are not yet fully eliminated. Applying the following Borel transform $\hat{B}_{4\tau}$:

$$\hat{B}_{4\tau} = \lim_{\substack{\Delta^2, n \to \infty \\ \Delta^2/n = 4\tau}} \frac{(-\Delta^2)^n}{(n-1)!} \left(\frac{d}{d\Delta^2}\right)^n, \tag{6.14}$$

and using

$$\hat{B}_{4\tau} \left[\frac{1}{(t - s)^2 + \Delta^2}\right] = \frac{1}{4\tau} e^{-\frac{(t-s)^2}{4\tau}}, \tag{6.15}$$

we get

$$\begin{aligned} 2\sqrt{\frac{\tau}{\pi}} \hat{B}_{4\tau} \left(\frac{\Pi_1(s + i\Delta) - \Pi_1(s - i\Delta)}{2i\Delta}\right) &= \frac{1}{\sqrt{4\pi\tau}} \int_0^\infty dt e^{-\frac{(t-s)^2}{4\tau}} \rho(t) \\ &\equiv G_\rho(s, \tau). \end{aligned} \tag{6.16}$$

Here, the subtraction terms have disappeared and the integral in the above equation is convergent.

As a next step, we will now show that Eq. (6.16) can also be rewritten by using the inverse Laplace transform $\hat{L}^{-1}$. Using

$$\begin{aligned} \frac{1}{(t - s)^2 + \Delta^2} &= \hat{L}[e^{-x(t-s)^2}] \\ &= \int_0^\infty e^{-\Delta^2 x} e^{-x(t-s)^2} dx, \end{aligned}$$

the kernel $e^{-x(t-s)^2}$ can be re-expressed by $\hat{L}^{-1}$:

$$\begin{aligned} e^{-x(t-s)^2} &= \hat{L}^{-1} \left[\frac{1}{(t - s)^2 + \Delta^2}\right] \\ &= \frac{1}{2\pi i} \int_{b-i\infty}^{b+i\infty} e^{\Delta^2 x} \frac{1}{(t - s)^2 + \Delta^2} d\Delta^2. \end{aligned} \tag{6.17}$$

Setting $x = \frac{1}{4\tau}$, the left-hand side of Eq. (6.16) can thus be rewritten as

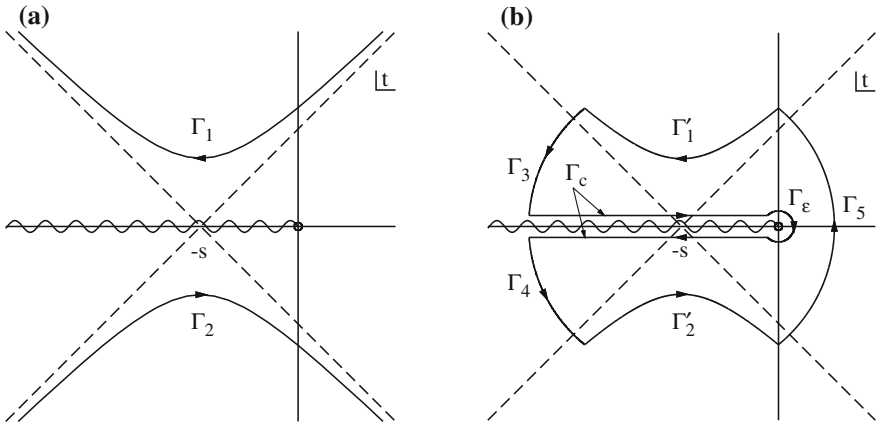

Fig. 6.1 **a** Contour of integration $\Gamma_1 + \Gamma_2$ in Eq. (6.18). The branch cut of $\Pi_1(-t)$ is represented by the *wavy line* on the negative real axis. **b** Closed contour $C(R)$ of Eq. (6.19). The *outer lines* Γ_3, Γ_4 and Γ_5 are a part of the circle of radius R centered at $-s$, while Γ_ε denotes a *circular line* with radius ε, centered at the origin. The branch cut of $\Pi_1(-t)$ is represented by the *wavy line* on the negative real axis and the contour Γ_c is the *straight line* of the contour above and below the branch cut. The contours Γ_1' and Γ_2' are parts of Γ_1 and Γ_2, respectively

$$\frac{1}{\sqrt{4\pi\tau}}\hat{L}^{-1}\left[\frac{\Pi_1(s+i\Delta)-\Pi_1(s-i\Delta)}{2i\Delta}\right]$$
$$=\frac{1}{\sqrt{4\pi\tau}}\frac{1}{2\pi i}\int_{b-i\infty}^{b+i\infty}\left[\frac{\Pi_1(s+i\Delta)-\Pi_1(s-i\Delta)}{2i\Delta}\right]e^{\frac{\Delta^2}{4\tau}}d\Delta^2.$$

Then, replacing $s+i\Delta$ by $-t$ in the first and $s-i\Delta$ by $-t$ in the second term, we get

$$G_\rho(s,\tau)=\frac{1}{\sqrt{4\pi\tau}}\frac{1}{2\pi i}\int_{\Gamma_1+\Gamma_2}\Pi_1(-t)\exp\left(-\frac{(t+s)^2}{4\tau}\right)dt \tag{6.18}$$

where the contour $\Gamma_1+\Gamma_2$ is shown in Fig. 6.1**a**.

Next, to obtain a sum rule that is practically usable, we consider the contour $C(R)$ shown in Fig. 6.1**b**. Taking the limit $R\to\infty$ and $\varepsilon\to 0$, we are led to the equation given below:

$$G_\rho(s,\tau)=-\lim_{R\to\infty}\frac{1}{\sqrt{4\pi\tau}}\frac{1}{2\pi i}\int_{\Gamma_c+\Gamma_\varepsilon}\Pi_1(-t)\exp\left(-\frac{(t+s)^2}{4\tau}\right)dt. \tag{6.19}$$

Substituting the right-hand side of Eq. (6.7) into Eq. (6.19) and examining the various terms, we see that the perturbative and dimension four terms only give contributions on the contour Γ_c. Meanwhile, the dimension six and eight terms do not have a branch discontinuity, but a pole at $t=0$ and therefore only contribute on Γ_ε. Using

$$\int_{-\infty}^{0} \exp\left(-\frac{(t+s)^2}{4\tau}\right) dt = \sqrt{\tau\pi}\left[1 + \mathrm{erf}\left(\frac{s}{2\sqrt{\tau}}\right)\right] \tag{6.20}$$

where the error function $\mathrm{erf}(x)$ is defined as

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-x^2} dx \tag{6.21}$$

and

$$-\frac{1}{2\pi i} \int_{\Gamma_\varepsilon} \frac{1}{t^n} \exp\left(-\frac{(t+s)^2}{4\tau}\right) dt = \lim_{t\to 0} \frac{1}{(n-1)!} \frac{d^{n-1}}{dt^{n-1}} \exp\left(-\frac{(t+s)^2}{4\tau}\right) \tag{6.22}$$

$$(n = 1, 2, \ldots),$$

we obtain

$$\begin{aligned} G_{\mathrm{OPE}}(s,\tau) = {} & \frac{5+2\beta+5\beta^2}{128(2\pi)^4} \\ & \times \left[\tau\left(1+\frac{s^2}{2\tau}\right)\left[1+\mathrm{erf}\left(\frac{s}{2\sqrt{\tau}}\right)\right] + \frac{s\sqrt{\tau}}{\sqrt{\pi}}\exp\left(-\frac{s^2}{4\tau}\right)\right] \\ & + \frac{5+2\beta+5\beta^2}{512(2\pi)^2}\langle\frac{\alpha_s}{\pi}G^2\rangle\left[1+\mathrm{erf}\left(\frac{s}{2\sqrt{\tau}}\right)\right] \\ & + \frac{1}{\sqrt{\pi\tau}}\frac{7-2\beta-5\beta^2}{48}\langle\overline{q}q\rangle^2\exp\left(-\frac{s^2}{4\tau}\right) \\ & + \frac{1}{\sqrt{\pi\tau}}\frac{13-2\beta-11\beta^2}{384}\langle\overline{q}q\rangle\langle\overline{q}g\sigma\cdot Gq\rangle\frac{s}{\tau}\exp\left(-\frac{s^2}{4\tau}\right). \end{aligned} \tag{6.23}$$

This then finally leads to the following form of the Gaussian sum rule, from which information of the spectral function $\rho(\omega)$ can be extracted:

$$G_{\mathrm{OPE}}(s,\tau) = G_\rho(s,\tau) = \frac{2}{\sqrt{4\pi\tau}} \int_0^\infty d\omega \cdot \omega e^{-\frac{(\omega^2-s)^2}{4\tau}} \rho(\omega). \tag{6.24}$$

Here, we have again set $t = \omega^2$.

6.3 Analysis Using the Borel Sum Rule

In this section, we will analyze the nucleon spectral function for the Borel sum rule. It is easily understood from dimensional considerations that unlike in the meson case, the contribution of the continuum states to the baryon spectral function is proportional

to ω^4 and thus strongly enhanced. As was done in similar studies using MEM and lattice QCD, we will therefore analyze $\rho(\omega)/\omega^4$ instead of $\rho(\omega)$ and hence from now on denote $\rho(\omega)/\omega^4$ as $\rho(\omega)$, leading to the equations below:

$$G_{\mathrm{OPE}}(M) = G_{\rho}(M) = \int_0^{\infty} d\omega K(M,\omega)\rho(\omega), \tag{6.25}$$

$$K(M,\omega) = \frac{2\omega^5}{M^2} e^{-\omega^2/M^2}. \tag{6.26}$$

6.3.1 Analysis Using Mock Data

In order to check the effectiveness of MEM to extract the spectral function of the nucleon, we first carry out an analysis using mock data. The employed mock spectral function is given below:

$$\rho_{\mathrm{mock}}(\omega) = \frac{\lambda^2}{2M_N^5}\delta(\omega - M_N) + \frac{5+2\beta+5\beta^2}{128(2\pi)^4}\frac{1}{1+e^{\frac{(\omega_0-\omega)}{\delta}}}, \tag{6.27}$$

where we use the following values for the various parameters:

$$M_N = 940\,\mathrm{MeV},\ \omega_0 = 1.3\,\mathrm{GeV},$$
$$\delta = 0.05\,\mathrm{GeV},\ \lambda^2 = \frac{0.19}{(2\pi)^4}\,\mathrm{GeV}^6. \tag{6.28}$$

Here, $\rho_{\mathrm{mock}}(\omega)$ is constructed to have a narrow ground state pole and a continuum, which approaches the perturbative value at high energy. Defining now

$$G_{\mathrm{mock}}(M) \equiv \int_0^{\infty} d\omega K(M,\omega)\rho_{\mathrm{mock}}(\omega), \tag{6.29}$$

we apply the MEM procedure to

$$G_{\mathrm{mock}}(M) = G_{\rho}(M) = \int_0^{\infty} d\omega K(M,\omega)\rho(\omega). \tag{6.30}$$

The residue of the nucleon pole λ was fitted so that the function $G_{\mathrm{mock}}(M)$ matches $G_{\mathrm{OPE}}(M)$ at $\beta = -1$ in the analyzed Borel mass region. For comparison, $G_{\mathrm{OPE}}(M)$ and $G_{\rho}(M)$ are shown in Fig. 6.2. The values of the parameters appearing in Eq. (6.7) are given in Table 6.1. As can be observed in Fig. 6.2, $G_{\mathrm{OPE}}(M)$ and $G_{\rho}(M)$ are consistent within the range of the error, which shows that the "pole + continuum" ansatz describes the OPE data well. However, this does not necessarily imply that the integral of Eq. (6.30) can be reliably inverted and that valid information on the

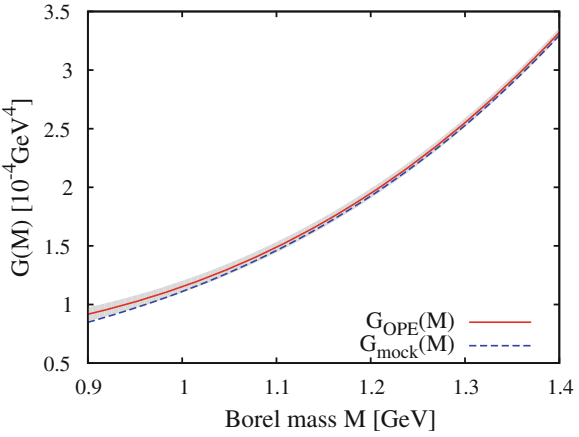

Fig. 6.2 Comparison of the mock data and the OPE data at $\beta = -1$. The *solid red line* shows $G_{\mathrm{OPE}}(M)$ of Eq. (6.10), with the respective error indicated by the *gray region*. The *dashed blue line* denotes $G_{\mathrm{mock}}(M)$ for which the mock spectral function of Eq. (6.27) was used

Table 6.1 Values of the parameters appearing in the OPE, taken from Colangelo and Khodjamirian (2001)

Condensates	$\langle\bar{q}q\rangle[\mathrm{GeV}^3]$	$\langle\frac{\alpha_s}{\pi}G^2\rangle[\mathrm{GeV}^4]$	$\langle\bar{q}g\sigma\cdot Gq\rangle/\langle\bar{q}q\rangle[\mathrm{GeV}^2]$
Used values	$-(0.24\pm0.01)^3$	0.012 ± 0.0036	0.8 ± 0.2

nucleon pole can be extracted. To investigate to what extent this is possible, we now analyze $G_{\mathrm{mock}}(M)$. For a realistic analysis, we here employ the error obtained from $G_{\mathrm{OPE}}(M)$, thus $\sigma_{\mathrm{mock}}(M) = \sigma_{\mathrm{OPE}}(M)$.

To use MEM, we must at first fix the default model $m(\omega)$. A reasonable choice for $m(\omega)$ should reflect our prior knowledge on the spectral function such as the asymptotic behavior at low or high energy. To test several possible choices, we here introduce three types of default models. The first one is a constant consistent only with the asymptotic behavior of the spectral function at low energy, therefore lying close to 0:

$$m_{\mathrm{flat}}(\omega) = 1.0\times10^{-6}. \tag{6.31}$$

The detailed value of m_{flat} is not so important, as long as it can be considered to be small enough compared to the asymptotic value at high energy. As we will discuss in more detail in the section dealing with the gaussian sum rule, we indeed have found that the position of the lowest lying pole and its residue of the obtained spectral functions depend on the value of m_{flat} only weakly. The second default model is also a constant which now reflects the asymptotic behavior at high energy:

$$m_{\mathrm{asym}}(\omega) = \frac{5+2\beta+5\beta^2}{128(2\pi)^4}. \tag{6.32}$$

The third one is a combination of the first two with the correct behavior at both high and low energy:

$$m_{\rm hybr}(\omega) = \frac{5+2\beta+5\beta^2}{128(2\pi)^4}\frac{1}{1+e^{(\omega_0-\omega)/\delta}}, \qquad (6.33)$$
$$\omega_0 = 1.8\,{\rm GeV},\ \delta = 0.1\,{\rm GeV}.$$

In the following, the analysis is carried out at $\beta = -1$. In the case of the conventional method, which assumes the spectral function to have a specific functional form such as the "pole + continuum" ansatz, the analyzed Borel mass region is restricted so that the ratio of the highest dimensional term is less than 0.1 of the whole $G_{\rm OPE}(M)$ to have some confidence on the OPE convergence. The Borel mass region is further limited by the condition that the contribution from the continuum states in $G_\rho(M)$ should be less than 0.5 to make sure that the lowest pole dominates the sum rule. The Borel mass region determined according to these restrictions becomes $0.91\,{\rm GeV} \le M \le 0.97\,{\rm GeV}$. As pointed out in Leinweber (1997), this region is very narrow and thus we expect that it will be very difficult to extract $\rho(\omega)$ in a wide range of ω with this small amount of available information. Although, when using MEM, we can in principle employ values of M above 0.97 GeV, we here first analyze the spectral function using the Borel mass region $0.91\,{\rm GeV} \le M \le 0.97\,{\rm GeV}$. The results are shown in Fig. 6.3.

It is clear from Fig. 6.3, that the obtained lowest peaks lie much above the input value of 940 MeV. Hence, as expected, we cannot extract much information on the

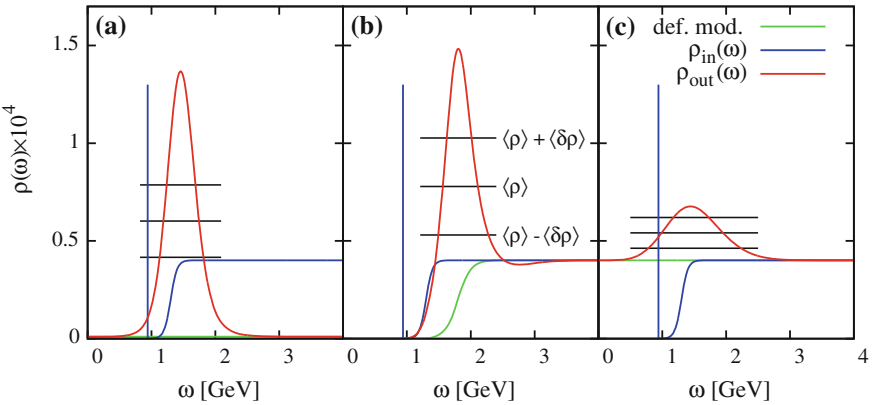

Fig. 6.3 Spectral functions extracted from $G_{\rm mock}(M)$ using various default models at $\beta = -1$. The analyzed Borel mass region is $0.91\,{\rm GeV} \le M \le 0.97\,{\rm GeV}$. The *red lines* show $\rho_{\rm out}(\omega)$, the *blue lines* depict the input mock spectral function and the *green lines* stand for the default model. As explained in Fig. 4.2 of Chap. 4, the three *horizontal lines* correspond to the value of the spectral function, averaged over the peak, $\langle\rho_{\rm out}\rangle_{\omega_1,\omega_2}$ and the respective upper and lower value, $\langle\rho_{\rm out}\rangle_{\omega_1,\omega_2} + \langle\delta\rho_{\rm out}\rangle_{\omega_1,\omega_2}$ and $\langle\rho_{\rm out}\rangle_{\omega_1,\omega_2} - \langle\delta\rho_{\rm out}\rangle_{\omega_1,\omega_2}$. Similar *horizontal lines* in following figures are understood to have the same meaning

nucleon pole from this sum rule. Especially, in the case of the default model $m_{\mathrm{asym}}(\omega)$, the spectral function in the low energy region approaches $m_{\mathrm{asym}}(0)$ and we can only observe a small lowest peak. A similar tendency was observed in the ρ-meson channel (see Chap. 5). We will therefore abandon this default model in the following. In the case of $m_{\mathrm{flat}}(\omega)$, although the high energy behavior wrongly approaches 0 at high energy, the low energy behavior, which is the main focus of our interest, seems to be reasonable. On the other hand, using $m_{\mathrm{hybr}}(\omega)$ leads to the correct behavior at both high and low energy. From these results, we can infer that the default model completely determines the asymptotic values of the spectral function. This behavior at high and low energy should therefore be considered to be an input in the current analysis. This can be understood from the properties of the kernel $K(M, \omega)$, leading to a function $G_\rho(M)$, which is insensitive to the values of $\rho(\omega)$ at large and small values of ω. From the behavior of $K(M, \omega)$, we can also expect that increasing the upper boundary of M allows the analysis to become sensitive to $\rho(\omega)$ at higher energy regions.

For investigating this case, we analyze the spectral function under the condition of $M_{\mathrm{min}} = 0.91\,\mathrm{GeV}$ and $M_{\mathrm{max}} = 1.4\,\mathrm{GeV}$. The results are shown in Fig. 6.4. When using $m_{\mathrm{flat}}(\omega)$, the resulting spectral function at high energy oscillates around the continuum value before approaching the default model. This is plausible, because the input spectral function has the "pole + continuum" structure, which the MEM analysis is trying to simulate with the given limited information of $G_{\mathrm{mock}}(M)$. Nevertheless, the MEM procedure cannot reproduce the mass of the nucleon, whatever default model or Borel mass range is used.

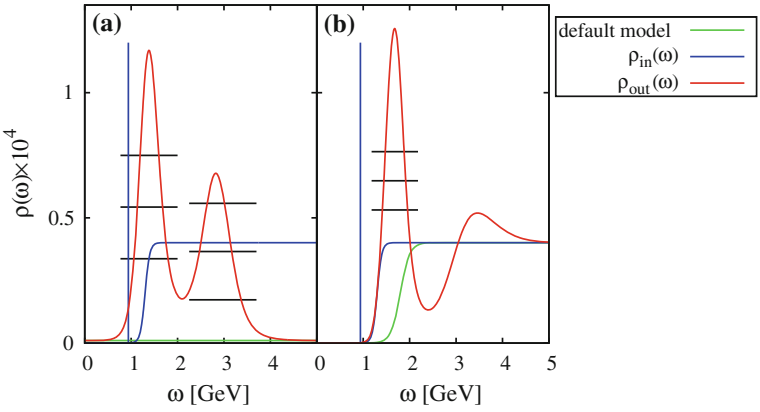

Fig. 6.4 Spectral functions extracted from $G_{\mathrm{mock}}(M)$ using various default models at $\beta = -1$. The analyzed Borel mass region is $0.91\,\mathrm{GeV} \leq M \leq 1.4\,\mathrm{GeV}$. The *red lines* show $\rho_{\mathrm{out}}(\omega)$, the *blue lines* depict the input mock spectral function and the *green lines* stand for the default model

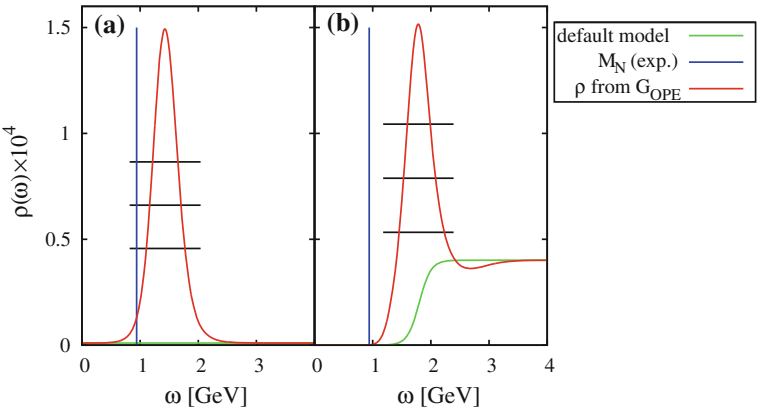

Fig. 6.5 Spectral functions extracted from $G_{\mathrm{OPE}}(M)$ using various default models at $\beta = -1$. The analyzed Borel mass region is $0.91\,\mathrm{GeV} \leq M \leq 0.97\,\mathrm{GeV}$. The *red lines* show $\rho_{\mathrm{out}}(\omega)$, the *vertical blue lines* indicate the position of the nucleon ground state and the *green lines* stand for the default model

6.3.2 Analysis Using OPE Data

Similar to the previous section, we now carry out the analysis using the real OPE data, $G_{\mathrm{OPE}}(M)$, even though from our experience of the mock data analysis, we cannot expect to obtain meaningful results. We analyze the spectral function by setting the analyzed Borel mass region to $0.91\,\mathrm{GeV} \leq M \leq 0.97\,\mathrm{GeV}$, as in the last section. The results are shown in Fig. 6.5. Using a wider Borel mass region leads to spectral functions very similar to the ones shown in Fig. 6.4. As in the mock data analysis, the MEM procedure does not succeed to reproduce the nucleon peak. The main reason for this failure can be traced back to the slow convergence of the OPE and to the large contribution of continuum to the sum rule. These factors severely reduce the information of the lowest nucleon pole that can be extracted from $G_{\mathrm{OPE}}(M)$.

6.4 Analysis Using the Gaussian Sum Rule

In case of the Gaussian sum rule, the analysis is carried out using the following equations:

$$G_{\mathrm{OPE}}(s,\tau) = G_{\rho}(s,\tau) = \int_0^{\infty} d\omega K(s,\tau,\omega)\rho(\omega), \tag{6.34}$$

$$K(s,\tau,\omega) = \frac{2\omega^5}{\sqrt{4\pi\tau}} e^{-\frac{(\omega^2-s)^2}{4\tau}}. \tag{6.35}$$

As before, we set $\beta = -1$ and the results will be shown in terms of the dimensionless spectral function.

6.4.1 Analysis Using Mock Data

As for the Borel sum rule, we use Eq. (6.27) as mock data and Eqs. (6.31–6.33) for the default model. As a first step, we must determine the ranges of s, τ used in the analysis. From the property of the kernel, we expect that G_ρ at small values of τ will be more sensitive to narrow structures such as the lowest peak, while G_ρ at larger values will to a large extent be fixed by the continuum. Hence, to extract as much information as possible from $G_{\mathrm{mock}}(s,\tau)$, we use several values of τ at the same time, which are $\tau = 0.5, 1, 1.5, 2\,\mathrm{GeV}^4$. From

$$\lim_{\tau\to 0} \frac{1}{\sqrt{4\pi\tau}} \mathrm{e}^{-\frac{(\omega^2-s)^2}{4\tau}} = \delta(\omega^2 - s), \tag{6.36}$$

one understands that in the limit $\tau \to 0$, $G_\rho(s,\tau)$ in principle should approach the spectral function as

$$\lim_{\tau\to 0} G_\rho(s,\tau) = s^2\rho(\sqrt{|s|}). \tag{6.37}$$

However, it is not possible to take this limit, because the OPE is an expansion in powers of $1/\sqrt{\tau}$ as can be seen in Eq. (6.23), meaning that the convergence of the expansion worsens with lower values of τ.

Turning now to the parameter s, similar to the Borel sum rule case, we determine its range from the convergence in $G_{\mathrm{OPE}}(s,\tau)$ and the contribution from the continuum in $G_\rho(s,\tau)$. The analyzed regions of s and τ are shown in Table 6.2. Here, we denote the upper and lower boundaries of s at each τ as s_{max} and s_{min}, respectively.

In the present analysis, we take 4 data points for $\tau (N_\tau = 4)$, and 25 data points for s ($N_s = 25$) at each τ^k and adjust ω_{min} and ω_{max} to 0 and 6 GeV, respectively. s^k_{max} and s^k_{min} at each τ^k are given in Table 6.2, and ds^k is $(s^k_{\mathrm{max}} - s^k_{\mathrm{min}})/N_s$.

