JOHANNES GUTENBERG
UNIVERSITÄT MAINZ

Electromagnetic form factors and radii of the nucleon from Lattice QCD and the proton radius puzzles

Dissertation
zur Erlangung des Grades
„Doktor der Naturwissenschaften“

am Fachbereich Physik, Mathematik und Informatik
der Johannes Gutenberg-Universität
in Mainz

Miguel Salg
geboren in Aschaffenburg

Mainz, den 15. Oktober 2024

Miguel Salg: Electromagnetic form factors and radii of the nucleon from Lattice QCD and the proton radius puzzles

Berichterstatter: Prof. Dr. Hartmut Wittig, Prof. Dr. Pedro Schwaller, Prof. Dr. Sara Collins

Tag der Promotionsprüfung: 14.10.2024

Abstract

In this thesis, we investigate the electromagnetic form factors of the proton and neutron in the framework of Quantum Chromodynamics (QCD). To perform a calculation of such low-energy quantities based on first principles, we employ the lattice regularization of QCD. In Lattice-QCD simulations of nucleon-structure observables, systematic errors are inherent due to the finite lattice spacing and volume as well as due to contamination by excited states. These can be controlled and removed in a systematic fashion: for the removal of excited-state contributions, a variety of dedicated methods exists, while discretization and finite-volume effects also need to be taken into account by performing a continuum and infinite-volume extrapolation. However, all previous lattice studies of the electromagnetic form factors of the nucleon have either neglected the numerically challenging quark-disconnected contributions or were not extrapolated to the continuum and infinite-volume limit.

We present results for the electromagnetic form factors of the proton and neutron computed on the $(2+1)$-flavor Coordinated Lattice Simulations (CLS) ensembles including both quark-connected and -disconnected contributions while, at the same time, controlling all sources of systematic uncertainties. For the excited-state analysis, we explore three complementary methods based on two-state fits to the effective form factors and on two different truncations of the summation method, respectively. The Q^2-, pion-mass, lattice-spacing and finite-volume dependence of our form factor data is fitted simultaneously to the expressions resulting from covariant baryon chiral perturbation theory including vector mesons amended by models for lattice artefacts. From these fits, we determine the electric, magnetic, Zemach and Friar radii as well as the magnetic moments of the proton and neutron. To assess the influence of systematic effects, we average over various cuts in the pion mass and the momentum transfer, as well as over different models for the lattice-spacing and finite-volume dependence, using weights derived from the Akaike Information Criterion (AIC).

Our *ab-initio* QCD results for the electromagnetic radii of the proton are of particular relevance in light of the so-called proton radius puzzle, *i.e.*, the observation of a large tension between different experimental measurements of the proton's electric radius, which is after more than a decade of vigorous research still not completely explained. Also for the magnetic radius, analyses based on different data sets find discrepant results. In this situation, a firm theoretical prediction of the proton radii can contribute towards the clarification of the origins of the discrepancies.

Our results for the magnetic moments of the proton and neutron are in good agreement with the experimentally very precisely known values. For the radii of the proton, we achieve, including all systematic errors, a precision which enables a meaningful comparison to the various experiments and data-driven evaluations. On the one hand, our result for the electric radius of the proton clearly points towards a small value, as favored by muonic hydrogen spectroscopy, the recent *ep*-scattering experiment by PRad and data-driven dispersive analyses. Our estimate for the magnetic radius, on the other hand, is well compatible with that inferred from the A1 *ep*-scattering experiment by a z-expansion analysis and in tension with z-expansion on the remaining world data as well as with dispersive approaches.

Zusammenfassung

In dieser Doktorarbeit werden die elektromagnetischen Formfaktoren des Protons und Neutrons im Rahmen der Quantenchromodynamik (QCD) untersucht. Um solche Niederenergiegrößen basierend auf ersten Prinzipien zu berechnen, wird die Gitterregularisierung der QCD verwendet. In Gitter-QCD-Simulationen von Nukleonstruktur-Observablen sind systematische Fehler inhärent aufgrund des endlichen Gitterabstands und -volumens sowie aufgrund der Kontamination durch angeregte Zustände. Diese können kontrolliert und auf systematische Weise entfernt werden: Für die Entfernung der Beiträge angeregter Zustände existieren eine Reihe dedizierter Methoden, während Diskretisierungs- und Endlich-Volumen-Effekte ebenfalls berücksichtigt werden müssen, indem man eine Extrapolation ins Kontinuum und zu unendlichem Volumen durchführt. Jedoch haben alle bisherigen Gitterstudien der elektromagnetischen Formfaktoren des Nukleons entweder die numerisch herausfordernden quark-unverbundenen Beiträge vernachlässigt oder waren nicht zum Kontinuums- und Unendlich-Volumen-Limes extrapoliert.

Wir präsentieren Resultate für die elektromagnetischen Formfaktoren des Protons und Neutrons, die auf den $(2 + 1)$-flavour Coordinated Lattice Simulations (CLS) Ensembles berechnet wurden. Sie berücksichtigen sowohl quark-verbundene als auch -unverbundene Beiträge und kontrollieren gleichzeitig alle Quellen systematischer Unsicherheiten. Für die Analyse angeregter Zustände werden drei komplementäre Methoden erforscht, die auf Zweizustandsfits an die effektiven Formfaktoren bzw. auf zwei verschiedenen Trunkierungen der Summationsmethode basieren. Die Q^2-, Pionmassen-, Gitterabstands- und Volumen-Abhängigkeit unserer Formfaktor-Daten wird simultan an die Ausdrücke gefittet, die aus kovarianter baryonischer chiraler Störungstheorie inklusive Vektormesonen resultieren, und die um Modelle für Gitterartefakte ergänzt werden. Mithilfe dieser Fits werden die elektrischen, magnetischen, Zemach- und Friar-Radien sowie die magnetischen Momente des Protons und Neutrons bestimmt. Um die Einflüsse systematischer Effekte zu beurteilen, wird über diverse Cuts in der Pionmasse und im Impulsübertrag sowie über verschiedene Modelle für die Gitterabstands- und Volumen-Abhängigkeit gemittelt. Hierzu werden Gewichte verwendet, die aus dem Akaike-Informationskriterium (AIC) hergeleitet sind.

Unsere *ab-initio*-Resultate für die elektromagnetischen Radien des Protons sind von besonderer Bedeutung im Lichte des sogenannten Proton-Radius-Puzzles, d. h. der großen Spannung zwischen verschiedenen experimentellen Messungen des elektrischen Radius des Protons, die nach mehr als einer Dekade intensiver Forschung immer noch nicht vollständig erklärt ist. Auch für den magnetischen Radius finden Analysen basierend auf verschiedenen Datensätzen diskrepante Ergebnisse. In dieser Situation kann eine solide theoretische Vorhersage der Protonradien dazu beitragen, die Ursprünge der Diskrepanzen zu klären.

Unsere Ergebnisse für die magnetischen Momente des Protons und Neutrons befinden sich in guter Übereinstimmung mit den experimentell sehr genau bekannten Werten. Für die Radien des Protons erreichen wir inklusive aller systematischer Fehler eine Präzision, die einen aussagekräftigen Vergleich mit den diversen Experimenten und datengetriebenen Bestimmungen erlaubt. Einerseits deutet unser Resultat für den elektrischen Radius des Protons klar auf einen kleinen Wert hin, wie er von der Spektroskopie myonischen Wasserstoffs, dem kürzlich erfolgten *ep*-Streuexperiment von PRad und datengetriebenen dispersiven Analysen bevorzugt wird. Andererseits ist unser Ergebnis für den magnetischen Radius gut kompatibel mit dem mithilfe einer z-expansion-Analyse aus dem A1 *ep*-Streuexperiment abgeleiteten Wert und in Spannung mit z-expansion auf den verbleibenden Welt-Daten sowie mit dispersiven Ansätzen.

Publications

Intermediate results of this thesis have appeared in the following conference proceedings,

- Dalibor Djukanovic *et al.* "Isoscalar electromagnetic form factors of the nucleon in $N_\mathrm{f} = 2+1$ lattice QCD." in: *Proceedings of The 38th International Symposium on Lattice Field Theory — PoS(LATTICE2021)*. Vol. 396. July 2022, p. 422. DOI: 10.22323/1.396.0422. arXiv: 2110.10626 [hep-lat],
- Miguel Salg *et al.* "Electromagnetic form factors of the proton and neutron from $N_f = 2+1$ lattice QCD." in: *Proceedings of The 39th International Symposium on Lattice Field Theory — PoS(LATTICE2022)*. Vol. 430. Apr. 2023, p. 121. DOI: 10.22323/1.430.0121. arXiv: 2211.17049 [hep-lat],
- Miguel Salg *et al.* "Proton and neutron electromagnetic radii and magnetic moments from $N_f = 2+1$ lattice QCD." in: *Proceedings of The 40th International Symposium on Lattice Field Theory — PoS(LATTICE2023)*. Vol. 453. Jan. 2024, p. 317. DOI: 10.22323/1.453.0317. arXiv: 2401.05404 [hep-lat],
- Miguel Salg *et al.* "Proton and neutron electromagnetic radii and magnetic moments from lattice QCD." in: *EPJ Web of Conferences* 303 (Sept. 2024), p. 02002. DOI: 10.1051/epjconf/202430302002. arXiv: 2401.09440 [hep-lat],

and central pieces of the work have been published in the following peer-reviewed articles,

- Dalibor Djukanovic *et al.* "Electromagnetic form factors of the nucleon from $N_f = 2+1$ lattice QCD." in: *Physical Review D* 109.9 (May 2024), p. 094510. DOI: 10.1103/PhysRevD.109.094510. arXiv: 2309.06590 [hep-lat],
- Dalibor Djukanovic *et al.* "Precision calculation of the electromagnetic radii of the proton and neutron from lattice QCD." in: *Physical Review Letters* 132.21 (May 2024), p. 211901. DOI: 10.1103/PhysRevLett.132.211901. arXiv: 2309.07491 [hep-lat],
- Dalibor Djukanovic *et al.* "Zemach and Friar radii of the proton and neutron from lattice QCD." in: *Physical Review D* 110.1 (July 2024), p. L011503. DOI: 10.1103/PhysRevD.110.L011503. arXiv: 2309.17232 [hep-lat].

Contents

1 Introduction

The elementary particles which make up the visible matter in our Universe are described theoretically by the Standard Model of Particle Physics: the fermionic quarks and leptons, and the bosonic exchange particles. They can interact through four fundamental forces: the strong, weak, electromagnetic and gravitational interactions. The mathematical formulation of the Standard Model is realized in terms of Quantum Field Theories. Three of the four aforementioned forces can be represented in this way; only gravity is not yet included.

The theory of the strong interaction in the context of the Standard Model is Quantum Chromodynamics (QCD). It contains fermions – the quarks – in six flavors (up, down, strange, charm, bottom and top), which additionally come in three different colors. The color charges of the quarks are assigned to the fundamental representation of the gauge group of QCD, which is SU(3). The exchange particles mediating the strong force are called gluons. They carry a color charge in the adjoint representation of SU(3) and thus appear in $3^2 - 1 = 8$ species.