A comment is in order here on the ranges of the variable s, shown in Table 6.2. It is seen that to the most part, we use negative values for s, which may seem somewhat counterintuitive as, naively, the kernel of Eq. (6.35) seems to have a peak at $\sqrt{s}$, when considered as a function of ω. Therefore one would expect values of s around 1 GeV to give a $G_\rho(s,\tau)$ that is most sensitive to the spectral function in the region

Table 6.2 The analyzed ranges of the variable s for the employed values of τ

$\tau[\mathrm{GeV}^4]$	0.5	1.0	1.5	2.0
$s_{\mathrm{min}}[\mathrm{GeV}^2]$	−0.79	−1.84	−2.96	−4.13
$s_{\mathrm{max}}[\mathrm{GeV}^2]$	0.94	0.02	−0.97	−2.00

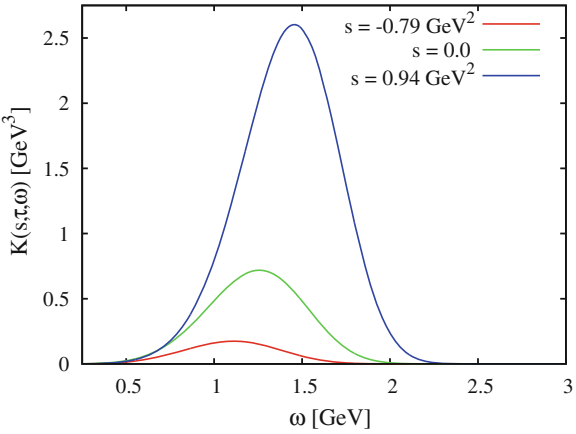

Fig. 6.6 Three examples of the kernel $K(s, \tau, \omega)$. The *red line* shows $K(s, \tau, \omega)$ for $s = -0.79\,\mathrm{GeV}^2$, the *green line* for $s = 0.0$ and the *blue line* for $s = 0.94\,\mathrm{GeV}^2$. In all three cases, τ is fixed to $\tau = 0.5\,\mathrm{GeV}^4$

of the nucleon pole. However, this is not the case because the kernel is distorted due to the ω^5-factor in front of the Gaussian of Eq. (6.35). For example, for constructing a kernel with a maximum value at 1 GeV which is the energy region which we are mostly interested in, s becomes $1\text{–}5\tau$ and, in the case of $\tau = 0.5, 1, 1.5, 2\,\mathrm{GeV}^4$, has a negative value. This is illustrated in Fig. 6.6, where one can see that only kernels with negative s have peaks around 1 GeV. From these arguments, we can understand that it is necessary and important to use negative values of s in the analysis of the Gaussian sum rules.

Let us now discuss the obtained spectral functions, which are shown in Fig. 6.7. Their detailed numerical results are given in Table 6.3. It is observed that, compared with the Borel sum rule, the reconstruction of the lowest lying peak has considerably improved. Nevertheless, it is seen in Table 6.3 that its position is shifted upwards about $120 \sim 180\,\mathrm{MeV}$, which should give an idea of the precision attainable with this method. To evaluate the residue of the first peak, we first have to define the region ω over which the peak can be considered to be the dominant contribution. The left edge of the first peak is determined to be the point whose height is 1/30 of the peak vertex and the right edge so that the center between the left and the right edge lies on the peak vertex. The residue is then obtained by simply integrating $\rho(\omega)$ over the peak region, indicated by the horizontal length of the error bars. Table 6.3 shows that the residue of the first peak gives about 80 % of the input value of the mock data. Since the default model $m_{\mathrm{flat}}(\omega)$ leads to a peak position and residue closest to the input values we will use only $m_{\mathrm{flat}}(\omega)$ as the default model in the following.

Here, we comment on the width of the peaks appearing in the obtained spectral functions. One might wonder why these peaks have a finite width even though the input mock data only contains a zero width pole. There are two important aspects related to this point. Firstly, one should note that the OPE side of the sum rule is rather

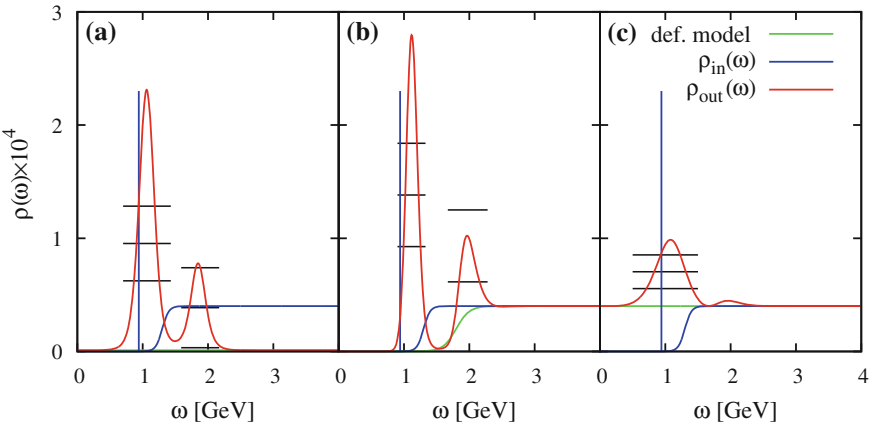

Fig. 6.7 Spectral functions extracted from $G_{\mathrm{mock}}(s,\tau)$ using various default models at $\beta = -1$. The analyzed regions of s and τ are shown in Table 6.2. The *red lines* show $\rho_{\mathrm{out}}(\omega)$, the *blue lines* depict the input mock spectral function and the *green lines* stand for the default model

Table 6.3 Position of the peaks (in MeV) and values of the corresponding residues (in 10^{-5} GeV), obtained from a mock data analysis with three default models at $\beta = -1$. Here, the given values of the mock data residue is defined as $\lambda^2/(2M_N^5)$. The errors are estimated by evaluating the half width of the obtained first peak

	Mockdata	m_{flat}	m_{hybr}	m_{asym}
Position of 1st peak [MeV]	940	1060 ± 140	1120 ± 100	1080 ± 410
Position of 2nd peak [MeV]	–	1850	1970	1960
Residue of 1st peak [10^{-5}GeV]	8.3	7.0	6.0	–

insensitive to the value of the width, because this information is largely lost after the integration over the spectral function in Eqs. (6.1) or (6.2). This is so as long as the parameters governing the scales of these integrals (M or $\tau^{1/4}$) are much larger than the width of the peak of interest. This point was discussed explicitly for the ρ-meson channel in Chap. 5 of this thesis (see also Gubler and Oka (2010)). Secondly, one should remember that MEM is a statistical method, that can only provide the most probable form of the spectral function, given the information available. This most likely spectral function depends not only on the input data but also on their error. Generally, a larger error makes a peak broader and smaller in magnitude, therefore introducing an artificial width connected to the error of the OPE data.

Next, to study the dependence of the results on the range of s, we fix s_{max} to $1.0\,\mathrm{GeV}^2$, $1.5\,\mathrm{GeV}^2$ and $2.0\,\mathrm{GeV}^2$ for all values of τ, and redo the analysis. The results are shown in Fig. 6.8 and Table 6.4. From Table 6.4, one can observe that the position and residue of the first peak depends on s_{max} only weakly, while the position of the second peak is quite sensitive to the value of s_{max}. Furthermore,

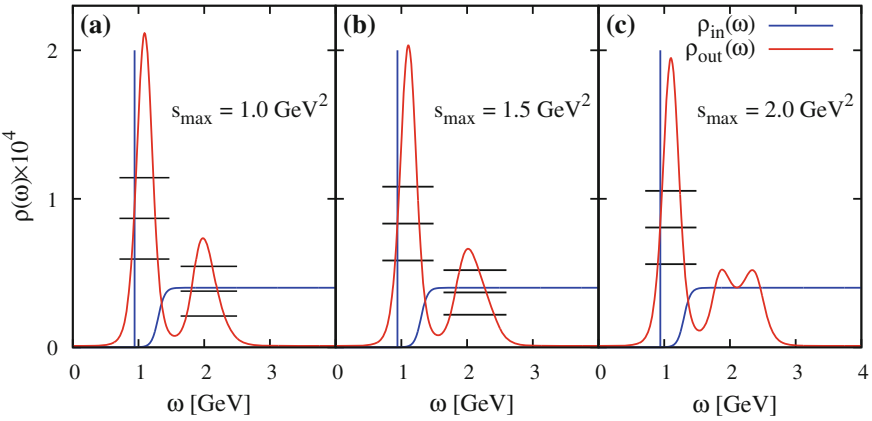

Fig. 6.8 Spectral functions extracted from $G_{\mathrm{mock}}(s,\tau)$ and $m_{\mathrm{flat}}(\omega)$ at $\beta=-1$. Various values of s_{max} are used as indicated in the figures. The values of s_{min} at different τ are shown in Table 6.2. The *red lines* show $\rho_{\mathrm{out}}(\omega)$ and the *blue lines* stand for the input mock spectral function

Table 6.4 Position of the peaks and values of the corresponding residues. Here, the default model $m_{\mathrm{flat}}(\omega)$ and various s_{max} values were used. The errors are estimated by evaluating the half width of the obtained first peak

s_{max}[GeV^2]	1.0	1.5	2.0
Position of 1st peak [MeV]	1090 ± 140	1100 ± 140	1100 ± 150
Position of 2nd peak [MeV]	1980	2020	–
Residue of 1st peak [10^{-5}GeV]	6.7	6.5	6.3

from the Fig. 6.8, it can be understood that the second peak rather represents the continuum, around which the MEM output oscillates.

Finally, we discuss the dependence of the spectral function on the parameters appearing in the default models. In the case of the flat default models as given in Eqs. (6.31) and (6.32), we have carried out analyses using several height values (spanning over several orders of magnitude). The results are shown in Fig. 6.9. It is clear that the detailed values have little influence on the position of the nucleon pole. Extracting the corresponding residues, it is understood that their values also depend on the height value only weakly. Specifically, the pole positions for the curves shown in Fig. 6.9 coincide within 1 % and the residues within 6 %. In the case of a hybrid default model as given in Eq. (6.33), we also carried out analyses using several parameter combinations, the result being that the positions of the obtained first peaks did depend on the parameter combinations only weakly.

In all, we conclude that using the Gaussian sum rule, the reconstruction of the properties of the ground state peak is largely improved compared to the Borel sum rule. This claim can be substantiated by considering the obtained pole of the nucleon, which is found at about 1.5 GeV for the Borel sum rule. This value is certainly not

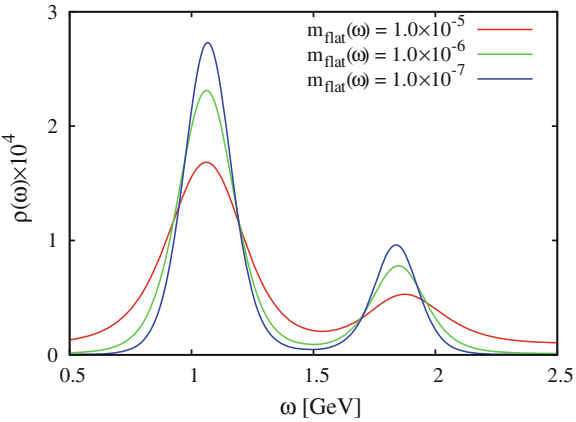

Fig. 6.9 The spectral functions extracted from $G_{\rm mock}(s,\tau)$, using various height values for the flat default model $m_{\rm flat}(\omega)$. The concrete numerical values are given in the figure

Table 6.5 Positions of the first and second peaks (in MeV) and the corresponding residue of the first peak (in 10^{-5} GeV) obtained from the analyses using OPE data at $\beta = -1$ and the default model $m_{\rm flat}(\omega)$. The errors are estimated by evaluating the half width of the obtained first peak

$s_{\rm max}$[GeV^2]	cont.	1.0	1.5	2.0
Position of 1st peak [MeV]	990 ± 130	1040 ± 140	1060 ± 150	1050 ± 150
Position of 2nd peak [MeV]	1840	1920	–	1800
Residue of 1st peak [10^{-5}GeV]	7.9	7.0	6.6	6.3

satisfactory in view of the precision of QCD sum rules, which is expected to be about $10 \sim 20\,\%$. On the other hand, in the case of the Gaussian sum rule, the position of the reconstructed nucleon pole agrees with the mock data within the range of $20\,\%$. Hence, the Gaussian sum rule is more suitable for the analysis of the nucleon channel, especially for the present MEM analysis, in which only Π_1 can be used.

6.4.2 Analysis Using OPE Data

Next, we carry out the analysis using $G_{\rm OPE}(s,\tau)$. We apply the MEM in the same regions of s and τ used in the mock data analysis. The results are shown in Fig. 6.10, while the numerical details are given in Table 6.5. The behavior of the results is similar to those of the mock data analysis. We observe that the positions of the lowest peak lie quite close to the experimental value. Besides a clearly resolved first peak, the spectral function exhibits one or two higher peaks which oscillate around the asymptotic high energy limit. In principle, this part could also contain nucleon resonances with both positive and negative parity. With the present resolution achievable with the MEM

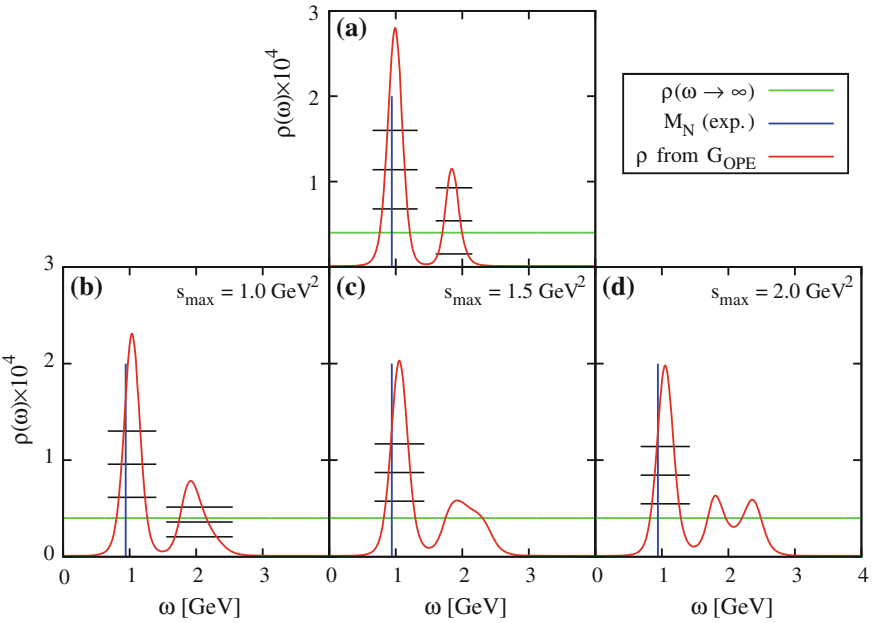

Fig. 6.10 Spectral functions extracted from the MEM analysis using $G_{\mathrm{OPE}}(s,\tau)$ and $m_{\mathrm{flat}}(\omega)$ at $\beta=-1$. The *red lines* show $\rho(\omega)$ and the *green lines* stands for the asymptotic value of the spectral function at high energy. (Note that $\rho(\omega)$ is normalized by $\frac{1}{\omega^4}$.) In the top figure (**a**), the values of s and τ given in Table 6.2 are used, while for the other figures (**b**, **c** and **d**) the values of s_{max} are fixed as indicated. The *vertical blue lines* are placed at the positions of the experimental mass value of the nucleon ground state

technique, it however seems to be rather difficult to resolve these resonances from the continuum. Using the parity projected sum rules (Jido et al. 1996), which can also be analyzed by the MEM approach, might improve the situation.

6.4.3 Investigation of the β Dependence

The coupling strengths of the nucleon ground state and its excited states depend on the choice of the interpolating operator, i.e., on the value of β. To investigate the nature of this dependence, we let β vary as $-1.5, -0.5, 0.0, 0.5, 1.0, 1.5, \infty$ and extract the corresponding spectral functions. For τ, we have chosen 0.5, 1.0, 1.5, 2.0 GeV^4 as before and s_{min} is determined from the OPE convergence condition at each β. In the case of $\beta=1$, s_{min} is determined from the dimension 4 term because the higher dimensional terms vanish for this choice. To obtain information not only on the spectral function around 1 GeV, but also in the region of possible excited states,

Table 6.6 Values of $s_{\min}$ [GeV^2] at each β and τ [GeV^4]. $s_{\max}$ is fixed to 2.0 GeV^2

β	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5	∞
$\tau = 0.5\,\text{GeV}^4$	-0.40	-0.79	-0.39	-0.35	-0.43	0.92	-0.23	1.04
$\tau = 1.0\,\text{GeV}^4$	-1.21	-1.84	-0.96	-0.87	-1.10	0.12	-0.93	-0.36
$\tau = 1.5\,\text{GeV}^4$	-2.12	-2.96	-1.63	-1.47	-1.87	-0.76	-1.74	-0.88
$\tau = 2.0\,\text{GeV}^4$	-3.07	-4.13	-2.34	-2.13	-2.68	-1.66	-2.60	-1.45

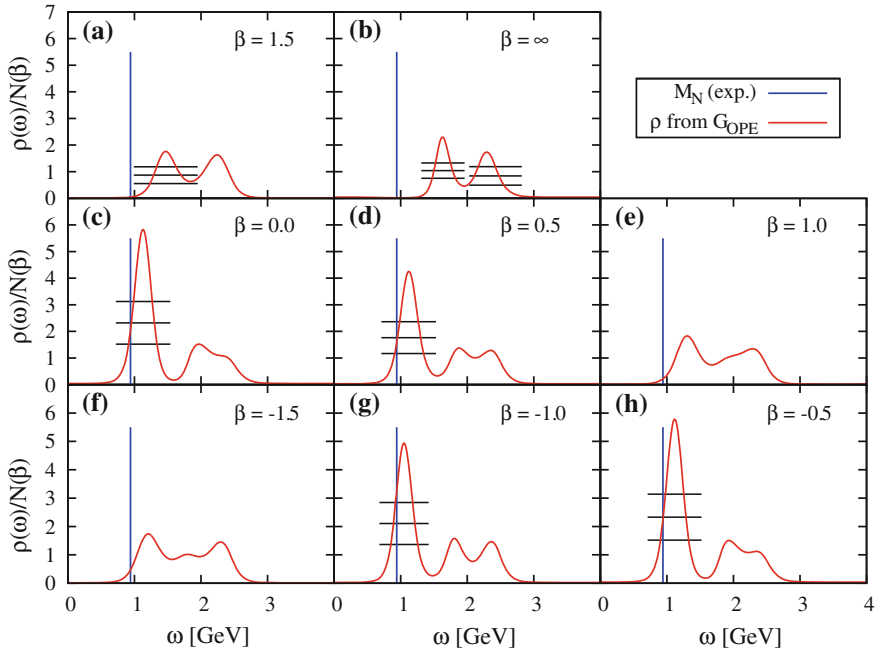

Fig. 6.11 Spectral functions (*red lines*) obtained from the MEM analyses of OPE data for various values of β, normalized by the factor $N(\beta)$. The values $s_{\min}$ are given in Table 6.6 and $s_{\max}$ is fixed to 2.0 GeV^2. The *vertical blue lines* are placed at the positions of the experimental mass value of the nucleon ground state

$s_{\max}$ is not determined from the dominance of the lowest lying state in $G_\rho(s, \tau)$ but is fixed to 2.0 GeV^2. The explicit values of $s_{\min}$ are given in Table 6.6.

Note that $\beta = \infty$ should not be taken literally, but just means that we use the correlator of only $\eta_2(x)$ (which is the coefficient of β^2 in Eq. (6.7)) for the analysis. The resulting spectral functions and their numerical properties are shown in Fig. 6.11 and Table 6.7. For a better comparison, we have normalized the spectral functions in Fig. 6.11 by dividing them by the factor $N(\beta) \equiv (5 + 2\beta + 5\beta^2)/(128(2\pi)^4)$, so that the continuum approaches unity in the high energy limit.

Considering first the lowest peaks at $\omega \sim 1$ GeV, we see that a peak is clearly resolved for $\beta = -1.0, -0.5, 0.0$ and 0.5, located at about the same position and

Table 6.7 Positions of the first peaks (in MeV) of the spectral functions shown in Fig. 6.11. For the cases in which the first peak can be interpreted as the nucleon ground state, the corresponding residue (in 10^{-5} GeV) is given as well

β	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5	∞
Position of 1st peak [MeV]	1210	1050	1120	1130	1120	1300	1470	1630
Residue of 1st peak [10^{-5}GeV]	–	6.3	5.0	4.8	5.3	–	–	–

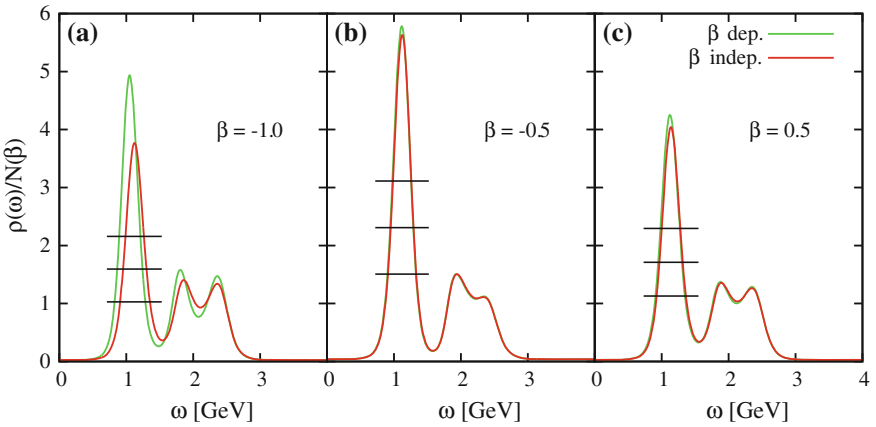

Fig. 6.12 Spectral functions obtained from a MEM analysis of OPE data with $m_{\text{flat}}(\omega)$ at $\beta = -1, -0.5$ and 0.5, normalized by $N(\beta)$. The *red lines* show the normalized output spectral function $\rho(\omega)$, for which a β independent "refined" criterion was used to determine $s_{\min}$. The details are explained in the text. For comparison, the spectral functions obtained with the previous β-dependent criterion are indicated by the *green lines*

with similar residue values. For the other values of β, no prominent peak is observed. These results are, however, obtained by using the $s_{\min}$ values determined separately for each β and one could suspect that the choice of $s_{\min}$ affects the properties of the first peak. To study this dependence and to get an idea of the stability of our results, we redo the analysis, now using values of $s_{\min}$ that are determined via a β independent criterion as advocated in Leinweber (1997, 1995). Taking a closer look at Table 6.6, it is observed that in the region between $\beta = -1.0$ and 0.5, where the ground state peak can be extracted, the OPE convergence is worst for $\beta = 0.0$, giving the largest value of $s_{\min}$. This implies that it is most reasonable to fix $s_{\min}$ to the values of $\beta = 0.0$, so that no values of s are used, where the convergence of the OPE might be questionable. Using this criterion, we repeat the analysis at $\beta = -1, -0.5, 0.5$. As before, we set $s_{\max}$ to $2.0\,\text{GeV}^2$. The results are shown and compared to the previous ones in Fig. 6.12. The numerical details are given in Table 6.8. We see from these results that some details of the output spectral functions (especially at $\beta = -1$) change when the new criterion is used. The qualitative structure of a clear lowest peak with a continuum structure at a somewhat higher energy is however not altered.

Table 6.8 Positions of the first peaks and the corresponding residues of the spectral functions shown in Fig. 6.12, where a β-independent criterion was used to fix $s_{\min}$. For comparison, the result for $\beta = 0$ is shown as well

β	-1.0	-0.5	0.0	0.5
Position of 1st peak [MeV]	1120	1120	1130	1140
Residue of 1st peak [10^{-5}GeV]	5.3	4.9	4.8	5.1

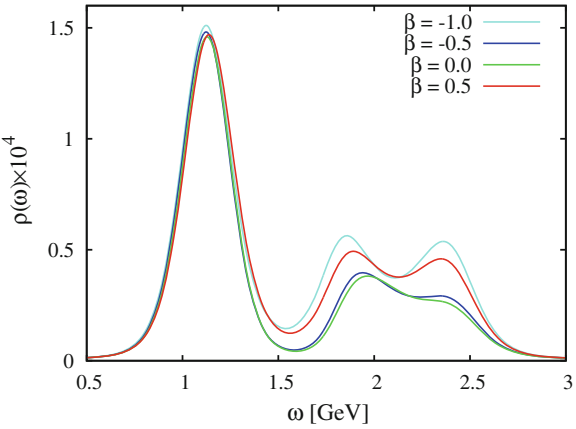

Fig. 6.13 Non-normalized spectral functions obtained from a MEM analysis using OPE data and $m_{\text{flat}}(\omega)$ at $\beta = 0.5, 0, -0.5$ and -1.0. The employed $s_{\min}$ is fixed to the one at $\beta = 0$ and $s_{\max}$ to $2.0\,\text{GeV}^2$

We also observe from Table 6.8 that the position and residue are almost independent of β. This fact becomes even more explicit when we plot the non-normalized spectral functions for several values of β around $\beta = 0$ as shown in Fig. 6.13. This figure clearly illustrates that, in contrast to the continuum spectra, the property of the lowest peak is essentially β-independent.

All these results can be naturally interpreted by presuming that the lowest lying pole couples strongly to η_1, and only very weakly to η_2. Therefore, the pole appears to be almost independent of β as long as $|\beta|$ is small, but disappears when $|\beta|$ becomes large and the contribution of η_2 dominates the spectral function. The nucleon ground state pole can be resolved as long as its strength is large enough compared to the continuum contribution. At about $|\beta| \geq 1.5$, however, this continuum contribution gets too large to extract information on the lowest peak. These conclusions agree with the findings of lattice QCD (Leinweber 1995; Sasaki et al. 2002; Melnitchouk 2003) and some earlier suggestions of QCD sum rule studies (Chung et al. 1984), in which however the experimental mass value of the nucleon ground state peak was used as an input.

Furthermore, by looking at Table 6.6 and Fig. 6.12, we can also understand why it is the Ioffe current ($\beta = -1.0$), rather than η_1, that seems to be most suitable

for QCD sum rule studies. This comes from the fact that the OPE convergence is considerably faster for the Ioffe compared to η_1, which allows the analysis to become more sensitive to the lower energy region of the spectral function, therefore providing more information on the nucleon peak.