At large distances or low energies, the coupling of QCD is large, which is connected to the phenomenon of confinement: the only finite-energy asymptotic states of the theory are those that are singlets under SU(3), or, in other words, color neutral. This means that color-charged particles, like the quarks and gluons, cannot be isolated and have to form composite particles, the hadrons. Due to the large coupling of QCD in the hadronic regime, a theoretical description of their structure necessitates a non-perturbative treatment by methods such as Lattice QCD. Here, one replaces space-time by a four-dimensional Euclidean lattice. It serves as a gauge-invariant UV-regulator for the Quantum Field Theory and makes the path integral finite dimensional and thus amenable to a numerical evaluation. Even though it introduces systematic errors due to the finite lattice spacing and volume, the method allows a systematic extrapolation to the continuum and infinite-volume limit.

In spite of its enormous success, the Standard Model cannot be complete for a number of different reasons. Firstly, the original Standard Model describes the neutrinos as massless. The observation of neutrino oscillations between different flavors (electron, muon and tau), however, requires them to have a finite mass. It is still unclear of which nature the neutrino mass term exactly is (Dirac or Majorana) and precisely how the correct extension of the Standard Model should hence look like. Another important shortcoming of the Standard Model is its failure to provide a sufficient source of CP and baryon-number violation to explain the large asymmetry between particles and antiparticles observed in the Universe. Futhermore, as mentioned in the beginning, General Relativity, which is the currently accepted theory of gravity, cannot easily be unified with the quantum-field-theoretical framework of the Standard Model. Apart from this, astrophysical observations of gravitational effects which cannot be explained by the visible matter in the Universe suggest that a large fraction of the mass in the Universe consists of dark matter which is not described by the Standard Model. Lastly, the accelerating expansion of the Universe requires the existence of dark energy which is even more mysterious as it needs to exert a negative pressure.

In general, two complementary approaches are being followed in the search for Beyond-the-Standard-Model (BSM) physics: direct searches for new particles at high energies and indirect searches by precision measurements of low-energy observables. The latter are based on the fact that in the presence of BSM physics, quantum loop corrections would also cause a slight difference to Standard-Model predictions at low energies. This approach therefore requires

theoretical calculations based on the Standard Model matching the precision of the experiments. A particularly exciting case is the anomalous magnetic moment of the muon, where currently a huge discrepancy between high-precision experimental measurements [8] and the theoretical prediction from the Standard Model [9] exists. The evaluation of the QCD contribution to the latter is, however, based on data-driven dispersive approaches. First-principles calculations using Lattice QCD, on the contrary, suggest that the discrepancy might be a lot smaller [10]. Since Ref. [10] is so far the only full lattice calculation of the muon anomalous magnetic moment with a precision that enables a comparison to experiment, this statement needs independent confirmation, though. Moreover, the precision of the lattice calculations has to be further increased in order to fully match the experimental one. Tremendous research efforts are ongoing in this direction [11].

Another long-standing discrepancy in subatomic physics is the so-called proton radius puzzle [12]. In the 1950s, Hofstadter et al. first showed experimentally that the proton is not a point-like particle, but has a finite extent [13]. Since then, measurements of the proton's size as parametrized by its radius have progressed immensely and are now a subject of precision science. These advancements have caused a number of so far unexplained tensions to appear. The most significant one concerns the electric (charge) radius of the proton: the value reported by the A1 collaboration based on *ep*-scattering data [$\sqrt{\langle r_E^2 \rangle^p} = (0.879 \pm 0.005\ (\text{stat}) \pm 0.006\ (\text{syst}))$ fm [14]], while in good agreement with hydrogen spectroscopy [$\sqrt{\langle r_E^2 \rangle^p} = 0.8758(77)$ fm [15]] at the time of publication, is incompatible with the most precise determination, which comes from the spectroscopy of muonic hydrogen [$\sqrt{\langle r_E^2 \rangle^p} = 0.84087(39)$ fm [16, 17]]. This has led to speculations about possible differences between measurements using either electrons or muons which could be caused by BSM physics [18].

The most recent experiments using electronic hydrogen spectroscopy favor the lower value [19–22], with the exception of Ref. [23] which reports a larger value in agreement with older measurements [15] and Ref. [24] which reports an intermediate value. The latest determinations from *ep* scattering yield differing results as well: while the A1 collaboration has essentially confirmed their previous result using the initial-state radiation technique [$\sqrt{\langle r_E^2 \rangle^p} = (0.878 \pm 0.011\ (\text{stat}) \pm 0.031\ (\text{syst}))$ fm [25]], the PRad experiment at Jefferson Lab has reported a smaller value [$\sqrt{\langle r_E^2 \rangle^p} = (0.831 \pm 0.007\ (\text{stat}) \pm 0.012\ (\text{syst}))$ fm [26]]. It is worth pointing out that dispersive analyses had already favored a smaller proton radius for a long time [27, 28], and continue to do so [29–32]. This applies in particular also to the dispersive analysis of the data taken by the A1 experiment [33, 34].

In an effort to resolve the still existing tensions, several new experimental efforts are underway: A new *ep*-scattering experiment, MAGIX [35], is being prepared at the Mainz-based accelerator MESA which is currently under construction. An upgrade of PRad, dubbed PRad-II, has been approved [36], while the ULQ^2 experiment at ELPH in Tohoku, Japan, is already taking data [37]. To complement the results from electronic and muonic hydrogen spectroscopy and *ep* scattering with a result from μp scattering, the MUSE collaboration aims to measure the μp cross section to sub-percent precision at PSI [38]. Furthermore, the AMBER experiment at CERN plans to determine the electric proton radius to a precision on the order of 0.01 fm using a similar method [39].

Recently, the goal of achieving a more complete description of the fundamental electromagnetic properties of the proton has come into focus. For instance, for the magnetic radius an analysis based on the z-expansion obtains two different numbers, depending on whether just the A1 data is analyzed [$\sqrt{\langle r_M^2 \rangle^p} = (0.776 \pm 0.034\ (\text{stat}) \pm 0.017\ (\text{syst}))$ fm [14, 40]] or the rest of

the *ep*-scattering world data excluding A1 [$\sqrt{\langle r_M^2 \rangle^p} = 0.914(35)$ fm [40]]. The tension is not as large as for the electric radius, but still, these two numbers are not compatible with each other. Dispersive analyses [30, 32, 34] arrive at magnetic proton radii significantly larger than the A1-data analyses [14, 40], but smaller than z-expansion on the remaining world data [40]. Therefore, the magnetic proton radius has received more attention in recent years [34], as have other derived definitions of radii relevant for atomic spectroscopy experiments, like the Zemach and Friar radii [41]. Of the newly planned experiments devoted to the electric proton radius, only MAGIX will immediately address the magnetic radius as well [42], while there is a prospect of AMBER measuring it in phase 2 of the experiment using inverse kinematics [43].

In order to understand whether the various discrepancies can be traced to the experimental data and their analyses, or whether they are really an indication for BSM physics [18], a firm Standard-Model prediction of the proton's radii is required. For this purpose, it is, as argued above, mandatory to apply a non-perturbative methodology such as Lattice QCD. In Lattice QCD as in the context of scattering experiments, radii are extracted from the derivative of the electromagnetic form factors $G_{E,M}(Q^2)$ at $Q^2 = 0$. A full calculation enabling a distinction between proton and neutron necessitates the evaluation of quark-disconnected diagrams which are computationally very expensive and notorious for their unfavorable signal-to-noise ratio. Previous lattice calculations of electromagnetic form factors and radii published in Refs. [44–61], with the exception of Refs. [54, 58], have neglected quark-disconnected contributions due to this technical complication. Furthermore, as pointed out above, results computed from Lattice QCD generally need to be extrapolated to the continuum and infinite-volume limit in order to allow a meaningful comparison with experiments, which has not been done in Refs. [54, 58].

In this thesis, we present results for the electromagnetic form factors of the proton and neutron computed from a set of $(2+1)$-flavor Lattice-QCD ensembles at four different lattice spacings and pion masses between 130 and 290 MeV. Our study improves on all previous calculations by explicitly evaluating both quark-connected and -disconnected contributions and, at the same time, taking into account all relevant systematic effects due to excited-state contamination, finite-volume effects and the extrapolation to the physical point. In addition to determining the shape of the form factors at moderate momentum transfers ($Q^2 \lesssim 0.6$ GeV^2), we extract the electric, magnetic, Zemach and Friar radii as well as the magnetic moments of the proton and neutron.

We start in chapter 2 with a brief recapitulation of the essentials of continuum QCD and its Euclidean formulation. Chapter 3 deals with the lattice discretization of QCD, methods for the numerical evaluation of the path integral and the construction of the discretized operators relevant to the observables of interest. Next, we consider in chapter 4 the specific set of simulations undertaken by the Coordinated Lattice Simulations (CLS) initiative, which we employ in this thesis. We give details about the action and boundary conditions used by CLS, the algorithmic setup of the simulations, the scale setting procedure and the location of the ensembles in parameter space. Chapter 5 is concerned with the calculation of the pertinent observables on these ensembles. After formally defining the form factors and radii, we discuss techniques for the efficient calculation of propagators, one of the most expensive parts of the computation. The remainder of the chapter is dedicated to the nucleon two- and three-point functions, their physical significance and their numerical implementation, finishing off with a brief note on the pion two-point function. In chapter 6, we report on our analysis and final results. To that end, we start with an overview of common data analysis techniques for Lattice QCD pertaining to error quantification and regression. Our own analysis begins with the determination of the pion and nucleon masses on the employed ensembles. Afterwards, we extract the effective form factors from ratios of nucleon three- and two-point functions. A particular emphasis of this thesis is on the removal of excited-state contributions, for which we

explore and compare different methods. Since only discrete Q^2-values are accessible in lattice simulations, we need to parametrize the momentum dependence of the form factors in order to extract the radii from them. This is combined with the extrapolation to the physical point by performing simultaneous fits to the Q^2-, pion-mass, lattice-spacing and finite-volume dependence of the form factors to the expressions resulting from covariant baryon chiral perturbation theory (BχPT), including vector mesons and amending the expressions by models for lattice artefacts. We compare this analysis to a more traditional approach based on a parametrization of the Q^2-dependence of the form factors on each ensemble followed by an extrapolation of the radii to the physical point. Moreover, we have developed a method to compute the Zemach and Friar radii based on the BχPT fits to the form factors and an extrapolation to arbitrarily large Q^2-values, which is presented next. Systematic errors are quantified using a model average, from which we also obtain our final results for the form factors, radii and magnetic moments. Chapter 7 finally draws some conclusions and gives an outlook to possible further improvements of our calculation as well as to potential future research opportunities.