As a last point, let us also consider the continuous structure above the nucleon ground state pole and possible excited states appearing there. It can be observed in Fig. 6.11 that for most values of β, the spectral function just oscillates around the high energy limit, similar to the results obtained from the mock data analysis shown in Fig. 6.8, where we have only included the continuum into the input spectral function without any resonances. On the other hand, for $\beta = 1.5$ and especially $\beta = \infty$, the behavior is somewhat different, showing a quite clear first peak at about 1.6 GeV. This peak could correspond to one of the nucleon resonances, N(1440), N(1535), N(1650) or a combination of them. Results of lattice QCD studies (Sasaki et al. 2002; Melnitchouk 2003) indicate that η_2 couples to the negative parity states N(1535) and N(1650), while the first excited positive parity state lies considerably above the Roper resonance N(1440). Even though the systematic uncertainties of these calculations are probably not yet completely under control, these findings suggest that the lowest peak seen at $\beta = \infty$ corresponds rather to N(1535) or N(1650) and not to N(1440). It, however, seems to be difficult at the present stage to make any conclusive statements on the nature of this peak. An analysis of the parity projected sum rules will hopefully help to clarify this issue.

6.5 Summary and Conclusion

In this chapter, we have applied the MEM technique, which is based on Bayes' theorem of probability theory to the analysis of the nucleon QCD sum rules. We have investigated two kinds of sum rules, namely the frequently used Borel sum rule and the less known Gaussian sum rule. Before analyzing the actual sum rules, we have first tested the applicability of the MEM approach by constructing and analyzing realistic mock data. Our findings show that due to the properties of the kernel and the slow OPE convergence, it appears to be difficult to extract much meaningful information on the nucleon ground state from the Borel sum rule. Another reason for this failure may also be that, because only spectral functions satisfying positivity can be analyzed with the currently available MEM procedure, we cannot use the chiral odd part $\Pi_2(q^2)$ of Eq. (6.6) in the present analysis, which has been claimed to be more reliable. For instance, in Leinweber (1997) analyses using only $\Pi_1(q^2)$ or $\Pi_2(q^2)$ were carried out and the respective Borel windows examined. As a result, in case of only using $\Pi_1(q^2)$, the Borel window is seen to be very narrow, making it difficult to obtain a reliable estimate for the nucleon mass. On the other hand, when using only $\Pi_2(q^2)$, the Borel window was shown to be sufficiently large, so that the nucleon mass can be reliably obtained. It therefore would be helpful if $\Pi_2(q^2)$ could be used. As long as one uses MEM, this, however, will only become possible when the parity projected sum rules are employed. As an alternative, we have formulated

the Gaussian sum rule and found that it allows us to extract more information on the spectral function and enables us to reconstruct the nucleon ground state with reasonable precision from both the mock and the OPE data.

As the analysis is done with the MEM technique, we obtain the spectral function directly and do not have to deal with quantities depending on unphysical parameters such as the Borel mass. Moreover, we do not have to restrict the spectral function to the traditional "pole + continuum" form, which allows us to investigate the spectral function of a large variety of interpolating fields. From this investigation, we have found that the nucleon pole is independent of the parameter β of Eq. (6.1) and vanishes when η_2 becomes the dominant contribution of the correlator. Thus we conclude that, in agreement with findings of lattice QCD, the nucleon ground state couples only to the interpolating field η_1, but not (or only very weakly) to η_2. Furthermore, a peak structure is seen around 1.6 GeV in the spectral function corresponding to η_2, which suggests that some nucleon resonance in this region couples to η_2. To clarify the nature of this peak, more thorough investigations are needed.

In all, we have shown that the MEM technique in combination with the Gaussian sum rule formulated in this paper is useful for extracting the properties of the nucleon ground state and may even make it possible to investigate possible excited states. There are, however, still several open questions to be answered. First of all, we have so far ignored all radiative α_s corrections to the Wilson coefficients. These corrections are known to be significant and it is therefore important to include them for a more quantitative analysis. Additionally, possible violations of the vacuum saturation approximation should also be considered. As a further point, it would be crucial to separate the contributions from positive and negative parity states to the spectral function, especially to investigate the excited nucleon resonances. These issues are left for future investigations.

References

Adami C, Zahed I (1992) Phys Rev D 45:4312
Belyaev VM, Ioffe BL (1982) Sov Phys JETP 56:493
Bertlmann RA, Launer G, de Rafael E (1985) Nucl Phys B250:61
Chung Y, Dosch HG, Kremer M, Schall D (1982a) Nucl Phys B 197:55
Chung Y, Dosch HG, Kremer M, Schall D (1982b) Z Phys C 15:367
Chung Y, Dosch HG, Kremer M, Schall D (1984) Z Phys C 25:151
Cohen TD, Furnstahl RJ, Griegel DK, Jin X (1995) Prog Part Nucl Phys 35:221
Colangelo P, Khodjamirian A (2001) At the frontier of particle physics/handbook of QCD, vol 3. World Scientific, Singapore, p 1495
Dosch HG, Jamin M, Narison S (1989) Phys Lett B 220:251
Drukarev EG, Ryskin MG, Sadovnikova VA (2010) arXiv:1012.0394 [nucl-th].
Efros VD, Leidemann W, Orlandini G (1994) Phys Lett B338:130
Efros VD, Leidemann W, Orlandini G, Barnea N (2007) J Phys G Nucl Part Phys 34:R459
Gubler P, Oka M (2010) Prog Theoret Phys 125:995
Hatsuda T, Hogaasen H, Prakash M (1991) Phys Rev Lett 66:2851
Ioffe BL (1981a) Nucl Phys B188:317

Ioffe BL (1981b) Nucl Phys B191, 591 (E).
Ioffe BL (1983) Z Phys C 18:67
Jamin M (1988) Z Phys C 37:635
Jido D, Kodama N, Oka M (1996) Phys Rev D 54:4532
Krasnikov NV, Pivovarov AA, Tavkhelidze NN (1983) Z Phys C 19:301
Leinweber DB (1990) Ann Phys (NY) 198:203
Leinweber DB (1995) Phys Rev D 51:6383
Leinweber DB (1997) Ann Phys 254:328
Melnitchouk W et al (2003) Phys Rev D 67:114506
Ohtani K, Gubler P, Oka M (2011) Eur Phys J A 47:114
Orlandini G, Steele TG, Harnett D (2001) Nucl Phys A686:261
Ovchinnikov AA, Pivovarov AA, Surguladze LR (1991) Int J Mod Phys A 6:2025
Sadovnikova VA, Drukarev EG, Ryskin MG (2005) Phys Rev D 72:114015
Sasaki S, Blum T, Ohta S (2002) Phys Rev D 65:074503
Shifman MA, Vainshtein AI, Zakharov VI (1979a) Nucl Phys B147:385
Shifman MA, Vainshtein AI, Zakharov VI (1979b) Nucl Phys B147:448
Shiomi H, Hatsuda T (1995) Nucl Phys A594:294
Yang KC, Hwang WYP, Henley EM, Kisslinger LS (1993) Phys Rev D 47:3001

Chapter 7
Quarkonium Spectra at Finite Temperature from QCD Sum Rules and MEM

7.1 Introduction

Since QCD was established to be the theory of strong interactions, quarkonium has often been used as a suitable probe of its dynamics, owing to the fact that in this system both perturbative and non-perturbative aspects of QCD play equally important roles (Novikov et al. 1978).

Especially, the behavior of charmonia in a hot or dense medium has attracted much interest, as it was suggested some time ago, that in the color-deconfined medium with a temperature above T_c charmonia will dissolve due to the color Debye screening, and thus serve as a signal for the formation of quark-gluon plasma (Matsui and Satz 1986). Testing these early suggestions from first principles of QCD has become feasible only recently, as new developments in lattice QCD have made it possible to access the charmonium spectral functions with the help of the maximum entropy method (MEM) (Asakawa and Hatsuda 2004; Datta et al. 2004; Umeda 2004; Jacovác et al. 2007). These studies found that the lowest charmonium states (J/ψ and η_c) survive up to temperatures as high as $\sim 1.5\, T_c$ or even higher.

On the other hand, bottomonium is expected to persist as a bound state up to higher temperatures, as it is a smaller system and will thus be affected by the Debye screening only if the temperature reaches values of a few T_c. Furthermore, bottomonia spectral functions have, in contrast to their charmonium counterparts, several excited states lying below the continuum threshold (for instance, $\Upsilon(2S)$ and $\Upsilon(3S)$ in the vector channel), which add novel and interesting features into the analysis, but also make an accurate extraction of the spectral functions more difficult, as we will discuss in the following sections. The behavior of these excited states is presently attracting considerable interest, as recent heavy-ion collision experiments at the LHC at CERN have found a significant reduction of the relative yield of $\Upsilon(2S)$ and $\Upsilon(3S)$ compared to $\Upsilon(1S)$, when switching from pp to Pb-Pb collisions, which is strong evidence for the disappearance the excited states at some lower temperature than the ground state (CMS Collaobation 2011, 2012). Also, first dedicated lattice calculations of

P. Gubler, *A Bayesian Analysis of QCD Sum Rules*, Springer Theses,
DOI: 10.1007/978-4-431-54318-3_7, © Springer Japan 2013

bottomonium states, which employ the NRQCD Lagrangian as their starting point, have recently appeared (Aarts et al. 2011a,b).

Besides lattice QCD, the method of QCD sum rules (Shifman et al. 1979a,b) provides another tool for investigating the properties of hadrons at finite temperature (Bochkarev and Shaposhnikov 1986; Hatsuda et al. 1993). Using this approach various charmonium and bottomonium channels were studied recently (Morita and Lee 2008a,b; Song et al. 2009; Morita and Lee 2010), and evidence for a considerable change of the spectral functions just above T_c (in the charmonium case) was found. To specify the nature of this change is the major goal of this chapter (see also Gubler et al. 2011; Suzuki et al. 2013). For this task we employ MEM, which, as we have shown in this thesis, is applicable to QCD sum rules and has the advantage that one does not have to introduce any strong assumption about the functional form of the spectral function.

7.2 Formalism

Let us first recapitulate what sort of information QCD sum rules can provide on the quarkonium spectral function at finite temperature (Bochkarev and Shaposhnikov 1986; Hatsuda et al. 1993). One considers the time-ordered correlator at finite temperature

$$\Pi^{\mathrm{J}}(q) = i \int d^4x e^{iqx} \langle T[j^{\mathrm{J}}(x) j^{\mathrm{J}}(0)] \rangle_T, \tag{7.1}$$

where $j^{\mathrm{J}}(x)$ stands for $\bar{h}\gamma_\mu h(x)$ and $\bar{h}\gamma_5 h(x)$ in the vector (V) and pseudoscalar (PS) channel, and for $\bar{h}h(x)$ and $(q_\mu q_\nu/q^2 - g_{\mu\nu})\bar{h}\gamma_5\gamma^\nu h(x)$ in the scalar (S) and axial-vector (AV) channel, respectively. Here, the operator h stands either for a charm or bottom quark. The expectation value $\langle \mathcal{O} \rangle_T$ is defined as $\langle \mathcal{O} \rangle_T \equiv \mathrm{Tr}(e^{-H/T}\mathcal{O})/\mathrm{Tr}(e^{-H/T})$. Throughout this work, we will set the spatial momentum of the quarkonium system relative to the thermal medium to be $\mathbf{0}$; thus, $q^\mu = (\omega, \mathbf{0})$. In this circumstance, there is only one independent component in the correlators of the vector and axial-vector channel. In what follows, we will use the dimensionless functions $\tilde{\Pi}^{\mathrm{V,AV}}(q^2) \equiv \Pi_\mu^{\mu,\mathrm{V,AV}}(q)/(-3q^2)$ and $\tilde{\Pi}^{\mathrm{PS,S}}(q^2) \equiv \Pi^{\mathrm{PS,S}}(q)/q^2$ for the analysis.

7.2.1 Formulation of the Sum Rule

Going to the deep Euclidean region $q^2 \equiv -Q^2 \ll 0$, one can calculate the correlation functions using the operator product expansion (OPE), giving an expansion in local operators O_n with increasing mass dimension n: $\tilde{\Pi}^{\mathrm{J}}(q^2) = \sum_n C_n^{\mathrm{J}}(q^2)\langle O_n \rangle_T$. As was first discussed in Hatsuda et al. (1993), as long as the temperature T lies below the separation scale of the OPE, which is of the order of $\sim m_h$, all the temperature

effects can be included into the expectation values of the local operators $\langle O_n \rangle_T$, while the Wilson coefficients $C_n^{\mathrm{J}}(q^2)$ are independent of T. Furthermore, to improve the convergence of the OPE and suppressing the influence of high energy states onto the sum rule, we apply the Borel transform to the correlator, leading to the final result of the OPE for $\nu \equiv 4m_c^2/M^2$, M being the Borel mass:

$$\begin{aligned}\mathscr{M}^{\mathrm{J}}(\nu) = e^{-\nu} A^{\mathrm{J}}(\nu)[1 + \alpha_s(\nu) a^{\mathrm{J}}(\nu) + b^{\mathrm{J}}(\nu)\phi_b(T) \\ + c^{\mathrm{J}}(\nu)\phi_c(T) + d^{\mathrm{J}}(\nu)\phi_d(T)].\end{aligned} \tag{7.2}$$

The first two terms in Eq. (7.2) are the leading order perturbative term and its first order α_s correction. The third and fourth terms contain the scalar and twist-2 gluon condensates of mass dimension 4:

$$\begin{aligned}\phi_b(T) &= \frac{4\pi^2}{9(4m_h^2)^2} G_0, \\ \phi_c(T) &= \frac{4\pi^2}{3(4m_h^2)^2} G_2,\end{aligned} \tag{7.3}$$

where G_0 stands for

$$G_0 = \left\langle \frac{\alpha_s}{\pi} G^a_{\mu\nu} G^{a\mu\nu} \right\rangle_T, \tag{7.4}$$

and G_2 is defined as

$$\left\langle \frac{\alpha_s}{\pi} G^{a\mu\sigma} G^{a\nu}_{\sigma} \right\rangle_T = \left(u^\mu u^\nu - \frac{1}{4} g^{\mu\nu} \right) G_2, \tag{7.5}$$

u^μ being the four velocity of the medium. For the detailed expressions of the Wilson coefficients of these terms, see Morita and Lee (2010). To evaluate the possible influence of higher order contributions, we include one more term, which is proportional to the scalar gluon condensate of dimension 6,

$$\phi_d(T) = \frac{1}{(4m_h^2)^3} \left\langle g^3 f^{abc} G^{a\nu}_{\mu} G^{b\lambda}_{\nu} G^{c\mu}_{\lambda} \right\rangle_T. \tag{7.6}$$

The Wilson coefficient of this term can be found in Marrow et al. (1987). Note that this term is highly suppressed in the bottom case, because of the large bottom quark mass. We can therefore safely neglect it in our investigation of bottomonium states and only include it for the charmonium sum rules.

The correlator can also be expressed by a dispersion relation, in terms of the spectral function $\rho^{\mathrm{J}}(\omega)$ of the channel specified by the operator $j^{\mathrm{J}}(x)$. After the Borel transform one obtains

$$\mathscr{M}^{\mathrm{J}}(\nu) = \int_0^\infty dx^2 e^{-x^2\nu} \rho^{\mathrm{J}}(2m_h x). \tag{7.7}$$

Equating Eqs. (7.2) and (7.7) then gives the final form of the sum rules. In the vector, axial-vector and scalar channel, an additional constant term contributes to Eq. (7.2), which originates from a pole at $\omega = 0$ in $\rho^{\mathrm{V}}(\omega)$ (Bochkarev and Shaposhnikov 1986). As this so-called scattering term considerably complicates the analysis, we eliminate it by taking the derivative of Eqs. (7.2) and (7.7) with respect to ν and analyze only the resulting derivative sum rule in these channels. For a discussion on the validity of this procedure in the heavy quark sum rules, see Morita and Lee (2012).

The usual strategy of analyzing QCD sum rules is to make some reasonable assumptions on the functional form of the spectral function, and then extract information on the lowest lying peak from Eqs. (7.2) and (7.7). This method, however, has several shortcomings. First of all, the widely used "pole + continuum" ansatz, which certainly works well at $T = 0$, may not be appropriate at temperatures above T_c, where the lowest lying state is expected to be modified and eventually melt into the c-$\bar{c}$ continuum, which could become the dominant contribution. Furthermore, it is not always possible to unambiguously fit a specific ansatz to the OPE results, because of the occurrence of equally valid solutions. Such a situation arose in Morita and Lee (2008a, 2010), where it was not possible to determine a unique solution for the used parametrization of the spectral function. To handle these problems, we propose to use MEM, which allows us to extract the spectral function from Eqs. (7.2) and (7.7) without prejudice on its functional form. Moreover, as it is shown in Appendix H, it can be proven that this method provides a unique solution for the spectral function.

7.2.2 The Temperature Dependence of the Condensates

Let us now describe how the temperature dependencies of the gluonic condensates are determined. For the scalar and twist-2 gluon condensates with mass dimension 4, we follow the approach proposed in Morita and Lee (2008a, 2010), where, in the quenched approximation, the energy momentum tensor, expressed using gluonic operators, was matched with the corresponding quantity written down in form of the energy density ε and the pressure p, leading to

$$\begin{aligned} G_0 &= G_0^{\mathrm{vac}} - \frac{8}{11}\big[\varepsilon(T) - 3p(T)\big], \\ G_2 &= -\frac{\alpha_s(T)}{\pi}\big[\varepsilon(T) + p(T)\big], \end{aligned} \tag{7.8}$$

for the scalar and twist-2 gluon condensates. The functions $\varepsilon(T)$, $p(T)$ and $\alpha_s(T)$ were then extracted from quenched lattice QCD data (Boyd 1996; Kaczmarek 2004). We will in this study use the same numerical values for the T dependent part of G_0 and G_2 as in Morita and Lee (2010). As is shown there, both G_0 and G_2 exhibit a sudden decrease in the vicinity of T_c.

Let us here make a brief comment on the meaning of the deconfinement temperature T_c. As we here use lattice data in the quenched approximation (which is consistent with the OPE, in which at the present order no light quarks appear), T_c takes a value of about 260 MeV (Boyd 1996), which is considerably larger than the corresponding value of full QCD, which is presently estimated to be in the region of $150 \sim 170$ MeV (Fukushima and Hatsuda 2011). Therefore, as long as we work in the quenched approximation (as we do in this work), T_c should have the corresponding value mentioned above. However, if one wants to compare our results with actual experiments, one should adapt the smaller value obtained in full QCD. When doing such a rescaling, one of course has to assume, that all quantities involving the temperature behave the same as a function of T/T_c. For the scalar gluon condensate, this is known to be approximately true (Miller 2007), while for the twist-2 gluon condensate it has not been checked yet. To make the findings of our method more robust, this point should certainly be studied in more detail in future investigations.

It was suggested in previous studies that the OPE of the charmonium sum rules could break down at temperatures above T_c as higher dimensional operators may become non-negligible (Morita and Lee 2010). To investigate this possibility, we include the scalar gluonic condensate with dimension 6, $\langle g^3 f^{abc} G_\mu^{a\nu} G_\nu^{b\lambda} G_\lambda^{c\mu} \rangle$, about which much less is known. To our knowledge, at $T = 0$, there exists only an estimate based on the dilute instanton gas model, giving

$$\langle g^3 f^{abc} G_\mu^{a\nu} G_\nu^{b\lambda} G_\lambda^{c\mu} \rangle = \frac{48\pi^2}{5\rho_c^2} \left\langle \frac{\alpha_s}{\pi} G^2 \right\rangle, \tag{7.9}$$

where ρ_c is a representative value for the instanton radius, for which we use the established value of 0.33 fm (Schafer and Shuryak 1998). Assuming that the relation above also holds at finite temperature, and taking into account the reduction of ρ_c above T_c (Chu and Schramm 1995), we, however, found that the dimension 6 term does not influence the behavior of the spectral function much in the temperature region investigated in this chapter. Therefore, we conclude that even though the relation obtained from the dilute instanton gas model can only be considered to be a crude estimate, as long as it gives the correct order of magnitude, the contribution of the dimension 6 condensate is small and does not alter the results obtained in this study.

7.3 Results of the MEM Analysis for Charmonium

7.3.1 Mock Data Analysis

Before presenting our MEM analysis results of the OPE data, we first show in this subsection a test analysis using mock data and realistic errors. This will firstly give an idea on the resolution achievable by the MEM analysis of this channel. Secondly,

and most importantly, it will also show whether our method is able to reproduce the potentially occurring melting process of the lowest charmonium peak. Here, we work entirely in the vector channel, for which detailed experimental information on the vacuum spectral function is available.

First, we have to specify the spectral function used for the construction of mock data. We include the three lowest charmonium peaks and a smooth continuum, which starts to appear in the region of the $D\bar{D}$ threshold and approaches the perturbatively calculable limit at high energy. Specificly, this gives

$$\rho_{\mathrm{mock}}(\omega) = \sum_{i=1}^{3} \frac{27}{32\pi\alpha^2} \Gamma_i(e^+e^-)\delta(\omega - m_i) + \frac{1}{1+e^{(\omega_0-\omega)/\delta}} \frac{1}{\pi} \mathrm{Im}\Pi^{\mathrm{V}}_{\mathrm{OPE}}(\omega^2_{\mathrm{max}}), \tag{7.10}$$

where in the first term on the right hand side we made use of the fact that in the vector channel, the strength of some specific peak is determined from its leptonic width. The values of the peak masses m_i and their leptonic widths $\Gamma_i(e^+e^-)$ can be found in the PDG (Nakamura 2010). Furthermore, for the parameters ω_0 and δ, which parametrize the emergence of the continuum, we use $\omega_0 = 4.34\,\mathrm{GeV}$ and $\delta = 0.2\,\mathrm{GeV}$.

The result of the MEM analysis of mock data constructed from Eq. (7.10) is shown in Fig. 7.1. For this analysis, we use exactly the same error that will be used in the investigation of the OPE data of the vector channel at $T = 0$. We also use the

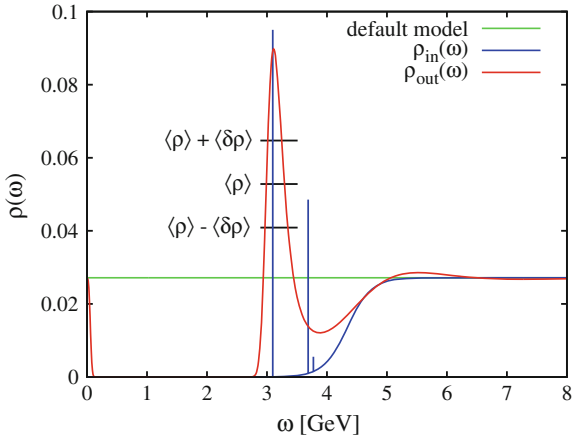

Fig. 7.1 The spectral function (*red line*) obtained from an MEM analysis of mock data constructed from the input spectral function of Eq. (7.10) (*blue line*). The *green line* shows the default model used in the MEM analysis. As explained in Fig. 4.2 of Chap. 4, the *three horizontal lines* correspond to the values of the spectral function, averaged over the peak $\langle\rho_{\mathrm{out}}\rangle_{\omega_1,\omega_2}$ and the respective upper and lower values, $\langle\rho_{\mathrm{out}}\rangle_{\omega_1,\omega_2} + \langle\delta\rho_{\mathrm{out}}\rangle_{\omega_1,\omega_2}$ and $\langle\rho_{\mathrm{out}}\rangle_{\omega_1,\omega_2} - \langle\delta\rho_{\mathrm{out}}\rangle_{\omega_1,\omega_2}$. Similar *horizontal lines* in following figures are understood to have the same meaning

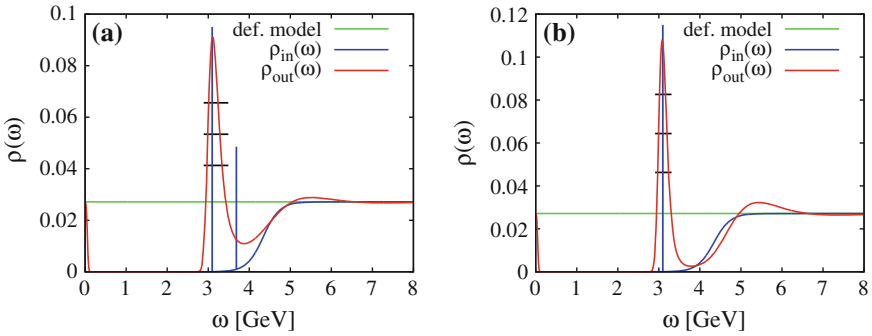

Fig. 7.2 Same as in Fig. 7.1, but with second (*left figure*) or first and second (*right figure*) excited states removed from the input spectral function

same range of ν, for which the details are specified in the next section. It can be seen in the figure that, the position of the lowest peak and the continuum are well reproduced, while the first and second excited states are completely washed out, so that it is impossible to make any sensible statement about their existence from the MEM result. Numerically, the position of the lowest peak is obtained as 3.11 GeV, which should be compared with the corresponding input value of 3.097 GeV.

To get an idea on the effect of the excited states on the extracted spectral function, we repeat the above analysis, but now remove one or both excited states from the input spectral function. The result of the corresponding analysis is given in Fig. 7.2. One can observe from these results, that the second excited state has only very little influence on the obtained spectral function, as the analysis results of Fig. 7.1 and the left plot in Fig. 7.2 almost completely coincide, the peak position still being at 3.11 GeV. On the other hand, the first excited state is responsible for the tail on the high energy side of the lowest peak in Fig. 7.1, which partly vanishes in the right plot of Fig. 7.2. Additionally, the peak position slightly moves downwards to 3.08 GeV in this case. In all, we however see that the position and existence of the lowest peak is quite stable against changes of the excited states and therefore reflects the properties of the ground state with good accuracy.