2 QCD in the continuum

Within the framework of the Standard Model of Particle Physics, the strong nuclear force is described theoretically by a Quantum Field Theory called Quantum Chromodynamics (QCD) that we have already mentioned in the introduction. In this chapter, QCD will be formally introduced and discussed in the continuum. Moreover, its Euclidean formulation will be explained, which will be required for the lattice discretization in chapter 3. This chapter is based on the standard textbooks [62–66], and is partly inspired by the PhD theses [67, 68].

2.1 The QCD action

QCD is a non-Abelian gauge theory with gauge group SU(3), coupled to fermions (quarks) in the fundamental representation [64]. The Lagrangian therefore consists of a fermionic part $\mathcal{L}_F$ and a gauge part $\mathcal{L}_G$,

$$\mathcal{L}_{\mathrm{QCD}} = \mathcal{L}_F + \mathcal{L}_G = \sum_{f=1}^{N_f} \bar{\psi}^f(x) \left(i\gamma^\mu D_\mu - m^f \right) \psi^f(x) - \frac{1}{4} F^a_{\mu\nu}(x) F^{a\mu\nu}(x). \tag{2.1}$$

Here, the sum in the first term runs over the individual flavors $f = 1, \ldots, N_f$ of the quarks which are described by Dirac 4-spinors $\psi^f(x)$. The eight massless gluons of QCD are represented by the gauge fields $A^a_\mu(x)$, contained in the covariant derivative and the field strength tensor [64],

$$D_\mu = \partial_\mu - igA^a_\mu(x)T^a, \quad T^a = \frac{\lambda^a}{2}, \tag{2.2}$$

$$F^a_{\mu\nu}(x) = \partial_\mu A^a_\nu(x) - \partial_\nu A^a_\mu(x) + g f^{abc} A^b_\mu(x) A^c_\nu(x). \tag{2.3}$$

The quark and gluon fields also carry Dirac indices and color indices in the fundamental representation, which have been suppressed in favor of a matrix/vector notation. g is the coupling constant of QCD; f^{abc} are the (totally antisymmetric) structure constants of SU(3) and T^a its generators, which satisfy the commutation relations $[T^a, T^b] = if^{abc}T^c$. Here, the upper Roman indices $a, b, c = 1, \ldots, 8$ index the adjoint representation of SU(3). The Dirac matrices obey the anticommutation relations $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$, where the metric is $g^{\mu\nu} = \eta^{\mu\nu}$ in Minkowski space.

Using this Lagrangian, it is possible to express the transition amplitude between two field configurations ϕ_1 and ϕ_2 at times t_1 and t_2, respectively, as a Feynman path integral [62],

$$\langle \phi_2, t_2 \,|\, \phi_1, t_1 \rangle = \int [d\bar{\psi}][d\psi][dA] \, \exp\left(i \int_{t_1}^{t_2} dt \int_{\mathbb{R}^3} d^3x \, \mathcal{L}_{\mathrm{QCD}} \right), \tag{2.4}$$

where the quantity in the exponent on the right-hand side is called the action. It is given by the space-time integral over the Lagrangian,

$$S_{\mathrm{QCD}} = \int dt \int d^3x \, \mathcal{L}_{\mathrm{QCD}}. \tag{2.5}$$

Two of the physically most interesting properties of QCD are asymptotic freedom and confinement. Both are linked to the dependence of the running coupling $\bar{\alpha}_s(Q^2) = \bar{g}^2(Q^2)/4\pi$ on the momentum transfer Q^2 [69]. This is described by the β-function [70, 71],

$$\beta(\bar{g}) = \frac{\partial \bar{g}(t)}{\partial t}, \quad t = \frac{1}{2}\ln\frac{Q^2}{\mu^2}, \tag{2.6}$$

where μ is the renormalization scale.

In the regime where the coupling is small, perturbation theory is applicable. To leading order (LO) in perturbation theory, the β-function of QCD is given by [70, 71]

$$\beta(g) = -\beta_0 g^3 + \mathcal{O}(g^5), \quad \beta_0 = \frac{1}{16\pi^2}\left(\frac{11}{3}N_c - \frac{2}{3}N_f\right), \tag{2.7}$$

where N_f is the number of active quark flavors, *i.e.*, flavors with $m^f \ll Q^2$ [69]. With $N_c = 3$, this means that as long as $N_f < 16.5$, $\beta(\bar{g}) = \partial\bar{g}(t)/\partial t < 0$ for small $\bar{g}$, so that $\bar{g} \to 0$ as $t \to \infty$. This is referred to as asymptotic freedom: at asymptotically large energy scales, the quarks behave like free particles. The LO solution of eq. (2.6) is then [70, 71]

$$\bar{g}^2(t) = \frac{g^2}{1 + 2\beta_0 g^2 t}, \tag{2.8}$$

where $g = \bar{g}(t = 0)$. Introducing the Λ-parameter of QCD at LO,

$$\Lambda^2_{\mathrm{QCD}} = \mu^2 e^{-\frac{1}{\beta_0 g^2}}, \tag{2.9}$$

and utilizing the definition of t in eq. (2.6), eq. (2.8) can be written as [69]

$$\bar{g}^2(Q^2) = \frac{1}{\beta_0 \ln(Q^2/\Lambda^2_{\mathrm{QCD}})}. \tag{2.10}$$

This shows again that $\bar{g}(Q^2) \to 0$ as $Q^2 \to \infty$.

For small momentum transfers, on the other hand, the coupling rises until perturbation theory and hence also eqs. (2.8) and (2.10) are no longer valid. The large coupling of QCD at low energies is connected to confinement [69]. This means that all physical states are (at low temperatures and baryon densities) singlets with respect to the color SU(3), so that quarks and gluons are not observable as final states [72, 73]. By contrast, they are confined within hadrons, bound states which form a singlet under SU(3). The classical examples are mesons ($q\bar{q}$) and baryons (qqq), but other combinations involving more than three quarks or only gluons are also possible. While numerical lattice studies clearly show that QCD is confining [74, 75], a rigorous mathematical proof is still outstanding.

To summarize the above discussion, the evolution of an experimentally measurable quantity called effective charge which has a strong connection with $\bar{\alpha}_s(Q^2)$ is shown in fig. 2.1.

2.2 Symmetries of the QCD action

The action of QCD can be constructed by requiring invariance under local SU(3) gauge transformations [64],

$$\psi^f(x) \to \Omega(x)\psi^f(x), \quad \bar{\psi}^f(x) \to \bar{\psi}^f(x)\Omega^\dagger(x), \quad \Omega(x) \in \mathrm{SU}(3). \tag{2.11}$$

It is straightforward to check that the mass term $m^f\bar{\psi}^f(x)\psi^f(x)$ is invariant. In order to make the derivative term gauge invariant, an operator $U(y, x)$ is required which ensures that

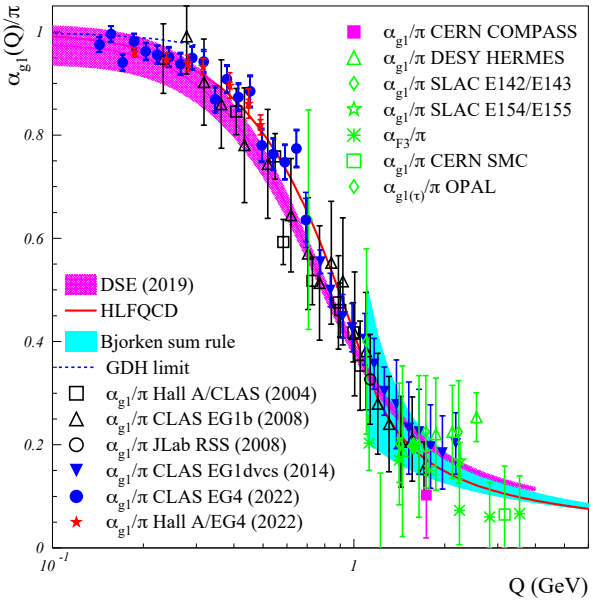


Figure 2.1: Experimental measurements of the effective charge $\alpha_{g_1}(Q)/\pi$ based on the generalized Bjorken sum rule and theoretical calculations of $\bar{\alpha}_s(Q)/\pi$ in the infrared regime [69]

$U(y,x)\psi(x)$ has the same transformation properties under gauge transformations as $\psi(y)$. Only in this way can two fields at different space-time points x and y be meaningfully compared. The transformation law of the above operator, which is called gauge transporter, hence needs to be [64]

$$U(y,x) \to \Omega(y)U(y,x)\Omega^\dagger(x). \tag{2.12}$$

Expanding the infinitesimal gauge transporter $U(x+\epsilon n, x)$ in terms of the separation ϵ, one introduces the gauge fields $A^a_\mu(x)$ [64],

$$U(x+\epsilon n, x) = 1 + ig\epsilon n^\mu A^a_\mu(x)T^a + \mathcal{O}(\epsilon^2). \tag{2.13}$$

Applying eq. (2.12) to the infinitesimal transporter in eq. (2.13), one obtains

$$\begin{aligned} U(x+\epsilon n, x) &\to \Omega(x+\epsilon n)U(x+\epsilon n, x)\Omega^\dagger(x) \\ &= \Omega(x+\epsilon n)\Omega^\dagger(x) + ig\epsilon n^\mu \Omega(x+\epsilon n)A^a_\mu(x)T^a\Omega^\dagger(x) + \mathcal{O}(\epsilon^2) \\ &= 1 + \epsilon n^\mu(\partial_\mu\Omega(x))\Omega^\dagger(x) + ig\epsilon n^\mu\Omega(x)A^a_\mu(x)T^a\Omega^\dagger(x) + \mathcal{O}(\epsilon^2) \\ &= 1 - \epsilon n^\mu\Omega(x)\partial_\mu\Omega^\dagger(x) + ig\epsilon n^\mu\Omega(x)A^a_\mu(x)T^a\Omega^\dagger(x) + \mathcal{O}(\epsilon^2). \end{aligned} \tag{2.14}$$

From the second to the third line, we have expanded $\Omega(x+\epsilon n) = \Omega(x) + \epsilon n^\mu\partial_\mu\Omega(x) + \mathcal{O}(\epsilon^2)$, and in the last step, we have made use of the fact that $(\partial_\mu\Omega(x))\Omega^\dagger(x) = \partial_\mu(\Omega(x)\Omega^\dagger(x)) - \Omega(x)(\partial_\mu\Omega^\dagger(x)) = \partial_\mu 1 - \Omega(x)(\partial_\mu\Omega^\dagger(x)) = -\Omega(x)(\partial_\mu\Omega^\dagger(x))$. From eq. (2.14), it is clear that the gauge fields transform according to [64]

$$A^a_\mu(x)T^a \to \Omega(x)\left(A^a_\mu(x)T^a + \frac{i}{g}\partial_\mu\right)\Omega^\dagger(x). \tag{2.15}$$

The infinitesimal gauge transporter can be used to define the covariant derivative as was done in eq. (2.2). With the help of this definition, it can be shown that $D_\mu\psi^f(x)$ has the same

transformation law as $\psi^f(x)$ itself, *i.e.*, it transforms in the fundamental representation as well,

$$\begin{aligned} D_\mu\psi^f(x) \to & \left[\partial_\mu - ig\Omega(x)\left(A^a_\mu(x)T^a + \frac{i}{g}\partial_\mu\right)\Omega^\dagger(x)\right]\Omega(x)\psi^f(x) \\ &= (\partial_\mu\Omega(x))\psi^f(x) + \Omega(x)\partial_\mu\psi^f(x) - ig\Omega(x)A^a_\mu(x)T^a\Omega^\dagger(x)\Omega(x)\psi^f(x) \\ &\quad + \Omega(x)(\partial_\mu\Omega^\dagger(x))\Omega(x)\psi^f(x) \\ &= \Omega(x)(\partial_\mu - igA^a_\mu(x)T^a)\psi^f(x) + [\partial_\mu\Omega(x) + \Omega(x)(\partial_\mu\Omega^\dagger(x))\Omega(x)]\psi^f(x) \\ &= \Omega(x)D_\mu\psi^f(x). \end{aligned} \tag{2.16}$$

As a consequence, the covariant derivative term in eq. (2.1) is also gauge invariant.

Noticing that the field strength tensor is given by the commutator of two covariant derivative operators [64],

$$\begin{aligned} [D_\mu, D_\nu] &= [\partial_\mu - igA^a_\mu(x)T^a, \partial_\nu - igA^b_\nu(x)T^b] \\ &= -ig(\partial_\mu A^b_\nu(x))T^b + ig(\partial_\nu A^a_\mu(x))T^a - g^2A^a_\mu(x)A^b_\nu(x)[T^a, T^b] \\ &= -ig\left(\partial_\mu A^a_\nu(x)T^a - \partial_\nu A^a_\mu(x)T^a + gA^a_\mu(x)A^b_\nu(x)f^{abc}T^c\right) \\ &= -ig\left(\partial_\mu A^a_\nu(x) - \partial_\nu A^a_\mu(x) + gA^b_\mu(x)A^c_\nu(x)f^{bca}\right)T^a = -igF^a_{\mu\nu}(x)T^a, \end{aligned} \tag{2.17}$$

it can be seen that it transforms in the adjoint representation, *i.e.*, according to

$$F^a_{\mu\nu}(x)T^a \to \Omega(x)F^a_{\mu\nu}(x)T^a\Omega^\dagger(x), \tag{2.18}$$

rendering $\mathcal{L}_G = -1/2\,\mathrm{tr}[F^a_{\mu\nu}(x)T^aF^{b\mu\nu}(x)T^b] = -1/4F^a_{\mu\nu}(x)F^{a\mu\nu}(x)$ also gauge invariant.[1] It is worth noting that it is not possible to formulate a kinetic term for the gauge fields without also including self-interactions [63]. In particular, writing out $\mathcal{L}_G$ in terms of the gauge fields $A^a_\mu(x)$, one discovers terms proportional to A^3 and A^4, leading to the aforementioned gluonic self-interactions [64]. Their appearance can thus be understood as a direct consequence of the non-Abelian nature of SU(3).[2]

In principle, by using higher-dimensional operators, one can form many gauge-invariant combinations in addition to the ones used in eq. (2.1). These are, however, not renormalizable and thus physically irrelevant at ordinary energies [63, 64].

Apart from the local SU(3) gauge invariance associated with color, the QCD Lagrangian in eq. (2.1) has a global U(1) symmetry for each quark flavor with a non-degenerate mass,

$$\psi^f(x) \to e^{-i\alpha^f}\psi^f(x). \tag{2.19}$$

The corresponding Noether current is $j^{\mu f}(x) = \bar{\psi}^f(x)\gamma^\mu\psi^f(x)$, implying the conservation of the net quark number of each such flavor [65].

If N_{fg} of the N_f quark masses are degenerate, there is a larger global symmetry for these flavors. It is useful to separate the U(1) and SU(N_{fg}) parts of these transformations,

$$\psi^f(x) \to e^{-i\alpha}\psi^f(x) \tag{2.20}$$

and

$$\psi^f(x) \to \psi'^{f'}(x) = \sum_f U^{f'f}\psi^f(x), \quad U \in \mathrm{SU}(N_{fg}), \tag{2.21}$$

[1] We have used that $\mathrm{tr}(T^aT^b) = 1/2\delta^{ab}$ [64].

[2] The self-interaction terms are proportional to the structure constants through the last term in eq. (2.3), which vanish for Abelian Lie groups such as U(1).

respectively [64]. Estimates of the masses of the two lightest quarks u and d [76],

$$m_u = 2.16^{+0.49}_{-0.26}\ \text{MeV}, \quad m_d = 4.67^{+0.48}_{-0.17}\ \text{MeV}, \tag{2.22}$$

show that they are nearly degenerate compared to the energy scale of the QCD interactions, $\Lambda_{\text{QCD}} \sim 0.3$ GeV [69]. The SU(N_{fg}) symmetry discussed above with $N_{fg} = 2$ can thus be identified with isospin, while the U(1) part corresponds to the conservation of the total net quark number [64].

For massless u- and d-quarks, the QCD Lagrangian is furthermore invariant under axial SU(2) × U(1) transformations,

$$\psi^f(x) \to e^{-i\alpha\gamma^5}\psi^f(x) \tag{2.23}$$

and

$$\psi^f(x) \to e^{-i\omega^a\tau^a\gamma_5}\psi^f(x), \tag{2.24}$$

respectively, where $\tau^a = \sigma^a/2$ are the generators of SU(2) [64]. Noticing that the massless QCD Lagrangian can be decomposed into separate terms for left- and right-handed quarks, the vector and axial transformations mix, leading to two independent chiral transformations for the left- and right-handed components. The total symmetry group (neglecting color) of the classical QCD Lagrangian for massless and degenerate u- and d-quarks is hence $\text{SU}(2)_L \times \text{SU}(2)_R \times \text{U}(1)_L \times \text{U}(1)_R$ [64].

The vacuum state, however, mixes the two quark helicities, so that the full symmetry group is spontaneously broken down to the subgroup of vector symmetries, $\text{SU}(2)_L \times \text{SU}(2)_R \times \text{U}(1)_L \times \text{U}(1)_R \to \text{SU}(2)_V \times \text{U}(1)_V$ [64]. According to Goldstone's theorem [77–79], this would imply the emergence of four massless spin-zero particles. The real strong interactions do not contain any massless particles because the u- and d-quarks are not exactly massless. Nevertheless, they do contain an isospin triplet of relatively light pseudoscalar mesons, the pions, which can be associated with the spontaneous breaking of $\text{SU}(2)_A$. On the other hand, a light isospin-singlet pseudoscalar meson which could be associated with the spontaneous breaking of $\text{U}(1)_A$ is not known to exist.

To shed more light onto this situation, it is important to note that our treatment of symmetries so far has been purely classical. In order for a symmetry of the action to be one of the full quantum theory, the measure of the functional integral in eq. (2.4) also needs to be invariant. For axial transformations, this requirement is in general not fulfilled [64]. This is referred to as axial anomaly. It can be shown, however, that the relevant group theory factor vanishes in QCD in the case of the axial isospin-triplet current, but not in the isospin-singlet case [64]. This means that while $\text{SU}(2)_A$ is in fact a symmetry of QCD with two massless and degenerate quarks, $\text{U}(1)_A$ is not.

In nature, all symmetries discussed above except for $\text{U}(1)_V$ and the color SU(3) receive a small explicit breaking due to the finite and not exactly degenerate masses of the u- and d-quarks. This thesis will employ the approximation of $m_u = m_d$ and thus an exact $\text{SU}(2)_V$ symmetry.

2.3 Wick's theorem

A useful tool for the evaluation of expectation values is Wick's theorem. In a free field theory, it allows one to express the vacuum expectation value of a time-ordered product of operators in terms of propagators [64]. Here, we review the derivation of Wick's theorem for bosons and fermions using the path integral quantization.

2.3.1 Wick's theorem for bosons

Consider a real bosonic field ϕ with a Lagrange density which is quadratic in the field,

$$\mathcal{L}[\phi] = \frac{1}{2}\phi(x)(-\partial^2 - m^2)\phi(x). \tag{2.25}$$

Introducing the generating functional

$$Z[J] = \int [d\phi]\, \exp\left[i \int d^4x \left(\mathcal{L}[\phi(x)] + J(x)\phi(x) \right) \right], \tag{2.26}$$

correlation functions can be computed by taking functional derivatives of $Z[J]$ with respect to the source field J [64],

$$\begin{aligned} \langle 0 \,|\, T\phi(x_1)\cdots\phi(x_n) \,|\, 0 \rangle &= \frac{1}{Z[0]} \int [d\phi]\, \phi(x_1)\cdots\phi(x_n) \exp\left[i \int d^4x\, \mathcal{L} \right] \\ &= \frac{(-i)^n}{Z[0]} \left. \frac{\delta^n Z[J]}{\delta J(x_1)\cdots\delta J(x_n)} \right|_{J=0}. \end{aligned} \tag{2.27}$$

In a free field theory defined as in eq. (2.25), $Z[J]$ can be brought into a more explicit form. For this purpose, we introduce the following change of variables,

$$\phi'(x) = \phi(x) - i \int d^4y\, D_F(x-y) J(y). \tag{2.28}$$

Here, $D_F(x-y)$ is the Feynman propagator, which is a Green's function of the operator $(-\partial^2 - m^2)$ [64],

$$(-\partial_x^2 - m^2) D_F(x-y) = i\delta^{(4)}(x-y). \tag{2.29}$$

We start by completing the square in eq. (2.26),

$$\begin{aligned} Z[J] = \int [d\phi]\, \exp\Bigg[i \int d^4x\, \frac{1}{2} \left(\phi(x) - i\int d^4y\, D_F(x-y)J(y) \right) (-\partial^2 - m^2) \\ \left(\phi(x) - i\int d^4y\, D_F(x-y)J(y) \right) \\ - \frac{i}{2} \int d^4x \int d^4y\, J(x)[-iD_F(x-y)]J(y) \Bigg]. \end{aligned} \tag{2.30}$$

Now we apply the change of variables from ϕ to ϕ' to the functional integral. As this is nothing but a constant shift, the Jacobian of the transformation is 1, so that the measure remains invariant. Hence, the result is

$$\begin{aligned} Z[J] &= \int [d\phi']\, \exp\left[i \int d^4x\, \frac{1}{2} \phi'(x)(-\partial^2 - m^2)\phi'(x) \right] \\ &\qquad \exp\left[-\frac{i}{2} \int d^4x \int d^4y\, J(x)[-iD_F(x-y)]J(y) \right] \\ &= Z[0] \exp\left[-\frac{1}{2} \int d^4x \int d^4y\, J(x) D_F(x-y) J(y) \right]. \end{aligned} \tag{2.31}$$

With this form of the generating functional at hand, the derivatives in eq. (2.27) can be evaluated, so that we obtain for the correlation function of an even number n of fields,

$$\langle 0 \,|\, T\phi(x_1)\cdots\phi(x_n) \,|\, 0 \rangle = \sum_{\sigma \in P_n} D_F(x_{\sigma(1)} - x_{\sigma(2)}) \cdots D_F(x_{\sigma(n-1)} - x_{\sigma(n)}). \tag{2.32}$$

Here, P_n is the set of all pairings of n numbers. This result is known as Wick's theorem, and the individual factors $D_F(x_{\sigma(k-1)} - x_{\sigma(k)})$ are also referred to as contractions of the fields $\phi(x_{\sigma(k-1)})$ and $\phi(x_{\sigma(k)})$ [64, 66]. From the above derivation it also follows that the correlation function of an odd number of fields vanishes.