Next, we also need to confirm if MEM is able to reproduce a spectral function with no peaks and only a smooth continuum. This check is important, as the goal of this study is to determine whether a charmonium peak at a specific temperature has disappeared from the spectral function or not, and to do this, the MEM technique has to be able to reproduce both cases (with or without peaks). To this end, we employ for the input spectral function the simple case of a continuum of non-interacting quarks, which can be written down as

$$\rho_{\mathrm{mock}}(\omega) = \frac{\theta(\omega - 2m_c)}{4\pi^2}\sqrt{1 - \frac{4m_c^2}{\omega^2}}\left(1 + \frac{2m_c^2}{\omega^2}\right). \tag{7.11}$$

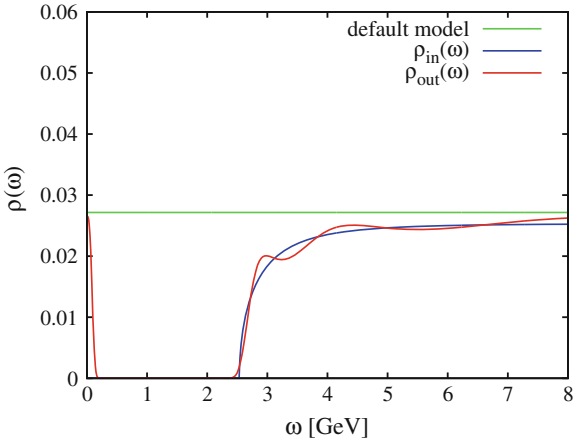

Fig. 7.3 The spectral function (*red line*) obtained from an MEM analysis of mock data constructed from the input spectral function of Eq. (7.11) (*blue line*). The *green line* shows the default model used in the MEM analysis

As before, we numerically construct mock data from this input spectral function and then carry out the MEM analysis using this input. The outcome of this procedure is given in Fig. 7.3. We see from this result, that even though the extracted spectral function shows some oscillating behavior around the input function, the reproduction of the continuum structure is reasonably accurate and the mentioned oscillations are small enough not to be confused with the charmonium peaks obtained earlier. Therefore, we can conclude that the MEM analysis of this channel is good enough to distinguish between spectral functions with a clear ground state peak or only a smooth continuum.

7.3.2 OPE Analysis at T = 0

Let us now turn to the MEM analysis of Eqs. (7.2) and (7.7). First, we investigate the spectral function at $T = 0$ for both the two S-wave (vector and pseudoscalar) and P-wave (scalar and axialvector) channels. To determine the upper boundary of the region of ν to be analyzed, we employ the criterion that the dimension 6 term should be smaller than 20 % of the whole OPE expression of Eq. (7.2), which gives $\nu_{\rm max}^{\rm V} = 8.03$ in the vector, $\nu_{\rm max}^{\rm PS} = 7.29$ in the pseudoscalar, $\nu_{\rm max}^{\rm S} = 5.67$ in the scalar and $\nu_{\rm max}^{\rm AV} = 4.74$ for the axialvector channel. We keep these values fixed when going to finite temperature. In fact, in the temperature region around T_c, the relative contribution of the dimension 6 term at $\nu_{\rm max}^{\rm V,PS,S,AV}$ is even smaller, namely, around 10 % or less. The lower boundary of ν is chosen to be $\nu_{\rm min}^{\rm V,PS,S,AV} = 0.78$, corresponding to a Borel mass of $M = 3.0\,$GeV. We have checked that the obtained

spectral functions do not depend much on this choice. For the value of the charm quark mass m_c, we use a recent estimate giving $\overline{m}_c(m_c) = 1.277 \pm 0.026\,\text{GeV}$ (Dehnadi 2012), for α_s we employ the newest world average $\alpha_s(M_Z) = 0.1184 \pm 0.0007$ (Bethke 2009), while for the vacuum gluon condensate G_0^{vac} the standard value $G_0^{\text{vac}} = 0.012 \pm 0.0036\,\text{GeV}^4$ (Shifman et al. 1979a,b; Colangelo and Khodjamirian 2001) is applied. For the default model $m(\omega)$, we use a constant matched to the perturbative value of the spectral function at high energy, as was done in similar studies using lattice QCD (Asakawa and Hatsuda 2004). Note, that we here do not have to introduce a default model, which approaches the expected value of the spectral function at high *and* low energy, which was necessary in the ρ meson case, because the charmonium sum rules contain more information on the region below the ground state as the quark mass scale m_c is fully included in the OPE. This allows us to use a default model which is only fitted to the high energy behavior of the spectral function.

The resulting spectral functions are given is Fig. 7.4. We observe in both S-wave channels a clear ground state peak, corresponding to η_c and J/ψ. The spectral functions also exhibit a second peak, which is, however, not statistically significant. These second peaks most likely reflect the existence of several excited states, which the MEM analysis is not able to resolve, quite similar to the situation encountered in lattice studies. For the P-wave channels we also see peaks close to the expected positions of χ_{c0} and χ_{c1}, which are, however, smaller than the errors of the same region.

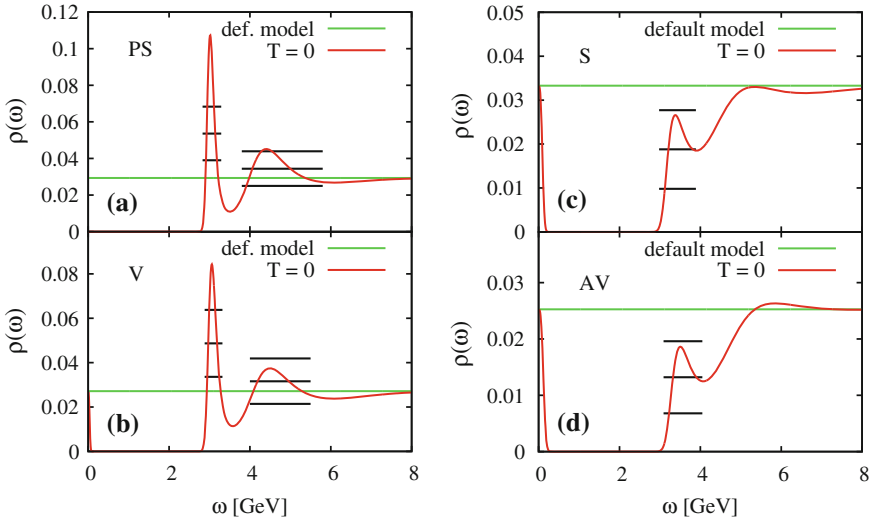

Fig. 7.4 The spectral functions in the pseudoscalar (*upper left plot*), vector (*lower left plot*), scalar (*upper right plot*) and axialvector (*lower right plot*) channel at $T = 0$ (*red lines*), with the errors of their averaged values in the peak regions. The horizontal extent and position of the error bars indicates the region over which the average is taken. The *green lines* show the default model used in the MEM analysis

This means, that, strictly speaking, we can not make any statement about the existence of these peaks in reality. The fact that they lie very close to the experimentally observed positions of χ_{c0} and χ_{c1} is nevertheless very suggestive and can at least be considered as evidence that the peaks we have obtained reflect to some degree the properties of these P-wave charmonia.

Furthermore, it is seen that the spectral function of the vector, scalar and axialvector channels approach the default model in the region close to $\omega = 0$, which, however, should not be confused with a contribution of the scattering term. This behavior is an artifact caused by our usage of the derivative sum rule in this channel and should thus not be considered to be a physical effect.

Numerically, for the S-wave channels the peak representing η_c lies at 3.02 GeV, while the one standing for J/ψ is found at 3.06 GeV. Thus, we see that the ground state in both channels reproduces the experimental value with a precision of the order of 50 MeV. In the vector channel, the residue can be related to the electronic width of the corresponding resonance. We can obtain this residue from Fig. 7.4 simply by integrating the spectral function in the region of the peak, which gives 0.162 GeV^2, which is in good agreement with the experimental value of 0.173 GeV^2. On the other hand, we observe that the hyperfine splitting between η_c and J/ψ is underestimated. For the P-wave channels, the peak corresponding to χ_{c0} lies at 3.38 GeV, while the one representing χ_{c1} is found at 3.50 GeV, which as discussed above compares very well with the experimental values.

These findings are in qualitative agreement with the results obtained in the conventional analysis of the charmonium sum rules. For instance, it was discussed in Marrow et al. (1987), that for reproducing the experimental values of the hyperfine splitting, one has to employ a value of the scalar dimension 4 gluon condensate which is 3–5 times larger than the standard value. We have found the same tendency in our MEM analysis. It is however not our goal here to fine tune the values of the gluon condensate and the quark mass to exactly reproduce the vacuum properties of η_c and J/ψ, as we are mainly interested in a possible change of the properties of these peaks around T_c. We will therefore use the spectral functions shown in Fig. 7.4 as a reference, to which the results at finite temperature will be compared.

7.3.3 OPE Analysis at $T \neq 0$

Next, we increase the temperature according to Eq. (7.2). The resulting spectral functions are shown in Fig. 7.5 at temperatures between 0.9 T_c and 1.2 T_c. It is seen in the figure that the behavior of the spectral functions changes abruptly in the vicinity of T_c. First, both S-wave ground state peaks experience a shift to lower energies of the order of 50 MeV, before dissolving quickly into the continuum above the critical temperature. Investigating the spectral functions in more detail, one observes that η_c disappears already at $T = 1.1\ T_c$, while J/ψ survives a bit longer, with still a small bump remaining at $T = 1.2\ T_c$. For the P-wave states, we do not see a downward shift, but only the abrupt disappearance of the peaks around $T = 1.0\ T_c$. It is

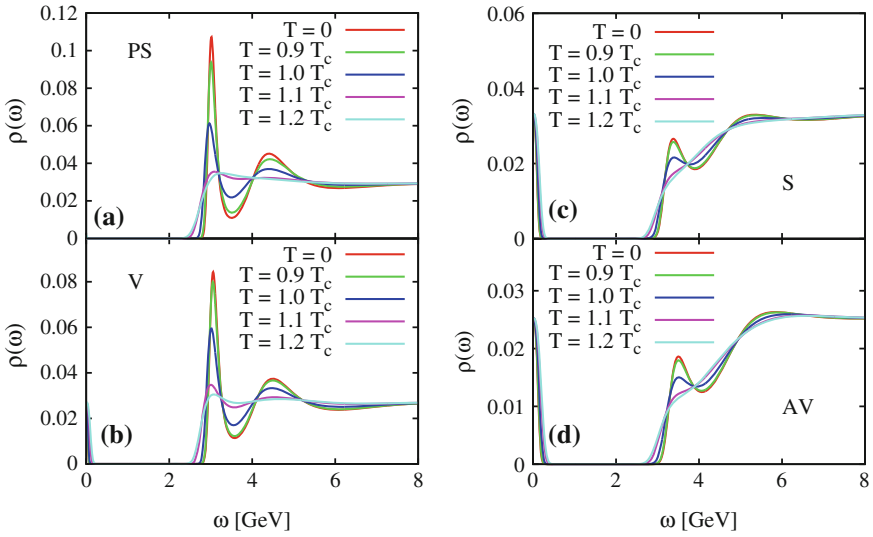

Fig. 7.5 The spectral functions of the pseudoscalar (*upper left plot*), vector (*lower left plot*), scalar (*upper right plot*) and axialvector (*lower right plot*) channel at various temperatures around T_c

reassuring to note that our results are consistent with the findings of Morita and Lee (2008a,b, 2010) in the sense that both observe a negative energy shift of the S-wave peaks around T_c. In these earlier works, it was, however, not possible to discuss the possible melting of the peaks because a relativistic Breit-Wigner form for the spectral function was assumed at all investigated temperatures.

This sudden qualitative change of all the investigated spectral functions mainly originates from the changes of the third and fourth terms in Eq. (7.2), which can be

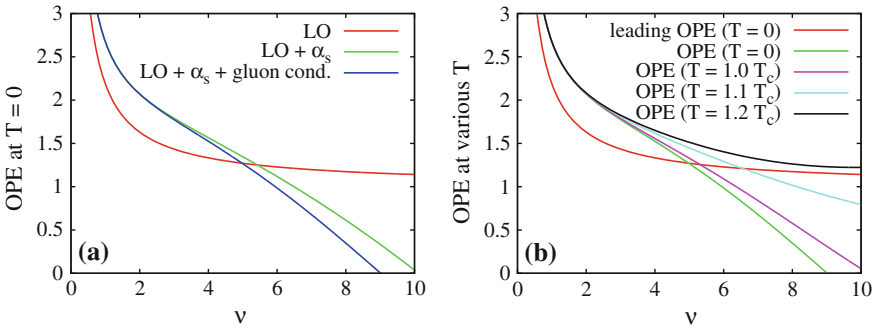

Fig. 7.6 **a** The OPE data of the vector channel at zero temperature are shown, split up into the various components. Here, for better visibility, the derivative of Eq. (7.2), divided by its leading term is shown. **b** The same quantity is given at various temperature values

traced back to the rapid adjustment of the thermodynamic quantities $\varepsilon(T)$ and $p(T)$ around T_c. To illustrate this point, we show the OPE data of Eq. (7.2) of the vector channel (all the other channels exhibit qualitatively the same behavior) in Fig. 7.6. Note that, as discussed before, we actually use the derivative of Eq. (7.2) for our analysis, which is also the quantity given in the figures. From the left plot of Fig. 7.6, we can observe that the gluon condensate terms are in fact considerably smaller than the α_s corrections, both contributions having however the same sign. The peak corresponding to the J/ψ obtained in Fig. 7.4 is caused by the effects of these two gluon condensate and α_s terms. Turning now to the right plot, we see that once the temperature is increased above T_c, the gluon condensate terms switch their sign and gradually start to cancel the α_s corrections (remember, that only the terms containing gluonic condensates change their values at finite T). Finally, at about $T = 1.2\ T_c$, both terms cancel each other to a large degree, leaving only the leading term as the dominant contribution and thus causing the J/ψ-peak to vanish.

For obtaining firm conclusions, one has to test the reliability of the MEM procedure at finite temperature, where systematic effects decrease the reproducibility and resolution of the spectral function obtained from MEM. In lattice studies, this reduced reliability is primarily caused by the reduction of the data points in the imaginary time correlator, due to periodicity and the reduction of the maximal time extent. In the case of QCD sum rules, this problem does not exist, as Eq. (7.2) is given as a continuous function at any temperature and therefore the same number of data points can be used. Nevertheless, the reliability of the MEM technique is still reduced at finite temperature due to the uncertainties of the thermodynamic functions $\varepsilon(T)$ and $p(T)$, whose contribution grows with temperature and therefore increases the error of Eq. (7.2). In order to confirm that the change of the spectral function in Fig. 7.5 is not an artifact, we reanalyze Eq. (7.2) at $T = 0$, but use the errors of $T \neq 0$ in the analysis. The results are then compared to the ones given in Fig. 7.5, to investigate the net temperature effect on the spectral function. Such a comparison is given in Fig. 7.7. As the most prominent change happens around $1.0 \sim 1.1\ T_c$, we show representative results of this temperature region. It can be seen in the figure that for the S-wave channels, even though the height of the peaks of the spectral functions at $T = 0$ is indeed reduced because of the increased error, this effect is much smaller than the actual reduction of the peaks around T_c, seen in Fig. 7.5. For the P-wave channels, the reduction of the peaks due to the increasing error is even smaller, as the properties of the peaks change only very little with changing error. We therefore conclude that the disappearance of the peaks observed in Fig. 7.5 is a physical effect and is not induced by an artifact of the MEM analysis.

7.3.4 Summary for Charmonium

We have extracted the spectral functions of the pseudoscalar, vector, scalar and axialvector channel at both zero and finite temperature using a combined analysis of QCD sum rules and MEM. At $T = 0$, the MEM technique is able to clearly resolve

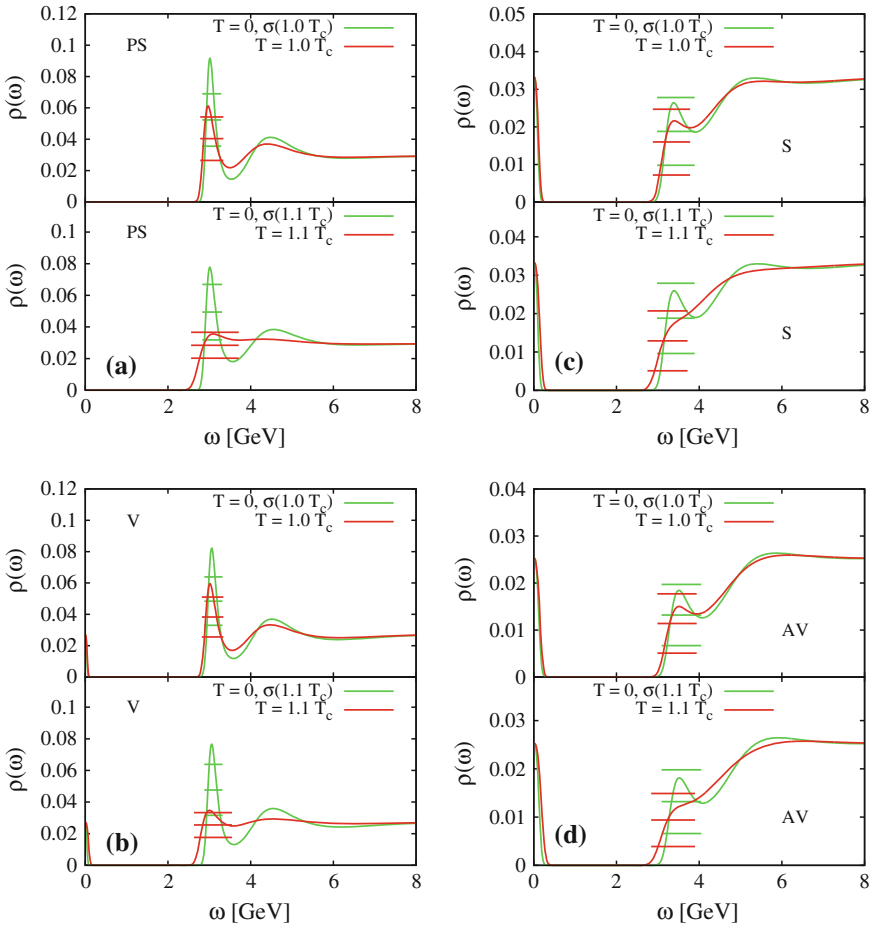

Fig. 7.7 Spectral functions of all four investigated channels at $T = 1.0\ T_c$ ($T = 1.1\ T_c$) and $T = 0$. The pseudoscalar channel is shown on the *top left* (**a**), the vector on the bottom left (**b**), the scalar on the *top right* (**c**) and the axialvector on the *bottom right* (**d**) plot. The *green lines* show the MEM results for the OPE at $T = 0$ with the error of $T = 1.0\ T_c$ ($T = 1.1\ T_c$) and the *red lines* the corresponding results for both OPE and error at $T = 1.0\ T_c$ ($T = 1.1\ T_c$)

the lowest energy peaks, corresponding to the η_c and J/ψ resonances. The positions of both peaks agree with the experimental values with a precision of about 50 MeV. On the other hand, the peaks representing the χ_{c0} and χ_{c0} resonances, even though they are visible in the respective spectral functions, are not statistically significant and we can thus not make strong statements about their properties. Nevertheless, as they lie close to the experimental values, we expect the obtained peaks to at least provide some hints about the fate of the P-wave states at finite T.

At finite temperature, we find that all observed states melt quickly after the temperature is raised above the deconfinement temperature T_c, caused by the sudden

change of the dimension-4, scalar and twist-2 gluon condensates in this temperature region. We have checked that this effect is not an artifact of the systematics of the MEM analysis. These results quantitatively disagree with the earlier findings of lattice studies which suggest that both η_c and J/ψ can survive at temperatures of up to 1.5 T_c or higher. It, however, has to be mentioned that our results are in fact consistent with the latest lattice results (Ding 2010), finding the peaks of η_c and J/ψ to be largely distorted between 0.73 T_c and 1.46 T_c. It remains to be seen whether or not the two methods will converge to compatible conclusions when more accurate analyses will become available in the future.

7.4 Results of the MEM Analysis for Bottomonium

Next, we move on to the MEM analysis of the bottomonium channels. The method is essentially the same as in the last section, we just have to change the quark mass from m_c to m_b. There are, however, several technical differences between the two cases. Firstly, the contributions of the terms containing gluonic condensates are strongly suppressed because of the heavier bottom quark mass. This leads to a very good convergence of the OPE, while the convergence of the α_s-expansion of the Wilson coefficients changes only little. This means, that in contrast to the charmonium case, the convergence of the OPE expression as a whole should be rather judged by the perturbative α_s-expansion, which is what we will do in this section. Secondly, the relative importance of the excited states grows for bottomoium, because they lie, relative to the bottom mass scale, quite close to the ground state and can thus not be ignored. As we will discuss in what follows, this introduces additional difficulties in interpreting the MEM analysis results, but also opens the possibility to make some predictions about the behavior of the excited states at finite temperature.

The results of this section, which are based on Suzuki et al. (2013), have been mainly obtained by K. Suzuki.

7.4.1 Mock Data Analysis

As before, we first investigate the ability of the MEM technique to reproduce the structures of interest in the bottomonium spectral functions. We here again only investigate the vector channel, because experimental information is available for this case. The used mock spectral function has the following form:

$$\rho_{\mathrm{mock}}(\omega) = \sum_{i=1}^{3} \frac{27}{8\pi\alpha^2} \Gamma_i(e^+e^-)\delta(\omega - m_i) + \frac{1}{1+e^{(\omega_0-\omega)/\delta}} \frac{1}{\pi} \mathrm{Im}\Pi^{\mathrm{V}}_{\mathrm{OPE}}(\omega^2_{\mathrm{max}}), \tag{7.12}$$

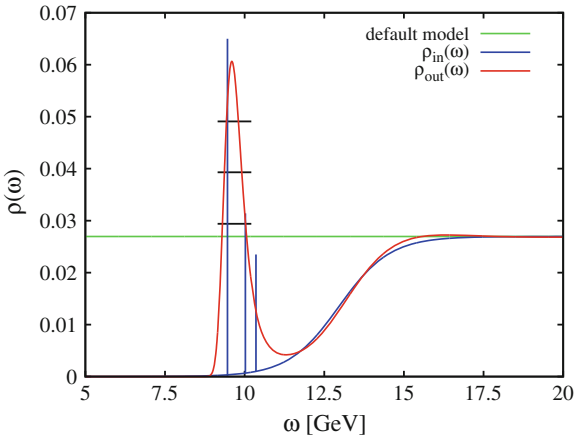

Fig. 7.8 The spectral function (*red line*) obtained from an MEM analysis of the mock data constructed from the input spectral function of Eq. (7.12) (*blue line*). The green line shows the default model used in the MEM analysis

Note the different factor in the first term ($\times 4$) compared to the charm case of Eq. (7.10), which originates from the fact that the absolute value the electro-magnetic charge of the bottom quark is only half as large as the one of the charm quark. We take the three lowest bottomonium states into account and use $\omega_0 = 12.96\,\text{GeV}$ and $\delta = 0.8\,\text{GeV}$ to describe the continuum part.

The result of the MEM analysis of the mock data constructed from Eq. (7.12) is given in Fig. 7.8, which should be compared with Fig. 7.1 for the respective charmonium outcome. We clearly see that here, the first and second excited states lie relatively close to the ground state and thus have quite a strong influence on the position and form of the lowest peak. Numerically, the position of the lowest peak that we can extract from our analysis turns out to be 9.60 GeV, which is about 140 MeV above the experimental value of the ground state, which lies at 9.46 GeV. This shift is caused by the presence of the excited states, which can be confirmed by repeating the analysis with mock data, in which the excited states are removed, The result of this check is shown in Fig. 7.9. From these figures, it is observed that once the second excited state is taken away, the lowest peak moves to 9.54 GeV, which is however still 80 MeV too high. Removing then also the first excited state, the lowest peak of the resulting MEM analysis is obtained at 9.45 GeV, which agrees with the input value within a precision of 10 MeV. In conclusion, we have found that the MEM analysis of this particular bottomonium channel is not able to disentangle the lowest three narrow states in the spectral function. Therefore, the lowest peak extracted from our MEM technique should (at least in the vector channel) be considered to be composed of three overlapping peaks. This of course makes the interpretation of the behavior of this peak at finite temperature more difficult as it will not always be possible to determine which of the three original peaks has caused a certain shape-change of the peak. One can, however, reasonably assume that higher peaks will receive stronger

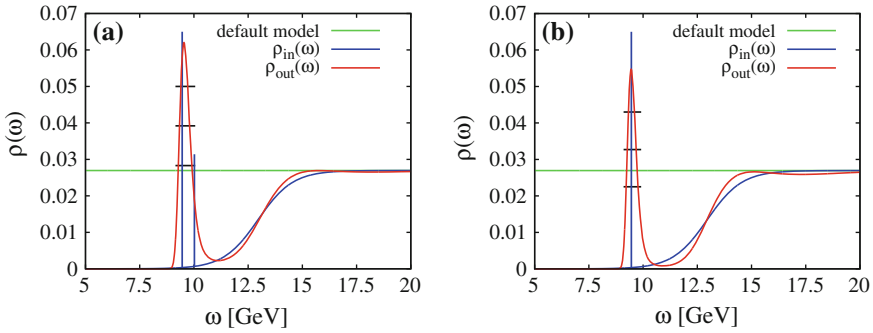

Fig. 7.9 Same as in Fig. 7.8, but with second (*left figure*) or first and second (*right figure*) excited states removed from the input spectral function

temperature effects because of their larger spatial extension and thus can still obtain meaningful conclusions from the MEM analysis. This will be our strategy of this section.

As a last point, let us, as before, check the validity of our analysis for the case, in which no existing peak can be found in the spectral function. The result of this test is given in Fig. 7.10. As for the charm case, the analysis works very well here and we can hence safely expect the MEM analysis to be able to distinguish between spectral functions with and without peaks.