2.3.2 Wick's theorem for fermions

For fermions, the derivation of Wick's theorem proceeds analogously to the one for bosons above. The main differences are that extra caution needs to be applied due to the anticommuting nature of fermions, and that two separate source fields are required for ψ and $\bar{\psi}$.

The generating functional for the free Dirac Lagrangian

$$\mathcal{L}[\bar{\psi}, \psi] = \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m)\psi(x) \tag{2.33}$$

is given by [64]

$$Z[\bar{\eta}, \eta] = \int [d\bar{\psi}][d\psi] \, \exp\left[i \int d^4x \left(\mathcal{L}[\bar{\psi}(x), \psi(x)] + \bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x) \right) \right], \tag{2.34}$$

where η and $\bar{\eta}$ are the Grassmann-valued source fields for $\bar{\psi}$ and ψ, respectively. The analogue of eq. (2.27) is [62, 64, 66]

$$\begin{aligned} \left\langle 0 \,\middle|\, T\bar{\psi}(x_1)\psi(y_1)\cdots\bar{\psi}(x_n)\psi(y_n) \,\middle|\, 0 \right\rangle &= \frac{1}{Z[0,0]} \int [d\bar{\psi}][d\psi] \, \bar{\psi}(x_1)\psi(y_1)\cdots\bar{\psi}(x_n)\psi(y_n) \\ &\qquad \exp\left[i \int d^4x \, \mathcal{L} \right] \\ &= \frac{(-1)^n (-i)^{2n}}{Z[0,0]} \left. \frac{\delta^{2n} Z[\bar{\eta}, \eta]}{\delta\eta(x_1)\delta\bar{\eta}(y_1)\cdots\delta\eta(x_n)\delta\bar{\eta}(y_n)} \right|_{\eta=\bar{\eta}=0}. \end{aligned} \tag{2.35}$$

In order to derive a more explicit form of the generating functional for the free Dirac theory, we introduce the following change of variables,

$$\psi'(x) = \psi(x) - i \int d^4y \, S_F(x-y)\eta(y), \quad \bar{\psi}'(x) = \bar{\psi}(x) - i \int d^4y \, \bar{\eta}(y) S_F(x-y), \tag{2.36}$$

where S_F is the fermionic Feynman propagator, which fulfills the analogue of eq. (2.29). Completing the square in eq. (2.34) leads to

$$\begin{aligned} Z[\bar{\eta}, \eta] = \int [d\bar{\psi}][d\psi] \, \exp\Bigg[i \int d^4x \Bigg(\bar{\psi}(x) - i \int d^4y \, \bar{\eta}(y) S_F(x-y) \Bigg) (i\gamma^\mu \partial_\mu - m) \\ \Bigg(\psi(x) - i \int d^4y \, S_F(x-y)\eta(y) \Bigg) \\ - i \int d^4x \int d^4y \, \bar{\eta}(x)[-iS_F(x-y)]\eta(y) \Bigg]. \end{aligned} \tag{2.37}$$

Now we apply the change of variables defined by eq. (2.36) to the functional integral. Again, this is just a constant shift, so that we obtain

$$\begin{aligned} Z[\bar{\eta}, \eta] &= \int [d\bar{\psi}'][d\psi'] \, \exp\left[i \int d^4x \, \bar{\psi}'(x)(i\gamma^\mu \partial_\mu - m)\psi'(x) \right] \\ &\qquad \exp\left[-i \int d^4x \int d^4y \, \bar{\eta}(x)[-iS_F(x-y)]\eta(y) \right] \\ &= Z[0,0] \exp\left[- \int d^4x \int d^4y \, \bar{\eta}(x) S_F(x-y)\eta(y) \right]. \end{aligned} \tag{2.38}$$

Evaluating the derivatives in eq. (2.35) then yields for the correlation function,

$$\begin{aligned} \left\langle 0 \,\middle|\, T\bar{\psi}(x_1)\psi(y_1)\cdots\bar{\psi}(x_n)\psi(y_n) \,\middle|\, 0 \right\rangle = (-1)^n \sum_{\sigma \in P_{2n}} \operatorname{sgn}(\sigma) S_F(y_{\sigma(1)} - x_{\sigma(2)}) \\ \cdots S_F(y_{\sigma(2n-1)} - x_{\sigma(2n)}) \end{aligned} \tag{2.39}$$

This is Wick's theorem for fermions [64, 66]. Here, the sign of the pairing σ is defined in the following way: Since pairings that only differ in the order of the pairs are not counted as different, summing over all such pairings of $2n$ numbers is equivalent to summing over the $n!$ permutations of n numbers (either the xs or the ys). The sign of σ is then defined as the sign of this permutation [62].

2.4 Euclidean formulation

It turns out that it is difficult to give a satisfactory mathematical meaning to the path integral which defines the transition amplitudes (*cf.*, eq. (2.4)) and the expectation values of observables (*cf.*, eqs. (2.27) and (2.35)) [80]. This is due to the fact that the integral is complex and strongly oscillating because of the complex weight e^{iS} in the integrand. For a numerical treatment of the theory this behavior is also highly problematic. It is therefore common to switch to a purely imaginary time coordinate by performing a so-called Wick rotation [80],

$$x^0 = -ix_0^{(E)}, \quad x^k = x_k^{(E)}, \tag{2.40}$$

$$p^0 = -ip_0^{(E)}, \quad p^k = p_k^{(E)}. \tag{2.41}$$

This is illustrated in fig. 2.2. Here, the curved arrows indicate how the time-ordered correlation functions of the Minkowski theory (defined on the real x^0-axis) can be obtained from the Euclidean correlation functions (defined on the imaginary x^0-axis) [80]. The analytic continuation of the Euclidean correlation functions back to Minkowski space is contingent upon a property called reflection positivity [80, 81],

$$\sum_{n,m} \int d^4x_1^{(E)} \cdots d^4x_n^{(E)} d^4y_1^{(E)} \cdots d^4y_m^{(E)}\, f_n^*(\theta x_1^{(E)}, \ldots, \theta x_n^{(E)}) f_m(y_1^{(E)}, \ldots, y_m^{(E)}) \\ \left\langle 0 \,\middle|\, \phi(x_1^{(E)}) \cdots \phi(x_n^{(E)}) \phi(y_1^{(E)}) \cdots \phi(x_m^{(E)}) \,\middle|\, 0 \right\rangle \geq 0. \tag{2.42}$$

Here, the f_n are arbitrary test functions which have support only in the positive-time, time-ordered subset of n coordinates ($0 < x_{1,0}^{(E)} < \cdots < x_{n,0}^{(E)} < \infty$). θ denotes Euclidean time reflection, $\theta(x_0^{(E)}, \mathbf{x}^{(E)}) = (-x_0^{(E)}, \mathbf{x}^{(E)})$. If the Euclidean correlation functions satisfy eq. (2.42) (and a set of additional mathematical axioms), the Osterwalder-Schrader theorem states that they can be analytically continued to Minkowski space, and that the resulting correlation functions can be used to define a Quantum Field Theory [81, 82].

The space-time metric for the coordinates $x_0^{(E)}, \ldots, x_3^{(E)}$ is Euclidean, $g_{\mu\nu} = \delta_{\mu\nu}$. For the covariant spatial components of the coordinate vectors, eq. (2.40) implies a sign change, $x_k = -x^k = -x_k^{(E)}$. Moreover, the zero-component of the Euclidean four-momentum vector can no longer be interpreted as energy of the particle, but rather $p_0^{(E)} = iE_{\mathbf{p}}$. Derivatives need to be rotated in the opposite direction,

$$\partial_0 = i\partial_0^{(E)}, \quad \partial_k = \partial_k^{(E)}. \tag{2.43}$$

The expression for the Euclidean transition amplitude looks like

$$\langle \phi_2, \tau_2 \,|\, \phi_1, \tau_1 \rangle = \int [d\bar{\psi}][d\psi][dA] \, \exp\left(- \int_{\tau_1}^{\tau_2} d\tau \int_{\mathbb{R}^3} d^3x^{(E)} \, \mathcal{L}_{\mathrm{QCD}}^{(E)} \right), \tag{2.44}$$

where $\tau = x_0^{(E)} \in \mathbb{R}$. The weight factor is now given by the exponential of the Euclidean action $S^{(E)} = -iS$ [80], the integral over Euclidean space-time of $\mathcal{L}^{(E)} = -\mathcal{L}$.

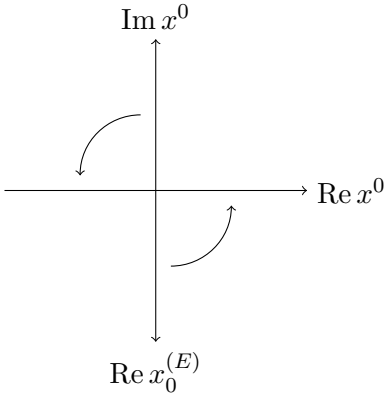


Figure 2.2: Illustration of the Wick rotation, indicated by curved arrows (adapted from Ref. [80])

In order to derive an explicit expression for the Euclidean QCD Lagrangian, the transformation of the Dirac matrices and the gauge fields under Wick rotation needs to be settled. The Euclidean Dirac matrices need to fulfill $\left\{\gamma_\mu^{(E)}, \gamma_\nu^{(E)}\right\} = 2\delta_{\mu\nu}$. This is satisfied by defining

$$\gamma_0 = \gamma_0^{(E)}, \quad \gamma_k = -\gamma^k = -i\gamma_k^{(E)}, \tag{2.45}$$

while we set

$$\gamma_5^{(E)} = \gamma_0^{(E)}\gamma_1^{(E)}\gamma_2^{(E)}\gamma_3^{(E)} = i\gamma^0\gamma^1\gamma^2\gamma^3 = \gamma^5. \tag{2.46}$$

These definitions imply that $\left(\gamma_\mu^{(E)}\right)^\dagger = \gamma_\mu^{(E)}$ for $\mu = 0, 1, 2, 3, 5$. For the derivative term in the free Dirac Lagrangian, we obtain from eqs. (2.43) and (2.45),

$$i\gamma^\mu\partial_\mu = i\gamma^0\partial_0 + i\gamma^k\partial_k = -\gamma_0^{(E)}\partial_0^{(E)} - \gamma_k^{(E)}\partial_k^{(E)} = -\gamma_\mu^{(E)}\partial_\mu^{(E)}. \tag{2.47}$$