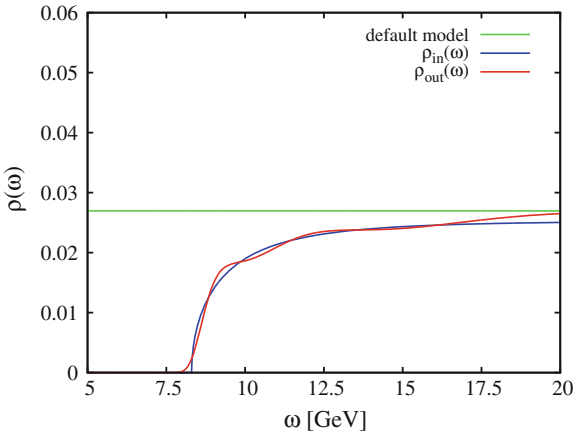

Fig. 7.10 The spectral function (*red line*) obtained from an MEM analysis of the mock data constructed from the input spectral function of Eq. (7.11), in which m_c is replaced by m_b (*blue line*). The *green line* shows the default model used in the MEM analysis

7.4.2 OPE Analysis at $T = 0$

Before showing the analysis results, let us note that, as already mentioned above, we here determine the upper boundary of the analyzed ν-region from the α_s expansion of the perturbative term of the OPE. This gives $\nu^{\mathrm{V}}_{\mathrm{max}} = 8.23$ in the vector, $\nu^{\mathrm{PS}}_{\mathrm{max}} = 9.33$ in the pseudoscalar, $\nu^{\mathrm{S}}_{\mathrm{max}} = 7.33$ in the scalar and $\nu^{\mathrm{AV}}_{\mathrm{max}} = 5.83$ for the axialvector channel. For the lower boundary, we adopt the value, for which the lowest peak becomes most pronounced, and hence find $\nu^{\mathrm{V}}_{\mathrm{min}} = 4.0$ in the vector, $\nu^{\mathrm{PS}}_{\mathrm{min}} = 4.3$ in the pseudoscalar, $\nu^{\mathrm{S}}_{\mathrm{min}} = 4.0$ in the scalar and $\nu^{\mathrm{AV}}_{\mathrm{min}} = 3.5$ for the axialvector channel. We have checked that our results depend only weakly on the detailed choice of this lower boundary. The values of the vacuum gluon condensate G_0^{vac} and the strong coupling constant α_s are taken to be the same as for the charmonium case. For the bottom quark mass, we use $\overline{m}_b(m_b) = 4.167 \pm 0.013\,\mathrm{GeV}$ (Narison 2012). As one further point, let us mention the dimension 6 gluon condensate term which was explicitly taken into account in the preceding charm section. Due to the large bottom quark mass, this term is suppressed by a factor of ∼4000 compared to the charmonium sum rules. Because this term was small already for the charm case, we can safely neglect it for the bottom sum rules in this section.

The obtained spectral functions of the vector, pseudoscalar, scalar and axialvector channels at zero temperature are shown in Fig. 7.11. The lowest peak of each

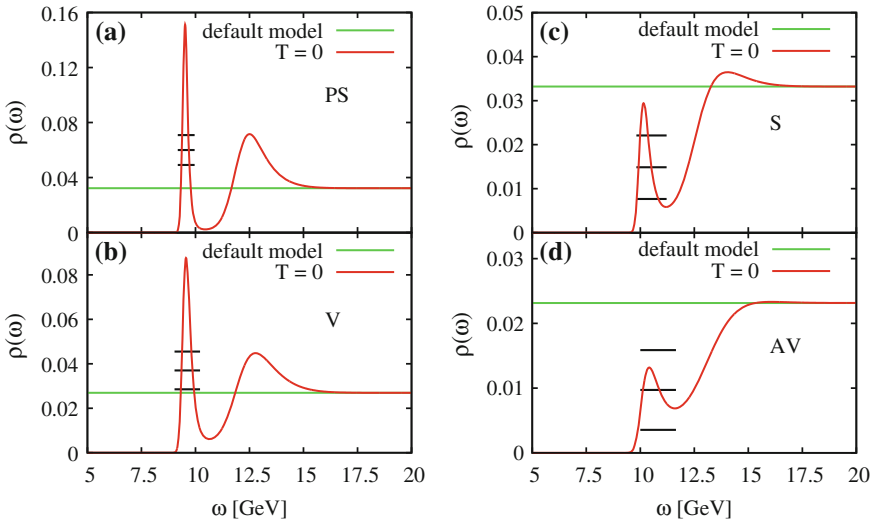

Fig. 7.11 The spectral functions in the pseudoscalar (*upper left plot*), vector (*lower left plot*), scalar (*upper right plot*) and axialvector (*lower right plot*) channel at $T = 0$ (*red lines*), with the errors of their averaged values in the lowest peak region. The horizontal extent and position of the error bars indicates the region over which the average is taken. The *green lines* show the default model used in the MEM analysis

channel is located at $m_\Upsilon = 9.56\,\text{GeV}$, $m_{\eta_b} = 9.51\,\text{GeV}$, $m_{\chi_{b0}} = 10.15\,\text{GeV}$, $m_{\chi_{b1}} = 10.42\,\text{GeV}$. These values are somewhat higher than the experimental masses, which was however to be expected from our analysis of mock data since the obtained peaks contain contributions from both the ground and excited states. In the vector channel, the validity of this picture is confirmed by evaluating the residue of the peak and comparing it with the values obtained from the leptonic decay width. Concretely, we obtain 0.0476 GeV for the residue of the lowest peak, while the three lowest peaks ($\Upsilon(1S)$, $\Upsilon(2S)$ and $\Upsilon(3S)$) summed give a total residue which is experimentally found to be 0.0483 GeV.

Another observation that can be made from Fig. 7.10 is that the S-wave peaks (Υ, η_b) are statistically significant, while for the P-wave channels this is true only marginally (χ_{b0}) or not at all (χ_{b1}). This means that the results obtained here for the later two cases are not conclusive and should be considered to have only indicative character. This situation will likely be improved once the OPE of these channels is known with better precision and for a wider range of the Borel mass.

7.4.3 OPE Analysis at $T \neq 0$

The results of the spectral functions at finite temperature are shown in Fig. 7.12. First, the peak positions for all the channels undergo shift to lower energy regions with increasing temperature. This behavior suggests that the excited states dissociate

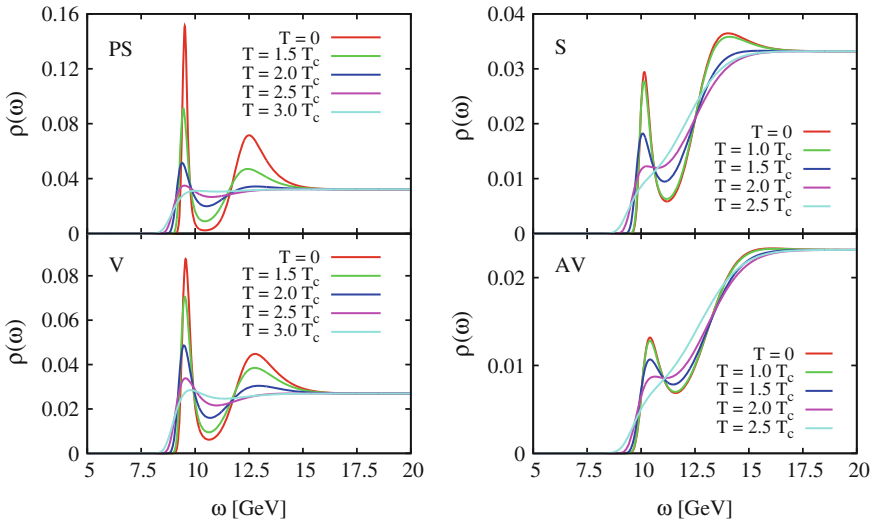

Fig. 7.12 The bottomonium spectral functions of the pseudoscalar (*upper left plot*), vector (*lower left plot*), scalar (*upper right plot*) and axialvector (*lower right plot*) channel at various temperatures below 3.0 T_c

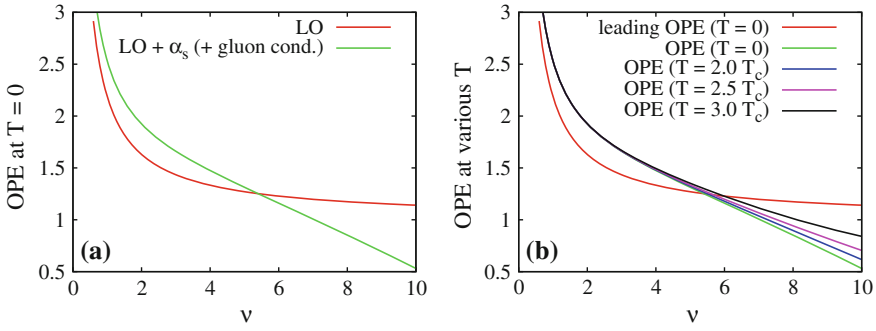

Fig. 7.13 In the *left plot*, the bottom OPE data at zero temperature are shown, split up into the various components. Here, for better visibility, the derivative of Eq. (7.2), divided by its leading term is shown. Note that the gluon condensate contribution at $T = 0$ is too small to have any visible effect in this figure. On the *right plot*, the same quantity is given at various temperature values

at lower temperature than the ground state. Next, the peaks gradually disappear, become broader and shift to slightly higher energies. At the same time, a continuum-like structure penetrates into the peak regions and moves downward. Furthermore, it is meaningful to mention the difference between the behavior of S-wave and P-wave channels. Υ and η_b exist as a clear peak up to $2.0\,T_c$ and still survive at $2.5\,T_c$. In the case of Υ, a bump remains even at $T = 3.0\,T_c$. On the other hand, one can observe that χ_{b0} and χ_{b1} completely disappear at $2.0 \sim 2.5\,T_c$.

Let us here briefly discuss what causes the various bottomonia states to persist as clear peaks up to higher temperatures than their charmonium counterparts. This different behavior can be explained by the fact that all temperature effects originate from the gluon condensate terms, which are inversely proportional to the fourth power of the quark mass. Therefore, as the bottom quarks are more than four times heavier than the charm quarks, the gluon condensate terms are relatively suppressed by a factor of more than $\sim$250 for the bottomonium channels, which necessitates much higher temperatures for having any sensible effect on the bottomonium spectral functions. We illustrate this point by showing the OPE of the bottomonium vector channel at 0 and finite temprature in Fig. 7.13, which should be compared with Fig. 7.6. In the left figure, one sees that at $T = 0$, the gluon condensate terms are almost completely negligible due to the suppression mentioned above. Furthermore, the right figure shows that the cancellation of the α_s-correction and the gluon condensate terms, which was the essential reason for the melting of the charmonium states, is only partially realized even at $T = 3.0\,T_c$.

In order to confirm that the obtained results are caused by genuine physical effects, we have to check possible contributions of MEM artifacts at finite temperature. First, as the contributions of the gluon condensates increase at finite temperature, their uncertainties magnify the OPE error. Therefore, it is expected that the resolution of the MEM is reduced and the peaks of the extracted spectral functions become broader. Thus, to investigate this effect, we reanalyze the spectral functions by using

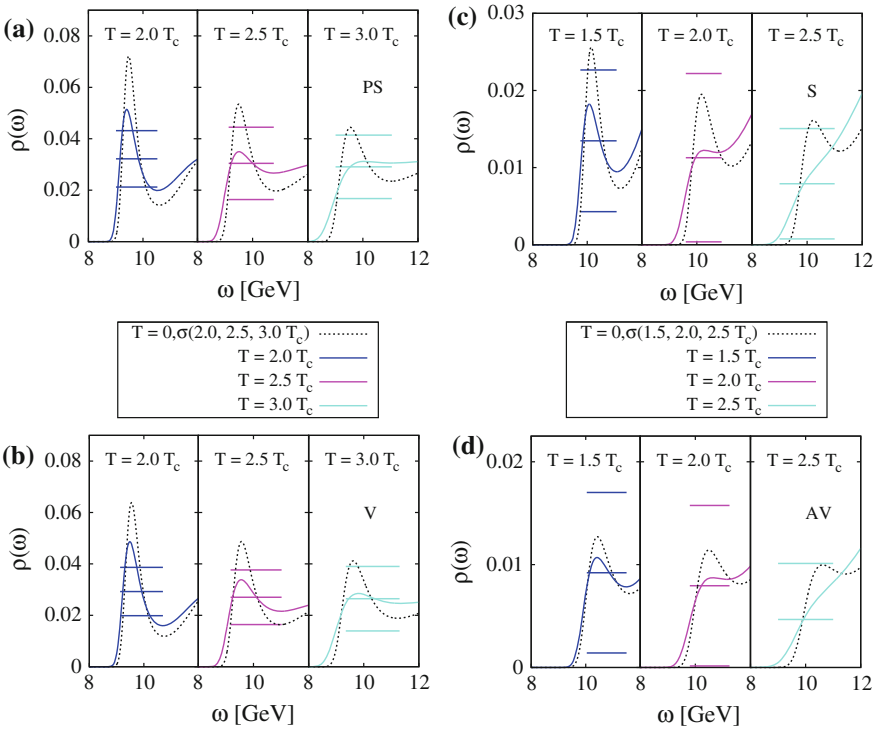

Fig. 7.14 *Solid lines* show the spectral functions of bottomonia obtained from finite temperature OPE data and the respective error bars. The *dashed lines* stand for spectral functions at zero temperature using error at corresponding finite temperature. *Upper left*: pseudo-scalar channel, *lower left*: vector channel, *upper right*: scalar channel, *lower right*: axial-vector channel

the OPE at $T = 0$ with error at finite temperature. The result for each channel is shown as dashed lines in Fig. 7.14. As one can see, although the heights of the peaks are reduced partly due to the MEM artifact described above, the peaks are still present in the S-wave (P-wave) channels with the error of $T/T_c = 3.0$ ($T/T_c = 2.5$). We also stress that the MEM artifact does not shift the peak position. From this analysis, we conclude that the disappearance of the peaks at the finite temperatures is caused by physical effects and not due to an MEM artifact.

Since the uncertainties involved in the calculation are quite large, as indicated by the error bars in Fig. 7.14, we can presently not make statements about specific numbers for the melting temperatures of the investigated states, but can only restrict the ranges of temperatures within which the peaks disappear. Concretely, we define the range of the melting temperatures as follows. The *upper limit* is determined as the temperature where the bump (extremum) disappears, while the *lower limit* taken as the temperature where the error bar exceeds the lowest-energy peak height of the spectral function, namely when the peak ceases to be statistically significant. The error bars for each temperature are shown in Fig. 7.14. The resultant upper and lower

Table 7.1 Upper and lower limits of the dissociation temperature ranges for the lowest bottomonium states. The precise definition of these limits is given in the text

Channel	Vector ($\Upsilon(1S)$)	Pseudoscalar (η_b)	Scalar (χ_{b0})	Axialvector (χ_{b1})
T/T_c	>2.3	>2.1	1.3–2.5	<2.5

limits of the dissociation temperatures are summarized in Table 7.1. Once we are able to calculate the OPE data with better precision, these temperature ranges will become narrower, thus allowing us a more accurate determination of the melting temperatures in the future.

7.4.3.1 Investigation of the Excited States of the Vector Channel

In our MEM analysis of mock data, we have shown that information on the excited states is included in the lowest peak of the spectral functions extracted from MEM. For obtaining information on the behavior of these excited states at finite temperature, we will now try to evaluate the residue of this lowest peak at various temperatures. However, for this evaluation, we cannot naively integrate the spectral function in the region of the peak because of the possible contributions of the continuum states which move downwards into the peak region from high energies as the temperature increases and which are thus not negligible.

To estimate the residue values with an excluded continuum background, we fit the obtained spectral functions by using a Breit-Wigner or Gaussian fitting function for the peak, while the continuum is parametrized by the leading order perturbative result. Specifically, we use

$$f(x) = \frac{|\lambda|^2}{2\pi} \frac{\Gamma}{(x-m)^2 + \Gamma^2/4} + \frac{1}{8\pi^2} \sqrt{1 - \frac{4a^2}{x^2}} \left(2 + \frac{4a^2}{x^2}\right), \tag{7.13}$$

for the Breit-Wigner + continuum form, in which we have four fitting parameters: $|\lambda|^2$, m, Γ, and a. They correspond to the residue, peak position, width, and continuum threshold ($2a$), respectively. These parameters are then fitted by the Levenberg-Marquardt method. Furthermore, in order to exclude a possible initial value dependence of the fitting procedure and to investigate the existence of local minima, we take 200 initial values generated randomly for the four fitting parameters at each temperature.

The temperature dependence of the residue of the Υ peak which is fitted by the Breit-Wigner + continuum form of Eq. (7.13) is shown on the left side of Fig. 7.15. For each temperature, 200 results corresponding to the different initial values are plotted. For some temperature, we find multiple solutions, which are supposed to be local minima solutions of the least-square function χ^2 in the L-M method. In the case of the Breit-Wigner + continuum fitting, we find that the local minimum form

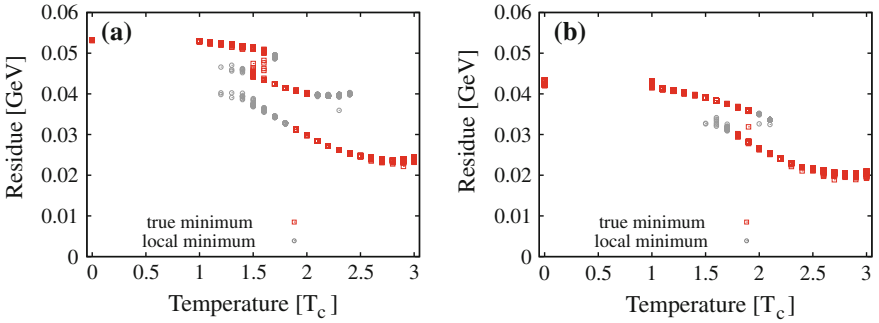

Fig. 7.15 Temperature dependence of the residue of the lowest peak for the vector channel peak fitted with a Breit-Wigner + continuum (**a**), and as Gaussian + continuum (**b**). For each temperature, we used 200 different initial values for the fitting parameters, which, as a result of the fit, give the points shown in the plots. This gives us information on the existence of local minima in the least-square function χ^2. The red points are corresponding to true minima (smallest χ^2), and gray points stand for local minima (larger χ^2). The fit range is fixed to 7.0–12.0 GeV

three clusters, top, middle, and bottom. For some T/T_c, one sees that the clusters are diffused and solutions are scattering to interpolate two clusters. There, the minimum valley of χ^2 seems to become flat between the two minima.

The red points at each T/T_c are the true minimum points, where χ^2 hits the minimum. One sees that at low T/T_c, the true minimum is located in the top cluster, while it moves down to the middle cluster at around $T/T_c \sim 1.5-1.6$, and further down to the bottom cluster at around $T/T_c \sim 1.9-2.0$. At $T/T_c \geq 2.5$, we have only one stable solution. The peak positions and the continuum thresholds from the fitting with the Breit-Wigner + continuum form is shown in Fig. 7.16. Both the peak positions and continuum threshold undergo a transition towards smaller values at $T/T_c \sim 1.5-2.0$.

To make sure that this fitting analysis is valid, we have repeated the same procedure with another fitting function, namely a Gaussian + continuum form:

$$f(x) = \frac{|\lambda|^2}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] + \frac{1}{8\pi^2}\sqrt{1-\frac{4a^2}{x^2}}\left(2+\frac{4a^2}{x^2}\right). \tag{7.14}$$

The respective result is shown in the right side of Fig. 7.15, which reveals a similar behavior as for the Breit-Wigner fit. It however shows only two clusters of local minima and the transition from the higher to the lower cluster at $T/T_c \sim 1.8-1.9$. The other fitting parameters also show patterns resembling the ones shown in Fig. 7.16, with small modifications. Therefore, comparing these two fitting results, we observe that the qualitative behavior of the residue does not depend on the functional form used for fitting the peak region.

It can be concluded from the above results that the residue of Υ peak decreases gradually with increasing temperatures and becomes a constant value at higher

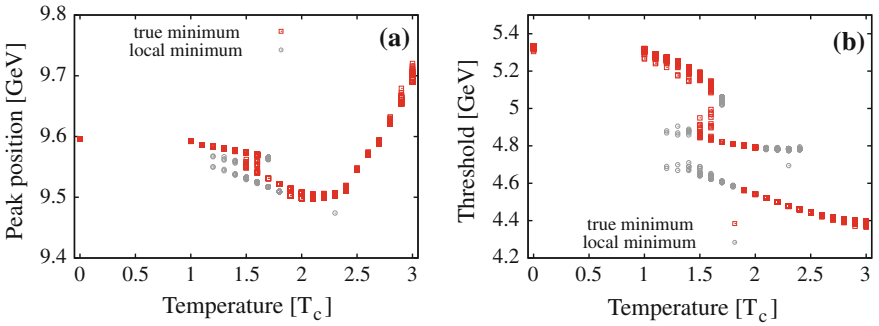

Fig. 7.16 Temperature dependence of the mass and threshold of the fit of Eq. (7.13) to the spectral function of the vector channel. On the left side, the mass (peak position) is shown while the right side shows the continuum threshold. As explained in the caption of Fig. 7.15, red points are corresponding to true minima (smallest χ^2), and gray points stand for local minima (larger χ^2). The fit range is fixed to 7.0–12.0 GeV

temperature. Especially, a rapid reduction of the residue is seen at $T/T_c = 1.5{-}2.0$. It should be noted that this behavior does not directly imply that the excited states, $\Upsilon(2S)$ and $\Upsilon(3S)$, disappear at lower temperatures than the ground state $\Upsilon(1S)$ because one in principle cannot exclude other possibilities such as simultaneous reduction of the ground and excited states. Nevertheless, if we assume that the excited states disappear at lower temperatures than the ground state, our results suggest that $\Upsilon(2S)$ and $\Upsilon(3S)$ vanish at $T/T_c = 1.5{-}2.0$, while $\Upsilon(1S)$ survives up to $T/T_c = 3.0$.

7.4.4 Summary for Bottomonium

In summary, we have analyzed the bottomonium spectral functions at zero and finite temperature by using a newly developed analysis method of QCD sum rules. The maximum entropy method (MEM) is adapted to extract the spectral function from the sum rule.

At $T = 0$, the lowest peak has been obtained for each channel corresponding to Υ, η_b, χ_{b0}, and χ_{b1}. Although these mass spectra agree qualitatively with the experimental values, their peak positions are slightly shifted to higher energies. By analyzing mock data for the vector channel and evaluating the obtained residue, we conclude that this disagreement is caused by the contribution of the excited states.

Next, we have investigated the temperature dependence of the spectral functions. Temperature dependences are taken into account in the gluon condensates, which are estimated from quenched lattice QCD data at finite temperature. As a result, we have found that the spectral functions of bottomonia are modified much slower as functions of T/T_c than those of charmonia, in which the lowest peak disappears

suddenly at the vicinity of T_c. Using the definitions of the upper and lower limits of the melting temperature given in Sect. 7.4.3, we find that Υ and η_b survive as a peak in the spectral functions up to some temperature restricted to the regions of $T/T_c > 2.3$ and $T/T_c > 2.1$, while the dissociation temperatures of χ_{b0} and χ_{b1} are confined to $T/T_c = 1.3-2.5$ and $T/T_c < 2.5$, respectively.

It should be noted, however, that our definition inevitably contains some ambiguity due to the limitation of the OPE and MEM. Therefore, respective results on the melting temperature should be regarded as qualitative guides. Furthermore, both P-wave peaks are not found to be fully significant statistically even at $T = 0$, which means that we can not make a definite conclusion about the fate of these states at finite temparature. To obtain more conclusive results on their behavior, further studies are needed once more precise information on the OPE is available. The current prediction of the melting temperatures depends on the extracted temperature dependences of the gluon condensates. For these we have used quenched lattice QCD data for the energy density and pressure. To go beyond the quenched approximation, a more detailed analysis will be required to include full QCD information on the gluon condensates, which will be the subject of a future investigation.

Our results are qualitatively consistent with previous QCD sum rule analyses employing a more conventional method (Morita and Lee 2010). As mentioned above, however, it turns out that our extracted lowest peaks contain excited states as well as the ground state, so that deformations of such peaks depend on the behavior of the excited states. Therefore, to extract more detailed information on the spectral function of the vector channel, we have investigated the temperature dependence of the residue of the lowest peak obtained from MEM for this channel. Doing this, we have observed that the residue decreases with increasing temperature, which is consistent with a picture in which the excited states, $\Upsilon(2S)$ and $\Upsilon(3S)$, dissociate at lower temperatures than the ground state $\Upsilon(1S)$. Assuming such a scenario, our results indicate that $\Upsilon(2S)$ and $\Upsilon(3S)$ disappear in the temperature region of $T/T_c = 1.5-2.0$.

References

Aarts G et al (2011a) JHEP 11:103
Aarts G et al (2011b) Phys Rev Lett 106:061602
Asakawa M, Hatsuda T (2004) Phys Rev Lett 92:012001
Bethke S (2009) Eur Phys J C 64:689
Bochkarev AI, Shaposhnikov ME (1986) Nucl Phys B268:220
Boyd G et al (1996) Nucl Phys B469:419
Chu M-C, Schramm S (1995) Phys Rev D 51:4580
CMS Collaobation, arXiv:1208.2826 [nucl-ex]
Colangelo P, Khodjamirian A (2001) At the frontier of particle physics/ handbook of QCD, vol 3. World Scientific, Singapore, p 1495
Collaobation CMS (2011) Phys Rev Lett 107:052302
Datta S et al (2004) Phys Rev D 69:094507
Dehnadi B et al arXiv:1102.2264 [hep-ph]

Ding H-T et al (2010) PoS LAT2010, 180
Fukushima K, Hatsuda T (2011) Rep Prog Phys 74:014001
Gubler P, Morita K, Oka M (2011) Phys Rev Lett 107:092003
Hatsuda T, Koike Y, Lee SH (1993) Nucl Phys B394:221
Jacovác A et al (2007) Phys Rev D 75:014506
Kaczmarek O et al (2004) Phys Rev D 70:074505
Marrow J, Parker J, Shaw G (1987) Z Phys C 37:103
Matsui T, Satz H (1986) Phys Lett B 178:416
Miller DE (2007) Phys Rept 443:55
Morita K, Lee SH (2008a) Phys Rev Lett 100:022301
Morita K, Lee SH (2008b) Phys Rev C 77:064904
Morita K, Lee SH (2010) Phys Rev D 82:054008
Morita K, Lee SH arXiv:1012.3110 [hep-ph]
Nakamura K et al (2010) J Phys G 37:075021
Narison S arXiv:1105.2922 [hep-ph]
Novikov VA et al (1978) Phys Rept 41:1
Schafer T, Shuryak EV (1998) Rev Mod Phys 70:323
Shifman MA, Vainshtein AI, Zakharov VI (1979) Nucl Phys B147:385
Shifman MA, Vainshtein AI, Zakharov VI (1979) Nucl Phys B147:448
Song Y, Lee SH, Morita K (2009) Phys Rev C 79:014907
Suzuki K, Gubler P, Morita K, Oka M (2013) Nucl Phys A897:28
Umeda T, Nomura K, Matsufuru H (2004) Eur Phys J C 39:9

Part III
Concluding Remarks

Chapter 8
Summary, Conclusion and Outlook

8.1 Summary and Conclusion

We have developed in this thesis a novel analysis method for QCD sum rules, which makes use of the maximum entropy method, have tested its validity in a simple toy model and have applied it to several physical channels, namely the ones for the ρ meson, nucleon and quarkonium at zero and finite temperature.