To derive the transformation of the gauge field, it is instructive to consider the infinitesimal gauge transporter $U(x+\epsilon n, x)$. According to eq. (2.13), demanding that $U(x+\epsilon n, x)$ is invariant under Wick rotation is equivalent to demanding that $\epsilon n^\mu A_\mu^a(x) = A_\mu^a(x)dx^\mu$ is; considering furthermore that

$$A_\mu^a dx^\mu = A_0^a dx^0 + A_k^a dx^k = -iA_0^a dx_0^{(E)} + A_k^a dx_k^{(E)} \tag{2.48}$$

leads to the conclusion that the Wick rotation of the gauge field must be

$$A_0^a = iA_0^{a(E)}, \quad A_k^a = A_k^{a(E)}. \tag{2.49}$$

As we demand that the covariant derivative transforms in the same way under Wick rotation as the partial derivative does, *i.e.*, $i\gamma^\mu D_\mu = -\gamma_\mu^{(E)} D_\mu^{(E)}$, we have to satisfy

$$i\gamma^0 D_0 = i\gamma^0(\partial_0 - igA_0^a T^a) = -\gamma_0^{(E)}\left(\partial_0^{(E)} - igA_0^{a(E)}T^a\right) = -\gamma_0^{(E)} D_0^{(E)}, \tag{2.50}$$

$$i\gamma^k D_k = i\gamma^k(\partial_k - igA_k^a T^a) = -\gamma_k^{(E)}\left(\partial_k^{(E)} - igA_k^{a(E)}T^a\right) = -\gamma_k^{(E)} D_k^{(E)}. \tag{2.51}$$

This means that we need to define $D_\mu^{(E)} = \partial_\mu^{(E)} - igA_\mu^{a(E)}T^a$, which has the same form as in Minkowski space (*cf.*, eq. (2.2)). For the transformation of the field strength tensor, one finds

according to eq. (2.3),

$$F^a_{0k} = \partial_0 A^a_k - \partial_k A^a_0 + g f^{abc} A^b_0 A^c_k = i\partial_0^{(E)} A_k^{a(E)} - i\partial_k^{(E)} A_0^{a(E)} + i g f^{abc} A_0^{b(E)} A_k^{c(E)} = i F_{0k}^{a(E)}, \tag{2.52}$$

$$F^a_{jk} = \partial_j A^a_k - \partial_k A^a_j + g f^{abc} A^b_j A^c_k = \partial_j^{(E)} A_k^{a(E)} - \partial_k A_j^{a(E)} + g f^{abc} A_j^{b(E)} A_k^{c(E)} = F_{jk}^{a(E)}, \tag{2.53}$$

if one defines the field strength tensor in Euclidean space to have the same form as in Minkowski space (*cf.*, eq. (2.3)). Analogously, one finds $F^{a0k} = -iF_{0k}^{a(E)}$ and $F^{ajk} = F_{jk}^{a(E)}$, while the diagonal components of $F^a_{\mu\nu}$ all vanish. This means that the product of two field strength tensors, which appears in the QCD Lagrangian, is invariant under Wick rotation, $F^a_{\mu\nu} F^{a\mu\nu} = F_{\mu\nu}^{a(E)} F_{\mu\nu}^{a(E)}$.

Noting furthermore that $\bar{\psi} = \psi^\dagger \gamma^0 = \psi^\dagger \gamma_0^{(E)} = \bar{\psi}^{(E)}$, the Euclidean QCD Lagrangian reads [83]

$$\mathcal{L}^{(E)}_{\mathrm{QCD}} = -\mathcal{L}_{\mathrm{QCD}} = \sum_{f=1}^{N_f} \bar{\psi}^{f(E)} \left(\gamma_\mu^{(E)} D_\mu^{(E)} + m^f \right) \psi^f(x) + \frac{1}{4} F_{\mu\nu}^{a(E)} F_{\mu\nu}^{a(E)}. \tag{2.54}$$

Another major advantage of the Euclidean formulation of Quantum Field Theories, in addition to the benefits outlined at the beginning of this section, is that it exhibits a structural equivalence with statistical mechanics [64, 66, 83]. The canonical partition function of a quantum statistical system has the form [66, 83]

$$Z(\beta) = \mathrm{tr}\, e^{-\beta H}, \tag{2.55}$$

where H is the Hamiltonian of the system and $\beta = 1/T$ the inverse temperature. This has a striking similarity with the path integral in eq. (2.44), if one makes the formal identification $\beta H \leftrightarrow S^{(E)}$ and imposes (anti-)periodic boundary conditions on the paths [66, 83]. Pursuing this formal identification, one arrives at the expression for the partition function of a thermal Euclidean Quantum Field Theory [66],

$$Z(\beta) = \int [d\phi]\, \exp\left(-S^{(E)}[\phi] \right), \quad S^{(E)}[\phi] = \int_{-\beta/2}^{\beta/2} d\tau \int_{\mathbb{R}^3} d^3x^{(E)}\, \mathcal{L}^{(E)}[\phi], \tag{2.56}$$

with periodic (antiperiodic) boundary conditions in Euclidean time on the bosonic (fermionic) field ϕ [83].

In this thesis, we are only interested in the zero-temperature properties of QCD. The corresponding partition function relevant for the computation of vacuum expectation values can be recovered from eq. (2.56) by taking the limit $\beta \to \infty$ [83],

$$Z_{\mathrm{QCD}} = \int [d\bar{\psi}][d\psi][dA]\, \exp\left(-S^{(E)}_{\mathrm{QCD}} \right), \quad S^{(E)}_{\mathrm{QCD}} = \int_{-\infty}^{\infty} d\tau \int_{\mathbb{R}^3} d^3x^{(E)}\, \mathcal{L}^{(E)}_{\mathrm{QCD}}. \tag{2.57}$$

Moreover, we will omit the superscripts (E) in the following, so that all quantities are implicitly Euclidean, unless otherwise stated.

3 QCD on the lattice

The Euclidean formulation of QCD cures the problem of having an imaginary exponent in the integrand as long as the Euclidean action S_{QCD} is real. The path integral in eq. (2.57), however, is still an infinite-dimensional integral. Therefore, one introduces a hypercubic lattice with a lattice spacing a,

$$\Lambda = \{\, x = a(n_0, n_1, n_2, n_3) \mid n_0 = 0, 1, \ldots, N_\tau - 1;\ n_{1,2,3} = 0, 1, \ldots, N_s - 1 \,\}\,, \tag{3.1}$$

to replace space-time. The spatial extent of the lattice is given by $L = aN_s$ and the temporal extent by $T = aN_\tau$. Hence, the lattice consists of a total of $N_s^3 N_\tau$ points.

The introduction of a finite lattice spacing a has two major advantages: On the one hand, the number of dimensions in the path integral of eq. (2.57) becomes finite, making it mathematically well-defined and numerically computable. On the other hand, the lattice provides a natural UV-regulator for the Quantum Field Theory by introducing a momentum cut-off $2\pi/a$, while preserving gauge invariance [80].

In order to extrapolate results computed on a finite lattice to the physically relevant ones, it is useful to distinguish two different limiting procedures: For the infinite-volume limit, one keeps the lattice spacing a fixed while taking N_s and thus $L = aN_s$ to infinity. In this limit, the allowed momenta become continuous, but the cut-off remains intact. For the continuum limit, one keeps the physical volume $L = aN_s$ fixed by simultaneously sending $N_s \to \infty$ and $a \to 0$, so that the momenta remain discrete, but the cut-off is removed. In order to suppress finite-temperature effects, the temporal extent $T = aN_\tau$ also needs to be large, regardless of the lattice spacing, as indicated by the boundaries of the τ-integral in eq. (2.57).

Another important point concerns the correct boundary conditions. For bosonic fields, the appropriate boundary conditions are periodic in all directions. For fermions, as has already been mentioned in section 2.4, the boundary conditions have to be chosen antiperiodic in time and periodic in space [80].

In the following, we will discuss the lattice discretization of the QCD action and the pertinent operators as well as methods for the numerical evaluation of the discretized path integral. Parts of the present chapter are inspired by the PhD theses [67, 68, 84].

3.1 Discretizing the Euclidean QCD action

The first critical step towards applying the lattice method to QCD is to find a suitable discretization of the QCD action defined by the Lagrangian in eq. (2.54). We will start the discussion with the gauge part in section 3.1.1, as this is relatively straightforward, and deal with the somewhat more involved introduction of fermion degrees of freedom in section 3.1.2.

3.1.1 The gauge action

For the transcription of the concept of a gauge field to a discrete lattice, it is useful to consider the gauge transporter introduced in eq. (2.12). On a lattice with shortest distance a, the elementary gauge transporters are associated with the links connecting nearest neighboring

points x and $x + a\hat{\mu}$. Following eq. (2.13), we introduce the lattice gauge field $\mathcal{A}^a_\mu(x)$ by defining the gauge transporter corresponding to the aforementioned link [80][1],

$$U_\mu(x) = e^{-iga\mathcal{A}^a_\mu(x)T^a} \in \mathrm{SU}(3). \tag{3.2}$$

Here, the path-ordered integral over $A^a_\mu T^a$ along the path from x to $x+a\hat{\mu}$ has been approximated by $a\mathcal{A}^a_\mu(x)T^a$, meaning that the lattice gauge field $\mathcal{A}^a_\mu$ is only equal to the continuum gauge field A^a_μ up to lattice artefacts of $\mathcal{O}(a^2)$ [83]. This is the reason why the gauge fields need to be formulated on the lattice in terms of SU(3) group elements instead of elements of the corresponding algebra as in the continuum: simply defining a covariant derivative analogously to eq. (2.2) would lead to violations of gauge invariance in higher orders of the lattice spacing [85]. Analogously to eq. (3.2), we define the gauge transporter from point x to $x - a\hat{\mu}$,

$$U_{-\mu}(x) = U^\dagger_\mu(x - a\hat{\mu}) = e^{iga\mathcal{A}^a_\mu(x-a\hat{\mu})T^a}. \tag{3.3}$$

Since the Lagrangian needs to be gauge invariant, it has to be constructed from gauge-invariant quantities built from the gauge links defined above. The simplest one is, according to eq. (2.12), a trace over a closed loop of gauge links.

Wilson gauge action

The action proposed by Wilson [73] for lattice gauge theory only depends on the shortest, non-trivial closed loop of gauge links, the so-called plaquette (*cf.*, fig. 3.1),

$$\begin{aligned} U_p = U_{\mu\nu}(x) &= U_\mu(x)U_\nu(x + a\hat{\mu})U_{-\mu}(x + a\hat{\mu} + a\hat{\nu})U_{-\nu}(x + a\hat{\nu}) \\ &= U_\mu(x)U_\nu(x + a\hat{\mu})U^\dagger_\mu(x + a\hat{\nu})U^\dagger_\nu(x). \end{aligned} \tag{3.4}$$

The Wilson gauge action is now given by a sum over all plaquettes, including every plaquette with only one orientation [83],

$$S_{G,W}[U] = \frac{2}{g^2}\sum_p \mathrm{Re}\,\mathrm{tr}(\mathbb{1} - U_p) = \frac{\beta}{3}\sum_{x\in\Lambda}\sum_{\mu<\nu}\mathrm{Re}\,\mathrm{tr}[\mathbb{1} - U_{\mu\nu}(x)]. \tag{3.5}$$

Here, the inverse lattice coupling $\beta = 6/g^2$ has been introduced. This action is by construction gauge invariant, real and positive [80].

In order to prove that in the naive continuum limit $a \to 0$, the Wilson gauge action eq. (3.5) converges to the gauge part of the Euclidean continuum QCD action defined by the second term in eq. (2.54), we need to expand the gauge links $U_\mu(x)$ in terms of the lattice gauge fields $\mathcal{A}^a_\mu(x)$ as given in eq. (3.2). Thus, the plaquette takes the form of a product of exponentials of matrices, to which we apply the Baker-Campbell-Hausdorff formula [85],

$$\begin{aligned} \exp(\epsilon A)\exp(\epsilon B) = \exp\Big\{&\epsilon A + \epsilon B + \frac{\epsilon^2}{2}[A,B] + \frac{\epsilon^3}{12}([A,[A,B]] + [B,[B,A]]) \\ &- \frac{\epsilon^4}{24}[B,[A,[A,B]]] + \mathcal{O}(\epsilon^5)\Big\}, \end{aligned} \tag{3.6}$$

where A and B are arbitrary matrices. By applying eq. (3.6) iteratively to eq. (3.4) and expanding $\mathcal{A}^a_\nu(x + a\hat{\mu}) = A^a_\nu(x) + a\partial_\mu A^a_\nu(x) + a^2\partial^2_\mu A^a_\nu(x) + a^3\partial^3_\mu A^a_\nu(x) + \mathcal{O}(a^4)$, one obtains an expression of the form [85]

$$U_{\mu\nu}(x) = \exp\left[iga^2\mathcal{F}^a_{\mu\nu}(x)T^a + a^3E^{(3)}_{\mu\nu}(x) + a^4E^{(4)}_{\mu\nu}(x) + \mathcal{O}(a^5)\right], \tag{3.7}$$

[1] Note that $U_\mu(x)$ corresponds to $U(x, x + a\hat{\mu})$, not $U(x + a\hat{\mu}, x)$, therefore the change of sign compared to eq. (2.13).

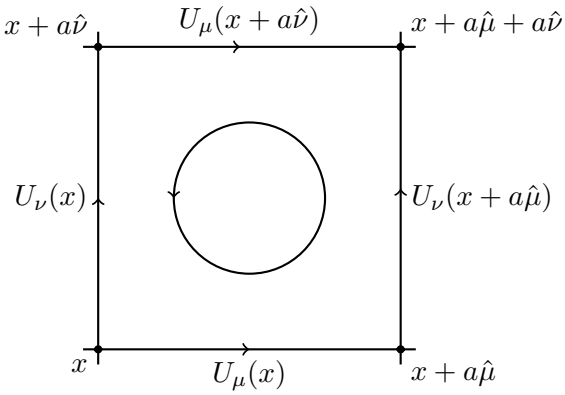


Figure 3.1: The four gauge links which build up the plaquette $U_{\mu\nu}(x)$. The circle indicates the order that the links are run through in the plaquette [83].

where the terms labelled $E^{(3,4)}_{\mu\nu}(x)$ are antisymmetric in μ and ν, *i.e.*, $E^{(3,4)}_{\mu\nu}(x) = -E^{(3,4)}_{\nu\mu}(x)$ [85], and $\mathcal{F}^a_{\mu\nu}(x)$ is the field strength tensor built like in eq. (2.3), but from the lattice gauge fields $\mathcal{A}$. Noting that $U_{\mu\nu}(x) \in \mathrm{SU}(3)$, we have that

$$\mathrm{tr}[U_{\mu\nu}(x)]^* = \mathrm{tr}[U^\dagger_{\mu\nu}(x)] = \mathrm{tr}[U_{\nu\mu}(x)]. \tag{3.8}$$

In an expansion of the exponential in eq. (3.7) for small a, the terms involving $E^{(3,4)}_{\mu\nu}(x)$ thus do not contribute to the real part of its trace [85],

$$\mathrm{Re}\,\mathrm{tr}[E^{(3,4)}_{\mu\nu}(x)] = \frac{1}{2}\left\{\mathrm{tr}[E^{(3,4)}_{\mu\nu}(x)] + \mathrm{tr}[E^{(3,4)}_{\mu\nu}(x)]^*\right\} = \frac{1}{2}\left\{\mathrm{tr}[E^{(3,4)}_{\mu\nu}(x)] + \mathrm{tr}[E^{(3,4)}_{\nu\mu}(x)]\right\} = 0. \tag{3.9}$$

For the Wilson gauge action, we hence obtain

$$\begin{aligned} S_{G,W}[U] &= \frac{\beta}{3}\sum_{x\in\Lambda}\sum_{\mu<\nu}\mathrm{Re}\,\mathrm{tr}[\mathbb{1} - U_{\mu\nu}(x)] \\ &= \frac{\beta}{3}\sum_{x\in\Lambda}\sum_{\mu<\nu}\mathrm{Re}\,\mathrm{tr}\Big[\cancel{-iga^2\mathcal{F}^a_{\mu\nu}(x)T^a} - \cancel{a^3E^{(3)}_{\mu\nu}(x)} - \cancel{a^4E^{(4)}_{\mu\nu}(x)} \\ &\qquad\qquad + \frac{1}{2}g^2a^4\mathcal{F}^a_{\mu\nu}(x)\mathcal{F}^b_{\mu\nu}(x)T^aT^b + \mathcal{O}(a^5)\Big] \\ &= \frac{a^4}{2}\sum_{x\in\Lambda}\sum_{\mu<\nu}\mathcal{F}^a_{\mu\nu}(x)\mathcal{F}^a_{\mu\nu}(x) + \mathcal{O}(a^5) = \frac{a^4}{4}\sum_{x\in\Lambda}\sum_{\mu,\nu}F^a_{\mu\nu}(x)F^a_{\mu\nu}(x) + \mathcal{O}(a^5), \end{aligned} \tag{3.10}$$

where in the second line, the first term cancels because $\mathrm{tr}(T^a) = 0$ and the two following ones according to eq. (3.9). In the last step, we have made use of the fact that the lattice field strength tensor is equal to the continuum one up to corrections of $\mathcal{O}(a^2)$.[2] Equation (3.10) clearly suggests that

$$\lim_{a\to 0} S_{G,W}[U] = \frac{1}{4}\int d^4x\, F^a_{\mu\nu}(x)F^a_{\mu\nu}(x) = S_G[A]. \tag{3.11}$$

This is the naive or classical continuum limit in contradistinction to the continuum limit of the quantized theory [73, 83].

[2]Note that such corrections to the first term in the second line of eq. (3.10) are proportional to T^a and thus vanish upon taking the trace.

Lüscher-Weisz gauge action

Performing the full continuum limit on the other hand is a very non-trivial task involving a series of simulations with increasingly smaller lattice spacings and an increasingly larger number of lattice points, as was mentioned at the beginning of this chapter. In order to facilitate this undertaking, one can try to find discretized actions which reduce the discretization error and thus speed up the approach to the continuum limit. A systematic approach to this is offered by the Symanzik improvement program [86, 87].

Starting from a simple, discretized version of the action, such as the Wilson gauge action in eq. (3.5), one collects a set of irrelevant terms which have the same symmetries as the original action (*i.e.*, invariance under gauge transformations, parity, and $\pi/2$ rotations [88]), ordered by their classical (inverse mass) dimension [89].[3] The plaquette term, from which the Wilson gauge action is built, is of dimension 4, as has been worked out in eq. (3.10). Since terms with an odd dimension are not allowed by the above symmetry constraints [89], the lowest-dimensional terms of interest are of dimension 6, *i.e.*, they contribute at $\mathcal{O}(a^2)$. As an incidental observation, this implies that discretization errors to the Wilson gauge action are only of $\mathcal{O}(a^2)$ [83]. One finds that there are three independent dimension-6 terms which satisfy the above symmetry constraints [88], so that the improved action can be written in the form [89]

$$S_{G,LW}[U] = \frac{\beta}{3}\sum_{i=0}^{3} c_i(g^2) \sum_{C\in\mathcal{S}_i} \mathrm{Re}\,\mathrm{tr}(\mathbb{1} - U_C). \tag{3.12}$$

Here, the $\mathcal{S}_i$s denote sets of closed loops of gauge links C with different geometries (*cf.*, fig. 3.2), and U_C the ordered product of the gauge links along C. Analogously to what has been demonstrated for the Wilson gauge action above, one can work out the classical small-a expansion of the additional terms with $i = 1, 2, 3$ in eq. (3.12). It turns out that the plaquette term needs to be subtracted with appropriate integer prefactors (the ones also appearing in eq. (3.13) below) from the remaining three terms in order to obtain pure dimension-6 terms. The resulting terms span a basis of all dimension-6 terms (up to total derivatives) fulfilling the above symmetry constraints [89].

In order to determine the coefficients $c_i(g^2)$ in eq. (3.12), one has to specify an improvement condition. An improvement condition is a statement about which quantities are to be improved. In the following, we will adopt the on-shell improvement condition from Ref. [89], which demands that the scaling violations near the continuum limit are of $\mathcal{O}(a^4)$ for all spectral quantities, *i.e.*, low-lying energy values of the system. Fixing (without loss of generality) the normalization convention

$$c_0(g^2) + 8c_1(g^2) + 8c_2(g^2) + 16c_3(g^2) = 1, \tag{3.13}$$

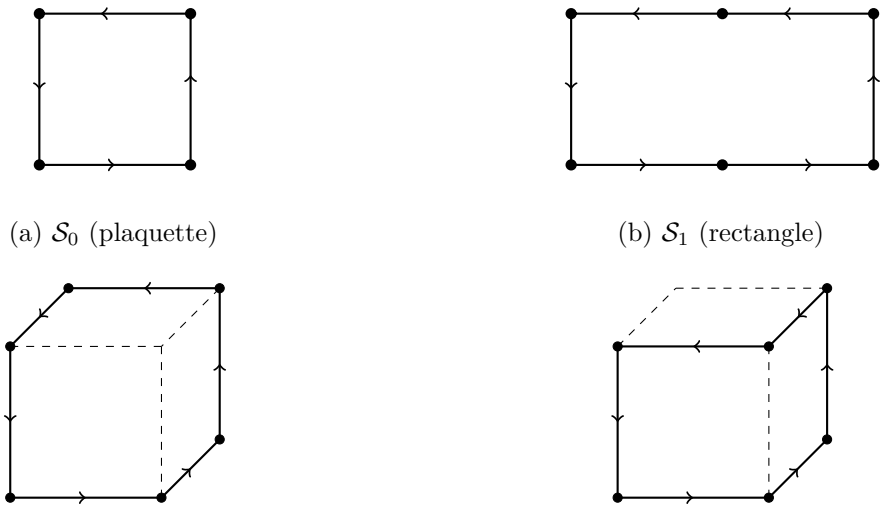
it has been found in Ref. [89] that to tree level in perturbation theory in g^2, all $\mathcal{O}(a^2)$-artefacts are removed from spectral quantities if the following two conditions are fulfilled,

$$c_1(0) - c_2(0) - c_3(0) = -\frac{1}{12}, \quad c_2(0) = 0. \tag{3.14}$$

This means that one coefficient, *e.g.*, $c_3(0)$, remains a free parameter which is only restricted by the requirement that the resulting action is positive. One allowed possibility which has the additional advantage that not only spectral quantities, but also many others, get improved at tree level is $c_3(0) = 0$ [89].[4] This choice obviously implies $c_0(0) = 5/3$ and $c_1(0) = -1/12$; this is also the convention used for the $(2+1)$-flavor CLS simulations [90] (*cf.*, chapter 4).