Our investigation of the toy model sum rule, which is derived from non-relativistic quantum mechanics, indicates that MEM indeed is a useful tool for analyzing the sum rules, as long as the uncertainties of the OPE input are not too large. We also find that there are considerable differences in the reproducibility of the MEM depending on which kind of kernel is employed. Specifically, using the Borel kernel of Eq. (4.23) one can only hope to extract the position of the lowest lying peak, while the Gaussian kernel of Eq. (4.29) allows one to determine the position of the first excited state as well. It should however be remembered that the spectral function of the analyzed toy model only contains well separated discrete peaks and no continuum, which can largely complicate the extraction of a possible second peak. As a last point, we found that due to the limited resolution of the MEM procedure, the reconstructed peaks usually have a finite width, which is not of physical origin, but rather an artifact of the analysis method. This is a general finding, that holds similarly for all channels studied in this thesis.

Next, as a first physical application, we have analyzed the Borel sum rule of the ρ meson channel. This case shows features that are quite similar to the toy model analyzed before. As was shown in Chap. 5 we can extract the position and residue of the peak corresponding to the ρ meson with reasonable precision. However, as the uncertainties of the condensates are quite large, and because the sum rule does not contain much information on the spectral function in the low energy region, this extraction works only with a suitable choice of the default model, which incorporates our prior knowledge about the spectral function both in the high- and low-energy region. Physically, the perhaps most interesting result of this analysis is the clear connection that we have shown to exist between the mass of the ρ meson and the

P. Gubler, *A Bayesian Analysis of QCD Sum Rules*, Springer Theses,
DOI: 10.1007/978-4-431-54318-3_8, © Springer Japan 2013

quark condensate, which is illustrated in Fig. 5.11. This finding shows the important role played by the quark condensate (and thus of the spontaneous breaking of the chiral symmetry) for explaining the properties of the ρ meson.

For the analysis of the nucleon channel, which we analyze both for the Borel and Gaussian sum rule, we observe that, due to the slow convergence of the OPE, the Borel sum rule does not contain enough information on the ground state peak to be accurately extracted by MEM. This result is not surprising as this sum rule does not work well even if one resorts to a more conventional analysis method (Leinweber 1997). However, as could be expected from our discussion of the toy model, our further analysis shows that the use of the Gaussian kernel improves the precision of the MEM analysis and makes it possible to study the ground state and even (with a suitably chosen interpolating field) its excited states. The results of this investigation are, however not yet conclusive, as it is possible to obtain more detailed information from the analysis of the parity projected sum rule (Ohtani et al. 2013), for which, as a by-product of the present study, we have resolved certain technical issues as explained in Sect. 3.4.

Finally, the last topic studied in this thesis deals with the Borel sum rules of quarkonia (charmonium and bottomonium) at both zero and finite temperature. As it was already discussed in the introduction, the analysis of hadrons in hot or dense matter is a subject, where our novel method can exhibit its full strength, as not much is known about the spectral function of hadrons in such an environment. Therefore, our technique of using MEM allows us to study such a system without prejudice, which has not been possible before. Indeed, we show in Chap. 7 that, after having tested that we are able to reproduce the quarkonia ground states of several channels at zero temperature, the corresponding peaks vanish from the spectral function at some point above the deconfinement temperature T_c, which indicates that these particles cease to exist as quark-antiquark bound states and instead melt into freely moving quarks. This finding has important phenomenological impacts, as, for instance the charmonium vector particle J/ψ (and its possible disappearance) is considered to be a useful probe for the existence of the quark-gluon plasma (QGP) that is believed to be produced in current heavy ion collisions at RHIC and the LHC. Considering furthermore the melting temperatures of the various charmonium and bottomonium states, one can even imagine that all these states constitute a crude thermometer of the QGP.

In all, after having studied all the channels mentioned above, we can now confidently conclude that that MEM is a useful tool for analyzing QCD sum rules, especially for the cases in which the frequently used "pole + continuum" ansatz is not applicable. On the other hand, one must always keep in mind, that this method does not exactly solve the sum rules, but only gives the most probable form of the spectral function given all the available information. Therefore, it will presumably never provide the true spectral function, but only a smeared version of it, as the MEM has only a limited resolution. Nevertheless, we have shown that the method is accurate enough to give valid information of the properties of the lowest lying structure of the spectral function, whether it may be a narrow peak or some smooth continuum.

8.2 Outlook

One can envisage several directions of development, starting from the work presented in this thesis. I will here mention a few of them.

First of all, there is still room for improvement for some of the systems studied in the last few chapters. Specifically, for the nucleon case, it is certainly necessary to carry out the parity projection, especially for investigating the properties of the excited states. Furthermore, one should include the α_s corrections to the Wilson coefficients, which are known to be large and could therefore substantially alter the extracted spectral function. Such work is being done presently (Ohtani et al. 2013). Moreover, one could extend the analysis to other members of the light baryon octet (or to members of the decuplet), especially to study its excited states. Of special interest here is the excited state of Λ with negative parity ($\Lambda(1405)$), which is believed to be rather a meson-baryon molecule than a genuine three-quark state (Hyodo and Jido 2012).

As for the quarkonium systems at finite T, we have in this thesis only presented the results for the OPE including first order α_s terms and gluon condensates up to mass dimension 6. For making our predictions more reliable, it would be important to incorporate further α_s- and power-corrections to the OPE side of the sum rule. For the vector channel, higher order perturbative contributions have in fact already been calculated (Chetyrkin 1996) and can thus in principle be included into the sum rules (Ioffe and Zyablyuk 2003). On the other hand, making estimates about non-perturbative corrections requires detailed knowledge of several gluonic condensates, including their behavior at finite temperature. Such information can only be acquired from dedicated lattice QCD simulations and it will thus take some time until such improvements can really be implemented.

Besides the above mentioned improvements, one can of course also consider applications to other channels. Clearly, there are many possibilities and I can here only name only those which I personally believe to be most promising. An interesting application could for instance be the study of light vector mesons at finite density. This subject in fact already has quite a long history (Hatsuda and Lee 1992) and there have been vigorous debates on how to interpret the OPE results in terms of the spectral function (Koike 1995; Hatsuda et al. 1995; Leupold et al. 1998). Thus, it would certainly be of great interest to see what the MEM analysis has to say about this issue. Furthermore, one could also study various baryons with one, two or three heavy quarks, because the sum rules of these channels potentially contain rich information on both the low- and high-energy limit of the spectral function due to the new scale introduced by the heavy quark mass (we have already observed such a case in the quarkonium sum rules of Chap. 7). By doing such analyses, one can hope to get more insights about the heavy quark symmetry and new constraints for the condensates. Finally, let me also mention the study of exotics states containing four or more quarks (see for instance Nielsen et al. 2010 for a recent review). Close to nothing is known about the spectral functions of these channels and the "pole + continuum" assumption is hence highly questionable, especially as exotic states

lie in most cases above a continuum threshold and one thus should be very careful not to misidentify a mere continuum as a resonance peak. In such a situation, MEM could provide a useful guide for determining what kind of structures really exist in the spectral functions of interest.

To conclude, there are still a large number of subjects in this field that can be (or better, have to be) studied in the future and we hope that the method outlined in this thesis will be used as a tool for such investigations.

References

Chetyrkin KG, Kühn JH, Steinhauser M (1996) Nucl Phys B 482:213
Hatsuda T, Lee SH, Shiomi H (1995) Phys Rev C 52:3364
Hatsuda T, Lee SH (1992) Phys Rev C 46:34
Hyodo T, Jido D (2012) Prog Part Nucl Phys 67:55
Ioffe BL, Zyablyuk KN (2003) Eur Phys J C27:229
Koike Y (1995) Phys Rev C 51:1488
Leinweber DB (1997) Ann Phys 254:328
Leupold S, Peters W, Mosel U (1998) Nucl Phys A 628:311
Nielsen M, Lee SH, Navarra FS (2010) Phys Rept 497:41
Ohtani K, Gubler P, Oka M (2013) Phys Rev D 87:034027

Appendix A
The Dispersion Relation

The dispersion relation can be derived by calculating the contour integral of Fig. A.1. Using the Cauchy formula for the function $\Pi(s)$ (which contains a cut in the positive region of the real axis and is analytic in the rest of the complex plane of s), we can obtain

$$\begin{aligned}\Pi(q^2) &= \frac{1}{2\pi i}\oint_C ds \frac{\Pi(s)}{s-q^2} \\ &= \frac{1}{2\pi i}\oint_{|s|=R} ds \frac{\Pi(s)}{s-q^2} + \frac{1}{2\pi i}\int_0^R ds \frac{\Pi(s+i\varepsilon)-\Pi(s-i\varepsilon)}{s-q^2}, \end{aligned} \tag{A.1}$$

where R is the radius of the outer circle in Fig. A.1. If R is taken to infinity, the first integral in Eq. (A.1) vanishes if $\Pi(q^2)$ decreases sufficiently fast at $|q^2| \sim R \to \infty$. If this is not the case, subtraction terms have to be considered, as will be discussed below. The second integral in Eq. (A.1) can be simplified with the help of the Schwarz reflection principle (Morse and Feshbach 1953; Arfken 1970):

$$\Pi(s+i\varepsilon) - \Pi(s-i\varepsilon) = 2i\mathrm{Im}\Pi(s+i\varepsilon). \tag{A.2}$$

Let us examine this relation in detail. For this, we consider a complex function $f(z)$ (representing $\Pi(s)$) which is real on the real axis below a certain threshold point x_{th}, but possesses an imaginary part above x_{th}. Moreover, it is analytic and continuous on the upper half of the imaginary plane (we call this region D_1) except the region on the real axis above x_{th}. Now, we define the function $g(z)$, defined in the lower half of the imaginary plane (D_2) such that $g(z) = \overline{f(z)}$ (Here, the bar stands for complex conjugation). It can be easily shown that the real and imaginary parts of $g(z)$ satisfy the Cauchy-Riemann conditions, which means that it is an analytic function in D_2. Furthermore, $g(z)$ equals $f(z)$ on the real axis below x_{th}, because $f(z)$ is real there. Now, from the theory of analytic continuation of analytic functions, one can proof the following theorem (Morse and Feshbach 1953):

P. Gubler, *A Bayesian Analysis of QCD Sum Rules*, Springer Theses,
DOI: 10.1007/978-4-431-54318-3, © Springer Japan 2013

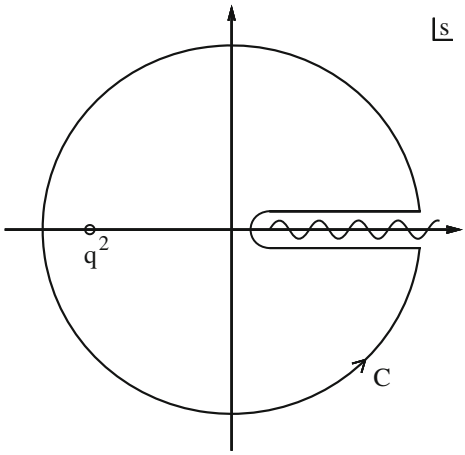

Fig. A.1 The contour integral C on the complex plane of the variable s, used for deriving the dispersion relation of Eq. (3.2). The *wavy line* denotes the non-analytic cut of $\Pi(s)$ on the positive side of the real axis. Note that here q^2 takes a negative value

If f is analytic in D_1 and g in D_2, if f equals g along their common boundary A, and if f and g are continuous along A, then g is the continuation of f in D_2 and vice versa.

In our current setting, this theorem immediately leads to the Schwarz reflexion principle, which states that $f(\overline{z}) = \overline{f(z)}$ and holds for the whole imaginary plane except non-analytic part on the real axis above $x_{\rm th}$.

Therefore, for $x > x_{\rm th}$ and ε being an infinitesimal real constant, we have

$$\begin{aligned} f(x - i\varepsilon) &= \overline{f(x + i\varepsilon)} \\ &= f(x + i\varepsilon) - 2i\,\mathrm{Im}\, f(x + i\varepsilon), \end{aligned} \tag{A.3}$$

which corresponds to Eq. (A.2). Thus we have finally obtained the dispersion relation of Eq. (3.2).

Next, let us consider the case, in which the integral Eq. (3.2) diverges. This problem can be fixed by using subtracted correlators as shown below. For instance, if the divergence is logarithmic, it suffices to employ the singly subtracted correlator:

$$\begin{aligned} \tilde{\Pi}(q^2) &\equiv \Pi(q^2) - \Pi(0) \\ &= \frac{q^2}{\pi} \int_0^\infty ds \frac{\mathrm{Im}\Pi(s + i\varepsilon)}{s(s - q^2)}. \end{aligned} \tag{A.4}$$

In this way, one power of s can be included into the denominator, therefore making the integral convergent. This procedure can be repeated arbitrarily many times, by subtracting more and more terms from the Taylor expansion of $\Pi(q^2)$ around $q^2 = 0$, thus it is possible to cure divergences of any power. We however note, that by applying the Borel transformation, which contains infinitely many differentiations of q^2, all subtraction terms (which are polynomials in q^2) vanish and the integral of Eq. (3.2) is automatically turned into convergent one. Therefore, as long as the Borel

transformation is applied, one usually does not have to worry about divergences and subtraction terms and can work directly with Eq. (3.2), the original form of the dispersion relation.

Appendix B
The Fock-Schwinger Gauge

In this appendix, the derivative expansions of the gluonic and quark fields, given in Eqs. (3.26) and (3.27), are derived. These expressions are valid only in the Fock-Schwinger gauge (Fock 1937; Schwinger 1954) (sometimes also referred to as the "fixed point gauge"),

$$(x - x_0)^\mu A^a_\mu(x) = 0, \tag{B.1}$$

where we set $x_0 = 0$ in the following. The derivation is based on the discussions given in Dubovikov and Smilga (1981) and Shifman (1980).

Multiplying $G^a_{\mu\nu}$, which is defined as

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f_{abc} A^b_\mu A^c_\nu \tag{B.2}$$

by x_μ and using Eq. (B.1), we get (with $x_0 = 0$)

$$x_\mu G^a_{\mu\nu}(x) = x^\mu \partial_\mu A^a_\nu(x) + A^a_\nu(x). \tag{B.3}$$

Then, x is replaced by αx after which we integrate by α:

$$\begin{aligned}\int_0^1 d\alpha \alpha x^\mu G^a_{\mu\nu}(\alpha x) &= \int_0^1 d\alpha \alpha \frac{d}{d\alpha} A^a_\nu(\alpha x) + \int_0^1 d\alpha A^a_\nu(\alpha x) \\ &= A^a_\nu(x).\end{aligned} \tag{B.4}$$

Here, the first line has been obtained by using

$$\frac{d}{d\alpha} A^a_\nu(\alpha x) = x^\mu \frac{\partial}{\partial(\alpha x^\mu)} A^a_\nu(\alpha x) = \frac{x^\mu}{\alpha} \partial x^\mu A^a_\nu(\alpha x). \tag{B.5}$$

Taylor expanding $G^a_{\mu\nu}(\alpha x)$ around $\alpha x = 0$ on the left hand side of Eq. (B.4) and carrying out the integration of α, we arrive at

P. Gubler, *A Bayesian Analysis of QCD Sum Rules*, Springer Theses,
DOI: 10.1007/978-4-431-54318-3, © Springer Japan 2013

$$A^a_\nu(x) = \frac{1}{2}x^\mu G^a_{\mu\nu}(0) + \frac{1}{3}x^\mu x^\alpha \partial_\alpha G^a_{\mu\nu}(0) + \frac{1}{8}x^\mu x^\alpha x^\beta \partial_\alpha \partial_\beta G^a_{\mu\nu}(0) + \cdots \quad \text{(B.6)}$$

For obtaining the final result, we have to show that the derivative ∂ can be replaced by the covariant derivative D in the above equation. The Fock-Schwinger gauge actually makes this substitution possible, as will be shown below.

The Taylor expansion of Eq. (B.1)

$$x^\mu A^a_\mu(x) = x^\mu \left[A^a_\mu(0) + x^\alpha \partial_\alpha A^a_\mu(0) + \frac{1}{2}x^\alpha x^\beta \partial_\alpha \partial\beta A^a_\mu(0) + \cdots \right] = 0, \quad \text{(B.7)}$$

which has to be valid for any value of x_μ, leads to the equations

$$\begin{aligned} x^\mu A^a_\mu(0) &= 0, \\ x^\mu x^\alpha \partial_\alpha A^a_\mu(0) &= 0, \\ &\cdots \end{aligned} \quad \text{(B.8)}$$

As the covariant derivative applied to gluon fields is defined as

$$D_\mu = \partial_\mu - igT^a A^a_\mu, \quad \text{(B.9)}$$

where T^a are the generators of $SU(3)$ in the adjoint representation and g is the strong coupling constant, we can derive the following relations:

$$\begin{aligned} x^\alpha \partial_\alpha G^a_{\mu\nu}(0) &= x^\alpha D_\alpha G^a_{\mu\nu}(0), \\ x^\alpha x^\beta \partial_\alpha \partial_\beta G^a_{\mu\nu}(0) &= x^\alpha x^\beta \partial_\alpha D_\beta G^a_{\mu\nu}(0) = x^\alpha x^\beta D_\alpha D_\beta G^a_{\mu\nu}(0), \\ &\cdots \end{aligned} \quad \text{(B.10)}$$

This shows that the derivatives can be substituted by the covariant derivatives, giving us thus the final result:

$$\begin{aligned} A^a_\mu(x) = \frac{1}{2}x^\nu G^a_{\nu\mu}(0) + \frac{1}{3}x^\nu x^\alpha [D_\alpha G_{\nu\mu}(0)]^a \\ + \frac{1}{8}x^\nu x^\alpha x^\beta [D_\alpha D_\beta G_{\nu\mu}(0)]^a + \cdots \end{aligned} \quad \text{(B.11)}$$

Next, we consider the quark fields. For this purpose, we simply Taylor expand the field $q(x)$ around $x = 0$, giving

$$q(x) = q(0) + x^\mu \partial_\mu q(0) + \frac{1}{2!}x^\nu x^\mu \partial_\nu \partial_\mu q(0) + \cdots \quad \text{(B.12)}$$

Now, the relations of Eq. (B.10) are valid also for the covariant derivative living in the fundamental representation,

$$D_\mu = \partial_\mu - ig\frac{\lambda^a}{2}A^a_\mu, \tag{B.13}$$

in which λ^a are the Gell-Mann matrices. Therefore, as for the gluonic fields above, we can simply interchange the derivatives of Eq. (B.12) with the covariant derivative, leading to the desired result:

$$q(x) = q(0) + x^\mu D_\mu q(0) + \frac{1}{2!}x^\nu x^\mu D_\nu D_\mu q(0) + \cdots \tag{B.14}$$

Appendix C
The Quark Propagator

To calculate the free quark propagator with no coupling to gluons and no long range correlations is most simple. It is given in standard textbooks of quantum field theory (such as Peskin and Schroeder 1995) and we here state only the result:

$$\begin{aligned}\langle 0_{\text{pert.}}|T[q(x)\overline{q}(0)]|0_{\text{pert.}}\rangle \equiv S_0(x) &= \int d^4p e^{-ipx}\frac{i}{\not{p}-m_q} \\ &\approx \int d^4p e^{-ipx}\left(\frac{i}{\not{p}}+\frac{im_q}{p^2}\right)+\mathcal{O}(m_q^2) \\ &\approx \frac{i}{2\pi^2}\frac{\not{x}}{x^4}-\frac{m_q}{4\pi^2}\frac{1}{x^2}+\mathcal{O}(m_q^2)\end{aligned} \tag{C.1}$$

$|0_{\text{pert.}}\rangle$ stands for the perturbative vacuum, where all condensates and expectation values of matter fields vanish. The last line of the above equation gives us the first two terms of Eq. (3.28).

C.1 Coupling with Gluon Fields

Here, the behavior of the quark propagator $S_A(x)$ in an external gluon field is discussed. Such a propagator satisfies the equation

$$(i\not{\partial}+g\not{A}-m_q)S_A(x)=i\delta^4(x), \tag{C.2}$$

and is expanded in powers of the external field $A^a_\mu(x)$. Expressed in Feynman diagrams, this expansion is shown in Fig. C.1, while mathematically it is given as

$$\begin{aligned}S_A(x) = S_0(x) &+ \int d^4y S_0(x-y)ig\not{A}(y)S_0(y) \\ &+ \int d^4y d^4z S_0(x-y)ig\not{A}(y)S_0(y-z)ig\not{A}(z)S_0(z)+\cdots\end{aligned} \tag{C.3}$$

P. Gubler, *A Bayesian Analysis of QCD Sum Rules*, Springer Theses,
DOI: 10.1007/978-4-431-54318-3, © Springer Japan 2013

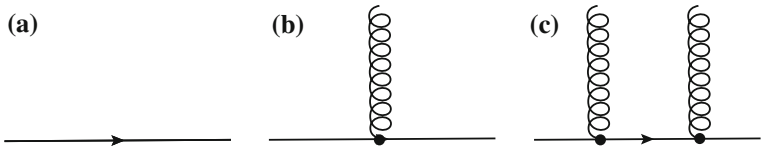

Fig. C.1 The first three diagrams of the external field expansion. **a** has already been calculated in Eq.(C.1). Here, we calculate **b** while **c** and the other higher order terms are neglected

Switching to the momentum representation and expressing the gluon field with the first term of Eq. (3.26), we get for the second term (Fig. C.1b):

$$
\begin{aligned}
&\int d^4y S_0(x-y) ig \not{A}(y) S_0(y) \\
&\approx \int d^4y \int \frac{d^4p}{(2\pi)^4} \int \frac{d^4q}{(2\pi)^4} e^{-ip(x-y)} e^{-iqy} \frac{i(\not{p}+m_q)}{p^2-m_q^2} ig\gamma^\mu \\
&\quad \times \left(\frac{y^\alpha}{2} G_{\alpha\mu}(0) \right) \frac{i(\not{q}+m_q)}{q^2-m_q^2} \\
&= \int \frac{d^4p}{(2\pi)^4} e^{-ipx} \frac{i(\not{p}+m_q)}{p^2-m_q^2} ig\gamma^\mu \left(\frac{i}{2} G_{\alpha\mu}(0) \right) \frac{\partial}{\partial q_\alpha} \left(\frac{i(\not{q}+m_q)}{q^2-m_q^2} \right) \Bigg|_{q=p} \\
&= -\frac{i}{4} g G^{\mu\nu}(0) \int \frac{d^4p}{(2\pi)^4} e^{-ipx} \frac{\sigma_{\mu\nu}(\not{p}+m_q) + (\not{p}+m_q)\sigma_{\mu\nu}}{(p^2-m_q^2)^2} \qquad \text{(C.4)} \\
&\approx -\frac{i}{4} g G^{\mu\nu}(0) \int \frac{d^4p}{(2\pi)^4} e^{-ipx} \frac{\sigma_{\mu\nu}\ \not{p} + \not{p}\sigma_{\mu\nu}}{p^4} \\
&\quad - \frac{i}{2} g m_q G^{\mu\nu}(0) \sigma_{\mu\nu} \int \frac{d^4p}{(2\pi)^4} \frac{1}{p^4} + \mathcal{O}(m_q^2) \\
&= -\frac{i}{32\pi^2} g G^{\mu\nu}(0) \frac{\sigma_{\mu\nu}\ \not{x} + \not{x}\sigma_{\mu\nu}}{x^2} \\
&\quad - \frac{1}{32\pi^2} g m_q G^{\mu\nu}(0) \sigma_{\mu\nu} \ln\left(-\frac{x^2\Lambda^2}{4} + 2\gamma_{EM} \right) + \mathcal{O}(m_q^2).
\end{aligned}
$$

This results provides us with terms number three and four of Eq. (3.28).

C.2 Non-Perturbative Contributions

As a next step, we have to consider long range fluctuations of quarks and gluons in the quark propagator, which are expressed by various condensates such as $\langle \overline{q}q \rangle$, $\langle \overline{q} g\sigma G q \rangle$ or $\langle \frac{\alpha_s}{\pi} G^2 \rangle$.

For this, we make use of Eq. (3.27), substitute it into

$$\langle q_i^a(x)\overline{q}_j^b(0)\rangle, \tag{C.5}$$

and investigate its first few terms one after the other. For making the following manipulations easier and more tractable, we have here explicitly denoted the color and spinor indices as a, b and i, j, respectively. Note also that the non-perturbative components of the quark (and gluon) fields behave as classical fields that satisfy the equations of motion. Therefore, is valid to omit the time ordering operator $T[\dots]$.

For the first term, since it is sandwiched between the vacuum, only the scalar and color-singlet component survives. We can hence write

$$\langle q_i^a(0)\overline{q}_j^b(0)\rangle = A\delta^{ab}\delta_{ij}. \tag{C.6}$$

Taking the contractions of color and spinor indices on both sides and using the fact that quarks are Fermions and therefore anti-symmetric, we get

$$A = -\frac{1}{12}\langle\overline{q}q\rangle \tag{C.7}$$

for A.

The second term can be expressed as

$$x^\mu\left\langle\left[D_\mu q_i(0)\right]^a\overline{q}_j^b(0)\right\rangle = x^\mu\frac{\delta^{ab}}{3}\langle D_\mu q_i(0)\overline{q}_j(0)\rangle, \tag{C.8}$$

because, like above, only the color-singlet term survives. Then, the Dirac indices i, j are expanded with the complete set of 1, γ^5, γ_ν, $\gamma^5\gamma_\nu$, $\sigma_{\nu\rho}$, which gives

$$\begin{aligned} x^\mu\langle\left[D_\mu q_i(0)\right]^a\overline{q}_j^b(0)\rangle = x^\mu\frac{\delta^{ab}}{3}\Bigg(&-\frac{\delta_{ij}}{4}\langle\overline{q}(0)D_\mu q(0)\rangle - \frac{\gamma^5_{ij}}{4}\langle\overline{q}(0)\gamma^5 D_\mu q(0)\rangle \\ &- \frac{\gamma^\nu_{ij}}{4}\langle\overline{q}(0)\gamma_\nu D_\mu q(0)\rangle + \frac{(\gamma^5\gamma^\nu)_{ij}}{4}\langle\overline{q}(0)\gamma^5\gamma_\nu D_\mu q(0)\rangle \\ &- \frac{\sigma^{\nu\rho}_{ij}}{4}\langle\overline{q}(0)\sigma_{\nu\rho}D_\mu q(0)\rangle\Bigg). \end{aligned} \tag{C.9}$$

In this equation, only the third term can have the same quantum numbers as the vacuum and thus all the other terms vanish. Moreover, the scalar component of $\gamma_\nu D_\mu$ can be obtained as $\frac{g_{\nu\mu}}{4}\not{D}$, which leads to

$$x^\mu\langle\left[D_\mu q_i(0)\right]^a\overline{q}_j^b(0)\rangle = -x^\mu\frac{\delta^{ab}}{48}(\gamma_\mu)_{ij}\langle\overline{q}(0)\not{D}q(0)\rangle. \tag{C.10}$$

The equation of motion $\not{D}q = -im_q q$ is then used to derive the final result:

$$x^\mu \langle \left[D_\mu q_i(0)\right]^a \overline{q}_j^b(0)\rangle = \frac{im_q}{48}(\not{x})_{ij}\delta^{ab}\langle \overline{q}q\rangle. \tag{C.11}$$

Next, we consider the third term of Eq. (3.27). As before, the only the color singlet part has to be retained and therefore

$$\frac{1}{2}x^\mu x^\nu \langle \left[D_\mu D_\nu q_i(0)\right]^a \overline{q}_j^b(0)\rangle = \frac{\delta^{ab}}{6}x^\mu x^\nu \langle D_\mu D_\nu q_i(0)\overline{q}_j(0)\rangle. \tag{C.12}$$

Then, we do the same as in Eq. (C.9) and expand the Dirac indices. As is readily understood, only the component proportional to δ_{ij} survives because it is the only one containing a scalar part with positive parity. We thus get

$$\frac{1}{2}x^\mu x^\nu \langle \left[D_\mu D_\nu q_i(0)\right]^a \overline{q}_j^b(0)\rangle = -\frac{\delta_{ij}\delta^{ab}}{24}x^\mu x^\nu \langle \overline{q}(0)D_\mu D_\nu q(0)\rangle. \tag{C.13}$$

Finally, using the fact that the scalar part of $D_\mu D_\nu$ is $\frac{g_{\mu\nu}}{4}D^2$ and the equation $D^2 = \frac{1}{2}g\sigma Gq - m_q^2 q$, which is easily derived from the equation of motion, the following result is obtained:

$$\begin{aligned}\tfrac{1}{2}x^\mu x^\nu \left\langle \left[D_\mu D_\nu q_i(0)\right]^a \overline{q}_j^b(0)\right\rangle &= -\tfrac{x^2}{96}\delta_{ij}\delta^{ab}\langle \overline{q}(0)D^2 q(0)\rangle \\ &\approx -\tfrac{x^2}{192}\delta_{ij}\delta^{ab}\langle \overline{q}g\sigma Gq\rangle + \mathcal{O}(m_q^2).\end{aligned} \tag{C.14}$$

The derivation of the fourth term and the fifth term of Eq. (3.27) is more involved, although the basic techniques are essentially the same. Here, only the results are stated:

$$\frac{1}{6}x^\mu x^\nu x^\rho \left\langle D_\mu D_\nu D_\rho q_i^a(0)\overline{q}_j^b(0)\right\rangle \approx \frac{im_q x^2}{2^7 3^2}(\not{x})_{ij}\delta^{ab}\langle \overline{q}g\sigma Gq\rangle + \mathcal{O}(m_q^2, g^2), \tag{C.15}$$

$$\begin{aligned}\tfrac{1}{24}x^\mu x^\nu x^\rho x^\sigma \left\langle D_\mu D_\nu D_\rho D_\sigma q_i^a(0)\overline{q}_j^b(0)\right\rangle &\approx -\tfrac{\pi^2 x^4}{2^8 3^3}\delta_{ij}\delta^{ab}\langle \overline{q}q\rangle \left\langle \tfrac{\alpha_s}{\pi}G^2\right\rangle \\ &\quad + \mathcal{O}(m_q^2, g^2).\end{aligned} \tag{C.16}$$

The explicit derivation of Eq. (C.15) can be found in Chap. 6 of Ioffe et al. (2010). Furthermore, we note that for deriving Eq. (C.16), in addition to the method explained above, the vacuum saturation approximation has been assumed and the contraction formula for gluon fields

$$\langle G^k_{\mu\nu} G^l_{\rho\sigma} \rangle = \frac{\delta^{kl}}{2^5 3} (g_{\mu\rho} g_{\nu\sigma} - g_{\mu\sigma} g_{\nu\rho}) \langle G^2 \rangle \tag{C.17}$$

has been used.

Altogether, the results of this section give the remaining non-perturbative terms of Eq. (3.28).

Appendix D
Non-Perturbative Coupling of Quarks and Gluons

In this appendix, we derive the form of the non-perturbative coupling between quarks and gluons, given in Eq. (3.30). Our starting point is the following expression,

$$\begin{aligned}\left\langle 0|q_i^a(x)gG_{\mu\nu}^k(0)\overline{q}_j^b(0)|0\right\rangle &= \left\langle 0|q_i^a(0)gG_{\mu\nu}^k(0)\overline{q}_j^b(0)|0\right\rangle \\ &\quad + x^\alpha\left\langle 0|D_\alpha q_i^a(0)gG_{\mu\nu}^k(0)\overline{q}_j^b(0)|0\right\rangle + \cdots, \end{aligned} \tag{D.1}$$

for which we treat each of the two terms separately.

D.1 The First Term

In contrast to the calculations in the preceding appendix, one here has to build a color octet from the quarks, which is then contracted with the gluon for constructing an overall color-singlet operator. Therefore, for the first term, we get

$$\begin{aligned}\langle 0|q_i^a(0)gG_{\mu\nu}^k(0)\overline{q}_j^b(0)|0\rangle &= -2\left(\frac{\lambda^l}{2}\right)^{ab}\left\langle 0|\overline{q}_j(0)\left(\frac{\lambda^l}{2}\right)gG_{\mu\nu}^k(0)q_i(0)|0\right\rangle \\ &= -\frac{1}{4}\left(\frac{\lambda^k}{2}\right)^{ab}\langle 0|\overline{q}_j(0)gG_{\mu\nu}(0)q_i(0)|0\rangle, \end{aligned} \tag{D.2}$$

after which the spinor indices are expanded as in Eq. (C.9). We then obtain

$$\begin{aligned}\left\langle 0|q_i^a(0)gG_{\mu\nu}^k(0)\overline{q}_j^b(0)|0\right\rangle &= -\frac{(\sigma_{\rho\sigma})_{ij}}{32}\left(\frac{\lambda^k}{2}\right)^{ab}\left\langle 0|\overline{q}(0)\sigma^{\rho\sigma}gG_{\mu\nu}(0)q(0)|0\right\rangle \\ &= -\frac{(\sigma_{\mu\nu})_{ij}}{2^6 3}\left(\frac{\lambda^k}{2}\right)^{ab}\langle\overline{q}g\sigma Gq\rangle, \end{aligned} \tag{D.3}$$

which is the final result.

P. Gubler, *A Bayesian Analysis of QCD Sum Rules*, Springer Theses,
DOI: 10.1007/978-4-431-54318-3, © Springer Japan 2013

D.2 The Second Term

The second term of Eq. (D.1) can be calculated in the same way, it is however somewhat more complicated. First, we construct a color octet from the quark fields and combine it with the octet from the gluon, as before:

$$\begin{aligned} & x^\alpha \langle 0|D_\alpha q_i^a(0) g G_{\mu\nu}^k(0) \overline{q}_j^b(0)|0\rangle \\ &= -2\left(\frac{\lambda^l}{2}\right)^{ab} x^\alpha \langle 0|\overline{q}_j(0)\left(\frac{\lambda^l}{2}\right) g G_{\mu\nu}^k(0) D_\alpha q_i(0)|0\rangle \\ &= -\frac{1}{4}\left(\frac{\lambda^k}{2}\right)^{ab} x^\alpha \langle 0|\overline{q}_j(0) g G_{\mu\nu}(0) D_\alpha q_i(0)|0\rangle. \end{aligned} \tag{D.4}$$

Next, we expand the spinor indices of the quark fields. Doing this, it is clear that only the terms with γ_μ or $\gamma_5\gamma_\mu$ can survive, because from all other terms it is not possible to construct a scalar operator. Thus we have

$$\begin{aligned} \langle 0|\overline{q}_j(0) g G_{\mu\nu}(0) D_\alpha q_i(0)|0\rangle = & \frac{(\gamma^\beta)_{ij}}{4}\langle 0|\overline{q}(0)\gamma_\beta g G_{\mu\nu}(0) D_\alpha q(0)|0\rangle \\ & -\frac{(\gamma_5\gamma^\beta)_{ij}}{4}\langle 0|\overline{q}(0)\gamma_5\gamma_\beta g G_{\mu\nu}(0) D_\alpha q(0)|0\rangle. \end{aligned} \tag{D.5}$$

Subsequently, we expand the remaining parts into their possible Lorentz structures. Parity considerations tell us that the first term can only be proportional to $g_{\mu\nu}g_{\alpha\beta}$, $g_{\mu\alpha}g_{\nu\beta}$ or $g_{\mu\beta}g_{\alpha\nu}$ and the second term only to $\varepsilon_{\mu\nu\alpha\beta}$, giving

$$\langle 0|\overline{q}(0)\gamma_\beta g G_{\mu\nu}(0) D_\alpha q(0)|0\rangle = A g_{\mu\nu}g_{\alpha\beta} + B g_{\mu\alpha}g_{\nu\beta} + C g_{\mu\beta}g_{\alpha\nu}, \tag{D.6}$$

and

$$\langle 0|\overline{q}(0)\gamma_5\gamma_\beta g G_{\mu\nu}(0) D_\alpha q(0)|0\rangle = D\varepsilon_{\mu\nu\alpha\beta}. \tag{D.7}$$

Contracting Eq. (D.6) with $g^{\mu\nu}g^{\alpha\beta}$, $g^{\mu\alpha}g^{\nu\beta}$ and $g^{\mu\beta}g^{\alpha\nu}$, we get three equations, which lead to

$$\begin{aligned} A &= 0, \\ B &= -\frac{1}{12}\langle 0|\overline{q}(0)\gamma^\rho g G_{\rho\sigma}(0) D^\sigma q(0)|0\rangle, \\ C &= \frac{1}{12}\langle 0|\overline{q}(0)\gamma^\rho g G_{\rho\sigma}(0) D^\sigma q(0)|0\rangle. \end{aligned} \tag{D.8}$$

On the other hand, contracting Eq. (D.7) with $\varepsilon^{\mu\nu\alpha\beta}$, we obtain

$$D = \frac{1}{24}\langle 0|\overline{q}(0)\gamma_5\varepsilon^{\mu\nu\alpha\beta}\gamma_\beta g G_{\mu\nu}(0) D_\alpha q(0)|0\rangle, \tag{D.9}$$

which can be rearranged by Eq. (E.21) of Appendix E and the equation of motion for quarks ($\not{D}q = -im_q q$). This gives

$$D = \frac{i}{12}\langle 0|\overline{q}(0)\gamma^\rho g G_{\rho\sigma}(0) D^\sigma q(0)|0\rangle + \frac{im_q}{24}\langle 0|\overline{q}(0) g\sigma G(0) q(0)|0\rangle. \tag{D.10}$$

Now, all we need is an expression of $\langle 0|\overline{q}(0)\gamma^\rho g G_{\rho\sigma}(0) D^\sigma q(0)|0\rangle$ in form of known condensates. The details of the manipulations necessary for this task are explained in Chap. 6 of Ioffe et al. (2010) and we here give only the final result:

$$\begin{aligned}&\langle 0|\overline{q}(0)\gamma^\rho g G_{\rho\sigma}(0) D^\sigma q(0)|0\rangle \\ &= -\frac{1}{2}\langle 0|\overline{q}(0)\gamma^\rho \left(\frac{\lambda^n}{2}\right) q(0) g\big[D^\sigma G_{\rho\sigma}(0)\big]^n|0\rangle \\ &\quad - \frac{m_q}{2}\langle 0|\overline{q}(0) g\sigma G(0) q(0)|0\rangle.\end{aligned} \tag{D.11}$$

The first term on the right hand side of the above equation can be rewritten using the equation of motion for gluons. This gives a term proportional to g^2, which we neglect here. If one wants to calculate higher orders of α_s, it however has to be retained.

Assembling the results of Eqs. (D.6)–(D.11) we finally get

$$\langle 0|\overline{q}_j(0) g G_{\mu\nu}(0) D_\alpha q_i(0)|0\rangle = \frac{m_q}{2^5 3}\big[(\gamma_\nu)_{ij} g_{\alpha\mu} - (\gamma_\mu)_{ij} g_{\alpha\nu}\big]\langle 0|\overline{q} g\sigma G q|0\rangle, \tag{D.12}$$

which leads to

$$\begin{aligned}&x^\alpha \langle 0| D_\alpha q_i^a(0) g G_{\mu\nu}^k(0) \overline{q}_j^b(0)|0\rangle \\ &\approx -\frac{m_q}{2^7 3}\left(\frac{\lambda^k}{2}\right)^{ab}\big[(\gamma_\nu)_{ij} x_\mu - (\gamma_\mu)_{ij} x_\nu\big]\langle 0|\overline{q} g\sigma G q|0\rangle + \mathcal{O}(g^2).\end{aligned} \tag{D.13}$$

The spinor part of this result $(\gamma_\nu x_\mu - \gamma_\mu x_\nu)$ can be further manipulated according to formula of Eqs. (E.21) and (E.22) in Appendix E as follows

$$\begin{aligned}\gamma_\nu x_\mu - \gamma_\mu x_\nu &= -x^\lambda(g_{\nu\lambda}\gamma_\mu - g_{\mu\lambda}\gamma_\nu) \\ &= -x^\lambda(i\varepsilon_{\mu\nu\lambda\rho}\gamma_5\gamma^\rho - g_{\mu\nu}\gamma_\lambda + \gamma_\mu\gamma_\nu\gamma_\lambda) \\ &= -\frac{i}{2}(\not{x}\sigma_{\mu\nu} + \sigma_{\mu\nu}\not{x}) + i\sigma_{\mu\nu}\not{x} \\ &= -\frac{i}{2}(\not{x}\sigma_{\mu\nu} - \sigma_{\mu\nu}\not{x}),\end{aligned} \tag{D.14}$$

which gives

$$x^{\alpha}\langle 0|D_{\alpha}q_i^a(0)gG_{\mu\nu}^k(0)\overline{q}_j^b(0)|0\rangle \approx \frac{im_q}{2^83}(\not{x}\sigma^{\mu\nu} - \sigma^{\mu\nu}\not{x})_{ij}\left(\frac{\lambda^k}{2}\right)^{ab}\langle\overline{q}g\sigma Gq\rangle + \mathcal{O}(g^2). \tag{D.15}$$

The final form of the non-perturbative coupling of quarks and gluons is then

$$\langle 0|q_i^a(x)gG_{\mu\nu}^k(0)\overline{q}_j^b(0)|0\rangle \approx -\frac{(\sigma_{\mu\nu})_{ij}}{2^63}\left(\frac{\lambda^k}{2}\right)^{ab}\langle\overline{q}g\sigma Gq\rangle + \frac{im_q}{2^83}(\not{x}\sigma^{\mu\nu} - \sigma^{\mu\nu}\not{x})_{ij}\left(\frac{\lambda^k}{2}\right)^{ab}\langle\overline{q}g\sigma Gq\rangle + \mathcal{O}(m_q^2, g^2), \tag{D.16}$$

which corresponds to Eq. (3.30) of the main text.

Appendix E
Gamma Matrix Algebra

When doing calculations in the QCD sum rule technique, various properties of gamma matrices are frequently used. A few of the most convenient formulae concerning these gamma matrices are given in this appendix. Note that we here use the convention $\varepsilon^{0123} = 1$ for the totally antisymmetric Levi-Civita tensor.

$$\{\gamma_\mu, \gamma_\nu\} = 2g_{\mu\nu} \tag{E.1}$$
$$\{\gamma_\mu, \gamma_5\} = 0 \tag{E.2}$$

$$\sigma_{\mu\nu} \equiv \frac{i}{2}[\gamma_\mu, \gamma_\nu] \tag{E.3}$$
$$\gamma_\mu\gamma_\nu = g_{\mu\nu} - i\sigma_{\mu\nu} \tag{E.4}$$
$$\gamma_5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{E.5}$$

$$C \equiv i\gamma^2\gamma^0 \text{ (charge conjugation matrix)} \tag{E.6}$$
$$C = C^* = -C^\dagger = -C^T = -C^{-1} \tag{E.7}$$
$$C^2 = -1 \tag{E.8}$$
$$C\gamma_5 = \gamma_5 C \tag{E.9}$$

$$C\Gamma^T C = +\Gamma \quad \text{for} \quad \Gamma = \gamma_\mu,\ \sigma_{\mu\nu},\ \gamma_5\sigma_{\mu\nu} \tag{E.10}$$
$$C\Gamma^T C = -\Gamma \quad \text{for} \quad \Gamma = \gamma_5,\ \gamma_5\gamma_\mu,\ (\not{x}\sigma^{\mu\nu} + \sigma^{\mu\nu}\not{x}) \tag{E.11}$$

$$\gamma^\mu\gamma_\nu\gamma_\mu = -2\gamma_\nu \tag{E.12}$$
$$\gamma^\mu\gamma_\alpha\gamma_\beta\gamma_\mu = 4g_{\alpha\beta} \tag{E.13}$$

P. Gubler, *A Bayesian Analysis of QCD Sum Rules*, Springer Theses,
DOI: 10.1007/978-4-431-54318-3, © Springer Japan 2013

$$\gamma^\mu \gamma_\alpha \gamma_\beta \gamma_\gamma \gamma_\mu = -2\gamma_\gamma \gamma_\beta \gamma_\alpha \tag{E.14}$$

$$\sigma^{\alpha\beta}\sigma_{\alpha\beta} = 12 \tag{E.15}$$

$$\sigma^{\alpha\beta}\gamma^\mu\gamma^\nu\sigma_{\alpha\beta} = 4\gamma^\nu\gamma^\mu + 8g^{\mu\nu} = 16g^{\mu\nu} - 4\gamma^\mu\gamma^\nu \tag{E.16}$$

$$\sigma^{\alpha\beta}(\text{odd number of } \gamma\text{-matrices})\sigma_{\alpha\beta} = 0 \tag{E.17}$$

$$(\not{x}\sigma^{\alpha\beta} + \sigma^{\alpha\beta}\not{x})(\not{x}\sigma_{\alpha\beta} + \sigma_{\alpha\beta}\not{x}) = 24x^2 \tag{E.18}$$

$$(\not{x}\sigma^{\alpha\beta} + \sigma^{\alpha\beta}\not{x})\gamma^\mu(\not{x}\sigma_{\alpha\beta} + \sigma_{\alpha\beta}\not{x}) = 8(x^2\gamma^\mu + 2x^\mu\not{x}) \tag{E.19}$$

$$\begin{aligned}(\not{x}\sigma^{\alpha\beta} + \sigma^{\alpha\beta}\not{x})\gamma^\mu\gamma^\nu(\not{x}\sigma_{\alpha\beta} + \sigma_{\alpha\beta}\not{x}) = \\ 8(4x^2 g^{\mu\nu} - 2x^\mu\gamma^\nu\not{x} + 2x^\nu\gamma^\mu\not{x} - x^2\gamma^\mu\gamma^\nu)\end{aligned} \tag{E.20}$$

$$\varepsilon_{\mu\nu\lambda\rho}\gamma^\rho = -i\gamma_5(g_{\mu\nu}\gamma_\lambda - g_{\mu\lambda}\gamma_\nu + g_{\nu\lambda}\gamma_\mu - \gamma_\mu\gamma_\nu\gamma_\lambda) \tag{E.21}$$

$$\not{x}\sigma_{\mu\nu} + \sigma_{\mu\nu}\not{x} = -2\varepsilon_{\mu\nu\alpha\beta}\gamma_5\gamma^\alpha x^\beta \tag{E.22}$$

$$\sigma_{\mu\nu} = \frac{i}{2}\varepsilon_{\mu\nu\rho\lambda}\gamma_5\sigma^{\rho\lambda} \tag{E.23}$$

$$\text{Tr}[\gamma_\mu\gamma_\nu] = 4g_{\mu\nu} \tag{E.24}$$

$$\text{Tr}[\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\lambda] = 4(g_{\mu\nu}g_{\rho\lambda} - g_{\mu\rho}g_{\nu\lambda} + g_{\mu\lambda}g_{\nu\rho}) \tag{E.25}$$

$$\text{Tr}[\text{odd number of } \gamma\text{-matrices}] = 0 \tag{E.26}$$

$$\text{Tr}[\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma\cdots] = \text{Tr}[\cdots\gamma_\sigma\gamma_\rho\gamma_\nu\gamma_\mu] \tag{E.27}$$

$$\text{Tr}[\gamma_5\gamma_\mu\gamma_\nu] = 0 \tag{E.28}$$

$$\text{Tr}[\gamma_5\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma] = -4i\varepsilon_{\mu\nu\rho\sigma} \tag{E.29}$$

$$\text{Tr}[\gamma_\mu\not{x}\gamma_\nu\not{x}] = 8x_\mu x_\nu - 4x^2 g_{\mu\nu} \tag{E.30}$$

$$\begin{aligned}\text{Tr}[\gamma_\mu(\not{x}\sigma_{\rho\lambda} + \sigma_{\rho\lambda}\not{x})\gamma_\nu\not{x}] = &-8ix^2(g_{\rho\mu}g_{\lambda\nu} - g_{\rho\nu}g_{\lambda\mu}) \\ &+ 8ix_\mu(x_\rho g_{\lambda\nu} - x_\lambda g_{\rho\nu}) \\ &+ 8ix_\nu(x_\lambda g_{\rho\mu} - x_\rho g_{\lambda\mu})\end{aligned} \tag{E.31}$$

$$\text{Tr}[\sigma_{\rho\lambda}\gamma_5\not{x}\gamma_\mu] = 4\varepsilon_{\alpha\rho\lambda\mu}x^\alpha \tag{E.32}$$

$$\text{Tr}[\sigma_{\rho\lambda}\not{x}\gamma_\mu] = 4i(g_{\rho\mu}x_\lambda - g_{\lambda\mu}x_\rho) \tag{E.33}$$

$$\text{Tr}[\sigma_{\mu\nu}\sigma_{\rho\lambda}] = 4(g_{\mu\rho}g_{\nu\lambda} - g_{\mu\lambda}g_{\nu\rho}) \tag{E.34}$$

$$\text{Tr}[(\not{x}\sigma_{\mu\nu} + \sigma_{\mu\nu}\not{x})(\not{x}\sigma_{\rho\lambda} + \sigma_{\rho\lambda}\not{x})] = -16\varepsilon_{\sigma\mu\nu\alpha}\varepsilon^\sigma{}_{\rho\lambda\beta}x^\alpha x^\beta \tag{E.35}$$

$$\varepsilon_{\sigma\mu\nu\alpha}\varepsilon^{\sigma}{}_{\rho\lambda\beta} = g_{\mu\rho}g_{\nu\lambda}g_{\alpha\beta} + g_{\mu\beta}g_{\nu\rho}g_{\alpha\lambda} + g_{\mu\lambda}g_{\nu\beta}g_{\alpha\rho} - g_{\nu\rho}g_{\mu\lambda}g_{\alpha\beta} - g_{\nu\beta}g_{\mu\rho}g_{\alpha\lambda} - g_{\nu\lambda}g_{\mu\beta}g_{\alpha\rho} \tag{E.36}$$

$$\varepsilon_{\alpha\beta\mu\nu}\varepsilon^{\alpha\beta}{}_{\rho\lambda} = 2(g_{\mu\lambda}g_{\nu\rho} - g_{\mu\rho}g_{\nu\lambda}) \tag{E.37}$$

Appendix F
The Fourier Transformation

When QCD sum rules of hadrons containing light quarks are considered, one usually carries out the OPE in coordinate space and Fourier transforms the result back into momentum space at the end of the calculation. We give in this appendix the necessary formulae for this task.

F.1 The Standard Case

For the standard Fourier transformation, one can derive (almost) all formulae needed in practical calculations from

$$\int d^4x e^{iqx} \frac{1}{(x^2)^n} = i(-1)^n \frac{2^{4-2n}\pi^2}{\Gamma(n-1)\Gamma(n)} (q^2)^{n-2} \ln(-q^2) + P_{n-2}(q^2), \quad \text{(F.1)}$$

which is valid for $n \geq 2$. The derivation of this equation can be found in Novikov et al. (1984). $P_m(q^2)$ stands for a polynomial of q^2 of order m. The coefficients of this polynomial are in fact divergent, but as they will in any case vanish when the Borel transform is applied, we omit them in the following discussion.

Variations of Eq. (F.1) with various tensor structures can be constructed by taking appropriate derivatives:

$$\int d^4x e^{iqx} \frac{x^\mu x^\nu \ldots}{(x^2)^n} = \left(\frac{\partial}{i\partial q_\mu}\right)\left(\frac{\partial}{i\partial q_\nu}\right) \ldots \int d^4x e^{iqx} \frac{1}{(x^2)^n}. \quad \text{(F.2)}$$

F.2 The "Old Fashioned" Case

Here, we evaluate the Fourier transforms of the various terms occurring in the "old fashioned" correlator of Eq. (3.53).

P. Gubler, *A Bayesian Analysis of QCD Sum Rules*, Springer Theses,
DOI: 10.1007/978-4-431-54318-3, © Springer Japan 2013

F.2.1 Dimension 0–5 Terms

For the terms appearing at dimensions 0–5, it is most convenient to work in coordinate space. Therefore, we directly use the expressions of Eq. (3.61) and substitute them into Eq. (3.53). For the dimension 5 term, this gives

$$\int d^4x\theta(x_0)e^{iqx}\frac{i}{(x^2-i\varepsilon)^2} = \int dx_0\theta(x_0)e^{iq_0x_0}\int d^3\mathbf{x}\frac{i}{(x_0^2-\mathbf{x}^2-i\varepsilon)^2}e^{-i\mathbf{q}\cdot\mathbf{x}}. \tag{F.3}$$

First, we calculate the integrals over the spacial angles θ and ϕ, leading to

$$\frac{2\pi}{|\mathbf{q}|}\int dx_0\theta(x_0)e^{iq_0x_0}\int_{-\infty}^{\infty}dr\frac{r}{(r-x_0+i\varepsilon)^2(r+x_0-i\varepsilon)^2}e^{i|\mathbf{q}|r}, \tag{F.4}$$

where we have used the definition $r \equiv |\mathbf{x}|$. Next comes the integral over r, which can be done in a standard way with the help of the Cauchy theorem. We thus obtain

$$\pi^2\int dx_0\theta(x_0)\frac{1}{x_0-i\varepsilon}e^{ix_0(q_0-|\mathbf{q}|)}. \tag{F.5}$$

At this point, we can drop $|\mathbf{q}|$, as there is no danger that the limit $|\mathbf{q}| \to 0$ leads to a divergence. Furthermore, we here introduce the Fourier transformed expression for the Heaviside step function:

$$\theta(x_0) = \frac{1}{2\pi i}\int dk_0\frac{1}{k_0-i\varepsilon}e^{ix_0k_0}. \tag{F.6}$$

We then get

$$\frac{\pi}{2i}\int dk_0\int dx_0\frac{1}{k_0-i\varepsilon}\frac{1}{x_0-i\varepsilon}e^{ix_0(q_0+k_0)}. \tag{F.7}$$

Making use of Eq. (F.6) now for the integral over x_0, giving the final result:

$$\pi^2\int_{-q_0}^{\infty}dk_0\frac{1}{k_0-i\varepsilon} = -\pi^2\ln(-q_0-i\varepsilon)+\pi^2\ln(\infty-i\varepsilon). \tag{F.8}$$

Here, we encounter a divergence in the second term, which, however, leads to no relevant contribution to the imaginary part of the correlator, which is the only quantity that is needed for the sum rules. We can therefore ignore it and hence have obtained the result used in Eq. (3.63).

The term of dimension 0, 3 and 4 can be calculated in a similar fashion. The main difference is that due to the larger powers in the denominator, the poles used in the Cauchy theorem leading to Eq. (F.5) are of a larger degree, which however does not

introduce any essential new difficulties. We here give only the results:

$$\begin{aligned}
\text{dim. } 0: &\int d^4x\theta(x_0)e^{iqx}\frac{\not{x}}{(x^2-i\varepsilon)^5} = -\frac{\pi^2}{2^9 3}\gamma_0 q_0^5 \ln(-q_0 - i\varepsilon),\\
\text{dim. } 3: &\int d^4x\theta(x_0)e^{iqx}\frac{i}{(x^2-i\varepsilon)^3} = \frac{\pi^2}{8}q_0^2 \ln(-q_0 - i\varepsilon),\\
\text{dim. } 4: &\int d^4x\theta(x_0)e^{iqx}\frac{\not{x}}{(x^2-i\varepsilon)^3} = -\frac{\pi^2}{4}\gamma_0 q_0 \ln(-q_0 - i\varepsilon).
\end{aligned} \tag{F.9}$$

We here, as above, have taken the limit $|\mathbf{q}| \to 0$. Note, however that in all the above calculations, this limit can be taken only *after* the integral over r has been carried out, as otherwise the factor $1/|\mathbf{q}|$ appearing in Eq. (F.4) can not be properly treated.

F.2.2 Dimension 6–10 Terms

For the terms with dimension larger than 5, the calculation is simpler if one starts from momentum space. This means that we take the expressions of Eq. (3.61) to substitute them into Eq. (3.53). In fact, the basic steps of the calculation are already given in Eqs. (3.54) and (3.55) of the main text and the result of the dimension 6 term can be directly deduced from these equations by setting $m^{\pm} = 0$. For the sake of illustration, we here show the calculation of one more term, the one of dimension 7.

For getting the result of this term we have to evaluate the following integral:

$$\int d^4x\theta(x_0)e^{iqx}\int\frac{d^4p}{(2\pi)^4}e^{-ipx}\frac{1}{p^2+i\varepsilon}. \tag{F.10}$$

Here, we first employ the expression of Eq. (F.6) and perform the integral over x. This yields

$$\frac{1}{2\pi i}\int dk_0\frac{1}{k_0 - i\varepsilon}\frac{1}{(k_0+q_0)^2 - \mathbf{q}^2 + i\varepsilon}. \tag{F.11}$$

For calculating the remaining integral over k_0, we note that there are three poles in the integrand, two in the upper half of the imaginary plane, and one in the lower half. Closing thus the contour in the lower half of the imaginary plane, we pick up the residue of this single pole, which originates from the second factor of the integrand. We then obtain the final result as

$$\frac{1}{2}\frac{1}{\sqrt{\mathbf{q}^2} - i\varepsilon}\frac{1}{q_0 - \sqrt{\mathbf{q}^2} + i\varepsilon}. \tag{F.12}$$

As is clear from this expression, we can at this point not take the limit $|\mathbf{q}| \to 0$ as it would lead to a divergence. This problem is only cured after the integral over q_0 is carried out as shown in Eq. (3.64).

The Fourier transforms of the higher order terms can be calculated analogously and we here show only the results.

$$\begin{aligned}
\text{dim. 8}: &\int d^4x\theta(x_0)e^{iqx}\int\frac{d^4p}{(2\pi)^4}e^{-ipx}\frac{\not{p}}{(p^2+i\varepsilon)^2}\\
&=\gamma_0\frac{1}{4}\frac{1}{\sqrt{\mathbf{q}^2}-i\varepsilon}\frac{1}{(q_0-\sqrt{\mathbf{q}^2}+i\varepsilon)^2},\\
\text{dim. 9}: &\int d^4x\theta(x_0)e^{iqx}\int\frac{d^4p}{(2\pi)^4}e^{-ipx}\frac{1}{(p^2+i\varepsilon)^2}\\
&=\frac{1}{4}\left[\frac{1}{(\sqrt{\mathbf{q}^2}-i\varepsilon)^2}\frac{1}{(q_0-\sqrt{\mathbf{q}^2}+i\varepsilon)^2}-\frac{1}{(\sqrt{\mathbf{q}^2}-i\varepsilon)^3}\frac{1}{q_0-\sqrt{\mathbf{q}^2}+i\varepsilon}\right].
\end{aligned} \tag{F.13}$$

Note that the result of the dimension 8 term in principle also contains expressions proportional to $\mathbf{q}\cdot\gamma$ as long as the $|\mathbf{q}| \to 0$ limit is not taken. These however vanish when the traces of Eq. (3.56) are taken and are therefore of no relevance here.

F.2.3 Evenness (Oddness) of Dimension 6, 8,…(7, 9,…) Terms

In this section, we proof the statement made in the main text, that the imaginary parts of the terms corresponding to dimensions $6, 8, \ldots$ $(7, 9, \ldots)$ in Eq. (3.63) are even (odd) functions of q_0, if one takes the limit $|\mathbf{q}| \to 0$.

First, by following the same steps that lead from Eq. (F.10) to Eq. (F.11), and setting $|\mathbf{q}| = 0$, we notice that all terms appearing at dimensions $6, 8, \ldots$ can generally be written down as

$$\begin{aligned}
&\frac{1}{2\pi i}\int dk_0\frac{1}{k_0-i\varepsilon}\frac{k_0+q_0}{[(k_0+q_0)^2+i\varepsilon]^n}\equiv F_1(q_0)\\
&=\frac{1}{2\pi i}\int dk_0\frac{1}{k_0-i\varepsilon}\frac{k_0+q_0}{(k_0+q_0+i\varepsilon)^n(k_0+q_0-i\varepsilon)^n}. \qquad (n=1,2,\ldots)
\end{aligned} \tag{F.14}$$

Here, we are ignoring any proportional real constant, including γ_0. Similarly, for dimensions $7, 9, \ldots$, we get

$$\begin{aligned}
&\tfrac{1}{2\pi i}\int dk_0\frac{1}{k_0-i\varepsilon}\frac{1}{[(k_0+q_0)^2+i\varepsilon]^n}\equiv F_2(q_0)\\
&=\tfrac{1}{2\pi i}\int dk_0\frac{1}{k_0-i\varepsilon}\frac{1}{(k_0+q_0+i\varepsilon)^n(k_0+q_0-i\varepsilon)^n}. \qquad (n=1,2,\ldots)
\end{aligned} \tag{F.15}$$

Next, we take the imaginary parts and, after some simple manipulations, get for $F_1(q_0)$

$$\begin{aligned}\mathrm{Im}F_1(q_0) &= \frac{1}{2i}\Big[F_1(q_0) - \overline{F_1(q_0)}\Big] \\ &= -\frac{1}{4\pi}\int dk_0\Big(\frac{1}{k_0 - i\varepsilon} + \frac{1}{k_0 - i\varepsilon}\Big)\Big[\frac{k_0 + q_0}{(k_0 + q_0 + i\varepsilon)^n(k_0 + q_0 - i\varepsilon)^n} \\ &\qquad - \frac{-k_0 + q_0}{(-k_0 + q_0 + i\varepsilon)^n(-k_0 + q_0 - i\varepsilon)^n}\Big],\end{aligned} \tag{F.16}$$

while the result for $F_2(q_0)$ is

$$\begin{aligned}\mathrm{Im}F_2(q_0) &= \frac{1}{2i}\Big[F_2(q_0) - \overline{F_2(q_0)}\Big] \\ &= -\frac{1}{4\pi}\int dk_0\Big(\frac{1}{k_0 - i\varepsilon}\Big)\Big[\frac{1}{(k_0 + q_0 + i\varepsilon)^n(k_0 + q_0 - i\varepsilon)^n} \\ &\qquad - \frac{1}{(-k_0 + q_0 + i\varepsilon)^n(-k_0 + q_0 - i\varepsilon)^n}\Big].\end{aligned} \tag{F.17}$$

Having the above equations at hand, it is now a trivial matter to show that

$$\mathrm{Im}F_1(-q_0) = \mathrm{Im}F_1(q_0), \tag{F.18}$$

and

$$\mathrm{Im}F_2(-q_0) = -\mathrm{Im}F_2(q_0), \tag{F.19}$$

which proofs our statement made in the main text, that the imaginary parts of the terms of dimensions 6, 8, ... (7, 9, ...) in the OPE of the "old fashioned" correlator are even (odd) functions of q_0 in the limit $|\mathbf{q}| \to 0$.

Appendix G
Derivation of the Shannon-Jaynes Entropy

In this appendix, we will provide two derivations for the Shannon-Jaynes entropy, given in Eq. (4.7), the first one making use of the law of large numbers, the second one being an axiomatic construction based on locality, system independence and scaling. We will mainly follow the explanations given in Asakawa et al. (2001).

G.1 Proof Based on the Law of Large Numbers

The proof of the Shannon-Jaynes entropy can be given by the so-called "monkey argument", which basically assumes that the probability of the spectral function $\rho(\omega)$ follows a certain Poisson distribution, as will be explained below.

What we need to derive is the probability of $\rho(\omega)$ to be in a specific region V of its allowed phase space. Formally, this probability can be denoted as

$$P(\rho \in V) = \frac{1}{Z(\alpha)} \int_V [d\rho] W(\alpha S(\rho)), \tag{G.1}$$

where $Z(\alpha)$ is simply a normalization constant, while α is just an arbitrary parameter, whose significance will be discussed in the main text. Furthermore, $S(\rho)$ is the entropy that we want to derive here. Also note that, as $P(\rho \in V)$ should have the maximum value where $S(\rho)$ is largest, the function W should be a monotone increasing function.

According to the monkey argument, we now divide the function $\rho(\omega)$ into N ω-regions of the same size and consider a monkey that throws M balls into them. The throwing process is not completely arbitrary, but is assumed to follow a certain pattern. Thus, each region has a probability p_i $(1 \leq i \leq N)$ to receive a ball, leading to an expectation value for the number of balls of $\lambda_i = M p_i$. Furthermore, we denote the actual number of balls that reaches a specific region as n_i. From probability theory, we know that if we take M to be very large and keep λ_i fixed, the probability of n_i to have a certain value will behave according to a Poisson distribution. Therefore, the

P. Gubler, *A Bayesian Analysis of QCD Sum Rules*, Springer Theses,
DOI: 10.1007/978-4-431-54318-3, © Springer Japan 2013

probability for a certain combination of $\mathbf{n} = (n_1, n_2, \ldots, n_N)$ to take place, can be written down as

$$P(\mathbf{n}) = \prod_{i=1}^{N} \frac{\lambda_i^{n_i} e^{-\lambda_i}}{n_i!}. \tag{G.2}$$

Here, the components of $\mathbf{n}$ are integers, and hence can not yet be considered to be a useful parametrization of the smooth function $\rho(\omega)$. We thus introduce a parameter q, with which we can make $\mathbf{n}$ proportional to the function $\rho(\omega)$:

$$\rho_i = q n_i, \tag{G.3}$$

where ρ_i stands for the value of $\rho(\omega)$ in the ith region of ω. Similarly, we can define the default model as

$$m_i = q \lambda_i. \tag{G.4}$$

We are now in a position to explicitly evaluate the probability of Eq. (G.1) as follows:

$$\begin{aligned} P(\rho \in V) &= \textstyle\sum_{\mathbf{n} \in V} P(\mathbf{n}) \simeq \frac{1}{q^N} \prod_{i=1}^{N} \int_V d\rho_i \, \frac{\lambda_i^{n_i} e^{-\lambda_i}}{n_i!} \\ &\simeq \frac{1}{(2\pi q)^{N/2}} \int_V \prod_{i=1}^{N} \frac{\rho_i}{\sqrt{\rho_i}} e^{S(\rho)/q}, \end{aligned} \tag{G.5}$$

where we have used the Stirling approximation $n! \simeq \sqrt{2\pi n} e^{n \log n - n}$ in the last line. $S(\rho)$ is given as

$$S(\rho) = \sum_{i=1}^{N} \Big[\rho_i - m_i - \rho_i \log \left(\rho_i / m_i \right) \Big], \tag{G.6}$$

which is equivalent to Eq. (4.7) of the main text. Furthermore, it is seen from Eqs. (G.1) and (G.5) that $q = 1/\alpha$ and that the measure $[d\rho]$ and the normalization constant $Z(\alpha)$ can be expressed as

$$[d\rho] = \prod_{i=1}^{N} \frac{\rho_i}{\sqrt{\rho_i}}, \qquad Z(\alpha) = \left(\frac{2\pi}{\alpha} \right)^{N/2}. \tag{G.7}$$

As a last point, we also observe that the function W of Eq. (G.1) is a simple exponential and therefore indeed a monotone increasing function as it should be.

G.2 Proof Based on an Axiomatic Construction

For illustration, we here give another proof for the Shannon-Jaynes entropy, which is based on the four axioms of locality, coordinate invariance, system independence

and scaling. This proof is less intuitive than the one given in the last section, it is, however, in some sense, more general, as it does not rely on the assumption of the Poisson distribution used in Eq. (G.2).

As in the last section, our task is to define a real functional, which satisfies the following condition:

$$\text{If } \rho_1 \text{ is a more probable function than } \rho_2, \text{ then: } S(\rho_1) > S(\rho_2). \tag{G.8}$$

Thus, the most probable of all functions can be found by looking for a stationary point in $S(f)$, which follows from the equation

$$\delta_\rho S(\rho) = 0. \tag{G.9}$$

Let us now derive the actual form of $S(\rho)$ by considering the four axioms mentioned above.

G.2.0.1 Locality

This axiom declares that the values of $\rho(\omega)$ at various values of ω should independently contribute to $S(\rho)$ without any correlation. Therefore, one can conclude that $S(\rho)$ should be a local function of $\rho(\omega)$ and can written down as

$$S(\rho) = \int d\omega m(\omega)\phi(\rho(\omega), \omega), \tag{G.10}$$

where $m(\omega)$ can be considered to be the integration measure and must be positive definite. Furthermore, ϕ is an arbitrary function of $\rho(\omega)$ and ω, but cannot contain any derivatives of $\rho(\omega)$, as they would lead to correlations between different values of ω.

G.2.0.2 Coordinate Invariance

The axiom of coordinate invariance demands that $S(\rho)$ does not depend on what sort of coordinates one uses for the function $\rho(\omega)$. In other words, $S(\rho)$ should be invariant under the coordinate transformation $\omega' = \omega'(\omega)$. Now, using $\rho(\omega)d\omega = \rho'(\omega')d\omega'$ and $m(\omega)d\omega = m'(\omega')d\omega'$, one can understand that the right hand side of Eq. (G.10) can only be invariant if the function $\rho(\omega)$ appears in ϕ divided by $m(\omega)$, because of the relation $\rho(\omega)/m(\omega) = \rho'(\omega')/m'(\omega')$. Hence, we can express Eq. (G.10) as

$$S(\rho) = \int d\omega m(\omega)\phi(\rho(\omega)/m(\omega)). \tag{G.11}$$

G.2.0.3 System Independence

This axiom states that in case of a function $\rho_c(\omega_1, \omega_2)$ having two independent variables ω_1 and ω_2, this function can be written as a product of two functions: $\rho_c(\omega_1, \omega_2) = \rho_1(\omega_1)\rho_2(\omega_2)$. Moreover, the corresponding integration measure can be divided in the same way: $m_c(\omega_1, \omega_2) = m_1(\omega_1)m_2(\omega_2)$. A further consequence of the axiom is that the variance of $S(\rho_c)$ w.r.t. $\rho_c(\omega_1, \omega_2)$ is given as

$$\frac{\delta S(\rho_c)}{\delta \rho_c(\omega_1, \omega_2)} = \alpha(\omega_1) + \beta(\omega_2), \tag{G.12}$$

where $\alpha(\omega_1)$ and $\beta(\omega_2)$ are functions related to the variance of $S(\rho_c)$ w.r.t. $\rho_1(\omega_1)$ and $\rho_2(\omega_2)$, respectively.

Now, using the form for $S(\rho_c)$ that we have obtained in Eq. (G.11), we can write down $S(\rho_c)$ as

$$S(\rho_c) = \int d\omega_1 \int d\omega_2 m_c(\omega_1, \omega_2)\phi(\rho_c(\omega_1, \omega_2)/m_c(\omega_1, \omega_2)), \tag{G.13}$$

which gives

$$\frac{\delta S(\rho_c)}{\delta \rho_c(\omega_1, \omega_2)} = \frac{d\phi}{dZ}\Big|_{Z=\rho_c(\omega_1,\omega_2)/m_c(\omega_1,\omega_2)} \equiv \sigma(Z = \rho_c(\omega_1, \omega_2)/m_c(\omega_1, \omega_2)). \tag{G.14}$$

Next, we act with $\partial^2/\partial\omega_1\partial\omega_2$ on the right hand sides of both Eqs. (G.12) and (G.14). As these should be equal, we are lead to the following equation for $\sigma(Z)$:

$$Z\frac{d^2\sigma(Z)}{dZ^2} + \frac{d\sigma(Z)}{dZ} = 0. \tag{G.15}$$

The above equation can be easily solved, giving $\sigma(Z) = c_1 \log(Z) + c_2$, from which we finally get the functional form of $\phi(Z)$ as

$$\phi(Z) = c_1 Z \log(Z) + (c_2 - c_1)Z + c_3, \tag{G.16}$$

where c_1, c_2 and c_3 are integration constants that are not yet determined at the current stage. Substituting the result of Eq. (G.16) into Eq. (G.11), $S(f)$ can now be given as

$$S(\rho) = \int d\omega \Big[c_1 \rho(\omega) \log\left(\frac{\rho(\omega)}{m(\omega)}\right) + (c_2 - c_1)\rho(\omega) + c_3 m(\omega) \Big]. \tag{G.17}$$

Using this equation, we get $\delta^2/\delta\rho^2 S(\rho) = c_1/\rho$ and thus observe that the sign of c_1 completely determines the curvature of $S(\rho)$, as ρ is a positive definite function. Therefore, in order for $S(\rho)$ to be bounded from above, one has to chose c_1 to be negative.

G.2.0.4 Scaling

According to this axiom, in case of no additional information available on $\rho(\omega)$ (for instance, from the likelihood function of Eq. (4.4)), the most probable form of $\rho(\omega)$ should be equal to the integration measure $m(\omega)$. Thus the maximum of $S(\rho)$ should be at $\rho(\omega) = m(\omega)$.

The maximum of $S(\rho)$ of Eq. (G.17) can be obtained from the solution of $\frac{\delta S(\rho)}{\delta \rho} = 0$, which gives

$$\rho(\omega) = m(\omega) e^{-c_2/c_1}. \tag{G.18}$$

From this result, we can immediately conclude that for satisfying the scaling axiom, we need to set $c_2 = 0$. Thus, the form of $S(\rho)$ is now

$$S(\rho) = -c_1 \int d\omega \left[\rho(\omega) - \rho(\omega) \log\left(\frac{\rho(\omega)}{m(\omega)}\right) - \frac{c_3}{c_1} m(\omega) \right]. \tag{G.19}$$

As a last task, we still have to determine c_1 and c_3. Considering first c_3, we see from the above equation that the term proportional to this constant does not depend on $\rho(\omega)$ and is therefore not of much relevance in the present discussion. In order for $S(\rho)$ to vanish when $\rho(\omega)$ equals $m(\omega)$, one usually chooses $c_3 = c_1$ for convenience. As for c_1, we have already mentioned above that it should have a negative value. As can be observed from Eq. (G.19), its magnitude just becomes an overall normalization factor in front of the integral over ω, which can be arbitrarily chosen. Usually, one takes $c_1 = -1$ for simplicity. We thus are lead to

$$S(\rho) = \int d\omega \left[\rho(\omega) - m(\omega) - \rho(\omega) \log\left(\frac{\rho(\omega)}{m(\omega)}\right) \right], \tag{G.20}$$

which is indeed the Shannon-Jaynes entropy of Eq. (4.7).

Appendix H
Uniqueness of the Maximum of $P[\rho|GH]$

It is important for the MEM procedure that there is only one solution for $\rho(\omega)$, which maximizes the conditional probability $P[\rho|GH]$. We will proof in this short appendix, that the solution is indeed unique if it exists, following the discussion given in Asakawa et al. (2001).

For proofing the uniqueness of the solution for $\rho(\omega)$, we first have to show the correctness of the following mathematical statement:

> Given a real and smooth function $F(x_1, x_2, \ldots, x_n)$ with real variables $(x_1, x_2, \ldots, x_n) \in \mathbf{R^n}$, for which the matrix $\partial^2 F/\partial x_i \partial x_j$ is negative definite, the solution of the equations $\partial F/\partial x_i = 0$ is unique if it exists.

Note here that the negative definiteness of $\partial F/\partial x_i \partial x_j$ can be denoted as

$$\sum_{i,j=1}^{n} y_i \frac{\partial^2 F}{\partial x_i \partial x_j} y_j < 0 \qquad ({}^{\forall} y_i \in \mathbf{R}/\{\mathbf{0}\}). \tag{H.1}$$

For showing the above statement, we assume that there are two solutions for $\partial F/\partial x_i = 0$, $\mathbf{x_1}$ and $\mathbf{x_2}$, and define $\mathbf{x}(t) \equiv \mathbf{x_1} + t(\mathbf{x_2} - \mathbf{x_1})$ and $G(t) \equiv F(\mathbf{x}(t))$. From these definitions, we can immediately see that $dG(t)/dt$ satisfies

$$\left.\frac{dG(t)}{dt}\right|_{t=0} = \left.\frac{dG(t)}{dt}\right|_{t=1} = 0. \tag{H.2}$$

Now, from the smoothness of F, the function $G(t)$ must be continuous and differentiable. We can therefore use Rolle's theorem, which states that between $t = 0$ and $t = 1$, there must be at least one t which satisfies

$$\frac{d^2 G(t)}{dt^2} = \sum_{i,j=1}^{n} y_i \left.\frac{\partial^2 F}{\partial x_i \partial x_j}\right|_{\mathbf{x}=\mathbf{x}(t)} y_j = 0, \tag{H.3}$$

P. Gubler, *A Bayesian Analysis of QCD Sum Rules*, Springer Theses,
DOI: 10.1007/978-4-431-54318-3, © Springer Japan 2013

which leads to a contradiction with Eq. (H.1). Therefore, the solution of $\partial F/\partial x_i = 0$ must be unique if it exists.

Thus, all we have to do for proofing the uniqueness of the solution for $\rho(\omega)$, is to show that $Q(\rho)$ of Eq. (4.9) satisfies the condition analog to Eq. (H.1). Using the (discretized forms of) the likelihood function and the prior probability of Eqs. (4.4) and (4.7), we can derive

$$\sum_{i,j=1}^{N_\omega} y_i \frac{\partial^2 Q}{\partial\rho_i \partial\rho_j} y_j = -\frac{\alpha}{\Delta\omega} \sum_{i=1}^{N_\omega} \frac{y_i^2}{\rho_i} - \frac{\Delta x}{x_{\max} - x_{\min}} \sum_{j=1}^{N_x} \sum_{i=1}^{N_\omega} \frac{\left[K(x_j, \rho_i) y_i\right]^2}{\sigma^2(x_j)}. \quad \text{(H.4)}$$

Here, ρ_i represents the discretized data points of $\rho(\omega)$: $\rho_i \equiv \rho(\omega_i)\Delta\omega$, with $\Delta\omega \equiv \frac{\omega_{\max}-\omega_{\min}}{N_\omega}$ and $\omega_i \equiv \frac{i}{N_\omega}(\omega_{\max} - \omega_{\min}) + \omega_{\min}$. Similarly, x_j stands for $x_j \equiv \frac{j}{N_x}(x_{\max} - x_{\min}) + x_{\min}$ and Δx for $\Delta x \equiv \frac{x_{\max}-x_{\min}}{N_x}$. While the second term in principle can become 0 for certain values of $\mathbf{y}$, the first term is always negative because of $0 < \alpha$ and $0 \leq \rho_i$. Therefore, we can conclude that

$$\sum_{i,j=1}^{N_\omega} y_i \frac{\partial^2 Q}{\partial\rho_i \partial\rho_j} y_j < 0 \qquad (^\forall y_i \in \mathbf{R}/\{\mathbf{0}\}), \quad \text{(H.5)}$$

which, together with the statement shown above, proofs the uniqueness of the maximum of $P[\rho|GH]$ if it exists.

References

Arfken G (1970) Mathematical methods for physicists, 2nd edn. Academic Press, New York
Asakawa M, Hatsuda T, Nakahara Y (2001) Prog Part Nucl Phys 46:459
Dubovikov MS, Smilga AV (1981) Nucl Phys B 185:109
Fock VA (1937) Sowj Phys 12:404
Ioffe BL, Fadin VS, Lipatov LN (2010) Quantum chromodynamics, perturbative and nonperturbative aspects. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, Cambridge
Morse PM, Feshbach H (1953) Methods of theoretical physics. McGraw-Hill Book Company, New York
Novikov VA, Shifman MA, Vainshtein AI, Zakharov VI (1984) Fortsch Phys 32:585
Peskin ME, Schroeder DV (1995) An introduction to quantum field theory. Westview Press, Boulder
Schwinger J (1954) Phys Rev 82:664
Shifman MA (1980) Nucl Phys B 173:13