[3] The classical dimension of a local lattice operator is defined as the smallest power of a appearing in its classical small-a expansion [89].

[4] One can even choose $c_3(g^2) = 0$ to all orders in perturbation theory [89].



(a) $\mathcal{S}_0$ (plaquette)

(b) $\mathcal{S}_1$ (rectangle)

(c) $\mathcal{S}_2$ (non-planar parallelogram)

(d) $\mathcal{S}_3$ (non-planar bent rectangle)

Figure 3.2: Sets $\mathcal{S}_i$ of elementary loops C occurring in the improved action eq. (3.12) [89]

3.1.2 Fermions on the lattice

Naive fermion action and doubling of fermion species

In order to construct a lattice action for the fermions, one needs to find a suitably discretized version of the Euclidean continuum fermion action defined by the first term in eq. (2.54). The space-time integral is replaced by a sum over the lattice, $\int d^4x \to a^4 \sum_{x\in\Lambda}$, analogously to the gauge action. For the partial derivative, we start with the symmetric finite-difference formula,

$$\partial_\mu \psi(x) = \frac{1}{2a}(\psi(x + a\hat{\mu}) - \psi(x - a\hat{\mu})) + \mathcal{O}(a^2). \tag{3.15}$$

To make this expression gauge invariant, we need to ensure that the terms involving fields located at $x \pm a\hat{\mu}$ transform in the same way under gauge transformations as $\psi(x)$. Analogously to the continuum case discussed in section 2.2, this proceeds by multiplying the respective fields with the appropriate gauge transporters,

$$S_{F,\text{naive}}[\bar{\psi}, \psi, U] = \sum_{f=1}^{N_f} a^4 \sum_{x\in\Lambda} \bar{\psi}^f(x) \left[\sum_\mu \gamma_\mu \frac{1}{2a}(U_\mu(x)\psi^f(x + a\hat{\mu}) - U_{-\mu}(x)\psi^f(x - a\hat{\mu})) + m^f \psi^f(x) \right]. \tag{3.16}$$

This is the so-called naive fermion action [83]. By multiplying the gauge links with the fermion fields, we have already introduced an interaction between them. Thus, on the lattice the gauge link plays a similar role as the covariant derivative does in the continuum. This connection can be made more explicit by expanding eqs. (3.2) and (3.3) for small a [83],

$$U_\mu(x) = \mathbb{1} - iga\mathcal{A}^a_\mu(x)T^a + \mathcal{O}(a^2) \qquad = \mathbb{1} - igaA^a_\mu(x)T^a + \mathcal{O}(a^2), \tag{3.17}$$

$$U_{-\mu}(x) = \mathbb{1} + iga\mathcal{A}^a_\mu(x - a\hat{\mu})T^a + \mathcal{O}(a^2) = \mathbb{1} + igaA^a_\mu(x - a\hat{\mu})T^a + \mathcal{O}(a^2). \tag{3.18}$$

Inserting these relations into eq. (3.16), we find

$$\begin{aligned} S_{F,\text{naive}}[\bar{\psi},\psi,U] &= \sum_{f=1}^{N_f} a^4 \sum_{x\in\Lambda} \bar{\psi}^f(x) \Bigg\{ \sum_\mu \gamma_\mu \Bigg[\frac{1}{2a}(\psi^f(x+a\hat{\mu}) - \psi^f(x-a\hat{\mu})) \\ &\qquad - \frac{ig}{2}(A^a_\mu(x) T^a \psi^f(x+a\hat{\mu}) + A^a_\mu(x-a\hat{\mu}) T^a \psi^f(x-a\hat{\mu})) \Bigg] \\ &\qquad + \mathcal{O}(a) + m^f \psi^f(x) \Bigg\} \\ &= \sum_{f=1}^{N_f} a^4 \sum_{x\in\Lambda} \bar{\psi}^f(x) \left[\sum_\mu \gamma_\mu (\partial_\mu \psi^f(x) - igA^a_\mu(x) T^a \psi^f(x)) + m^f \psi^f(x) + \mathcal{O}(a) \right], \end{aligned} \tag{3.19}$$

where in the last step we have used that $\psi^f(x \pm a\hat{\mu}) = \psi^f(x) + \mathcal{O}(a)$ and $A^a_\mu(x - a\hat{\mu}) = A^a_\mu(x) + \mathcal{O}(a)$. We have thus established that

$$\lim_{a\to 0} S_{F,\text{naive}}[\bar{\psi},\psi,U] = \sum_{f=1}^{N_f} \int d^4x\, \bar{\psi}^f(x)(\gamma_\mu D_\mu + m^f)\psi^f(x) = S_F[\bar{\psi},\psi,A], \tag{3.20}$$

i.e., that the naive fermion action has the correct classical continuum limit.

Nevertheless, the question remains whether the naive fermion action also describes the correct quantum physics on the lattice. In order to shed more light onto the situation, we derive the fermion propagator and analyze its pole structure which determines the particle content of the theory. To that end, we first need to rewrite the naive fermion action in a quadratic form [83],

$$S_{F,\text{naive}}[\bar{\psi},\psi,U] = \sum_{f=1}^{N_f} a^4 \sum_{x,y\in\Lambda} \bar{\psi}^f(x) Q^f_{\text{naive}}[U](x,y)\psi^f(y), \tag{3.21}$$

where the naive Dirac operator is given by [83]

$$Q^f_{\text{naive}}[U](x,y) = \sum_\mu \gamma_\mu \frac{1}{2a}(U_\mu(x)\delta_{x+a\hat{\mu},y} - U_{-\mu}(x)\delta_{x-a\hat{\mu},y}) + m^f \delta_{x,y}. \tag{3.22}$$

In the following, we focus on the case of free fermions, *i.e.*, trivial gauge fields $U_\mu(x) = \mathbb{1}$. The subsequent analysis is most easily performed in momentum space, so that we apply a Fourier transformation to the Dirac operator [83],

$$\begin{aligned} \tilde{Q}^f_{\text{naive,free}}(p,q) &= \frac{1}{N_s^3 N_\tau} \sum_{x,y\in\Lambda} e^{-ip\cdot x} Q^f_{\text{naive,free}}(x,y) e^{iq\cdot y} \\ &= \frac{1}{N_s^3 N_\tau} \sum_{x\in\Lambda} e^{-i(p-q)\cdot x} \left[\sum_\mu \gamma_\mu \frac{1}{2a}(e^{iq_\mu a} - e^{-iq_\mu a}) + m^f \right] \\ &= \delta_{p,q} \tilde{Q}^f_{\text{naive,free}}(p). \end{aligned} \tag{3.23}$$

The Fourier transform of the naive free lattice Dirac operator is hence defined by [83]

$$\tilde{Q}^f_{\text{naive,free}}(p) = m^f + \frac{i}{a} \sum_\mu \gamma_\mu \sin(p_\mu a), \tag{3.24}$$

and the corresponding propagator is given by its inverse [83],

$$\left(\tilde{Q}^f_{\text{naive,free}}(p)\right)^{-1} = \frac{m^f - ia^{-1}\sum_\mu \gamma_\mu \sin(p_\mu a)}{(m^f)^2 + a^{-2}\sum_\mu \sin^2(p_\mu a)}. \tag{3.25}$$

This expression can easily be verified by multiplying both sides with $\tilde{Q}^f_{\text{naive,free}}(p)$ as given in eq. (3.24) and considering that $\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}$.

In the preceding relations, the allowed momenta p_μ are elements of the dual lattice [83]

$$\tilde{\Lambda} = \left\{ p = (p_0, p_1, p_2, p_3) \;\middle|\; p_\mu = \frac{2\pi}{aN_\mu}(k_\mu + \theta_\mu),\, k_\mu = -\frac{N_\mu}{2}+1, -\frac{N_\mu}{2}+2, \ldots, \frac{N_\mu}{2} \right\}, \tag{3.26}$$

where $N_0 = N_\tau$ and $N_j = N_s$ for $j = 1, 2, 3$. The boundary phase θ_μ is defined by the boundary conditions in direction μ [83],

$$f(x + aN_\mu \hat{\mu}) = e^{i2\pi\theta_\mu} f(x), \tag{3.27}$$

so that $\theta_\mu = 0$ for periodic and $\theta_\mu = 1/2$ for antiperiodic boundary conditions. As mentioned at the beginning of this chapter, for fermions one has $\theta_0 = 1/2$ and $\theta_j = 0$. The boundary phase θ_μ needs to be included in the definition of the lattice momenta in eq. (3.26) so that the plane waves given by $\exp(ip \cdot x)$ also satisfy the boundary conditions eq. (3.27) [83].

Since the denominator of eq. (3.25) is periodic in $p_\mu a$ with period π, the naive free fermion propagator has $2^4 = 16$ poles within the Brillouin zone defined by eq. (3.26), one for each equivalent corner of the Brillouin zone. Because physical states correspond to poles of the propagator, this implies that the naive fermion action describes the propagation of 16 fermions. Compared to the situation in the continuum, where there is only one pole, we have thus identified 15 additional, unphysical poles, the so-called doublers. In an interacting theory, even if the external particles in some process correspond to the desired state at the 0-corner of the Brillouin zone, the doublers can appear in virtual loops and thus influence the physics [80]. This clearly means that the doublers need to be removed in order to obtain a reasonable lattice theory of fermions.

Wilson fermion action

A possible solution to the aforementioned doubling problem has been suggested by Wilson [91]. It consists in adding an extra term to the Dirac operator, so that, in the free case, the momentum-space Dirac operator reads [83]

$$\tilde{Q}^f_{W,\text{free}}(p) = m^f + \frac{i}{a}\sum_\mu \gamma_\mu \sin(p_\mu a) + \frac{1}{a}\sum_\mu (1 - \cos(p_\mu a)). \tag{3.28}$$

The extra term, the so-called Wilson term, acts like an additional mass term for the doublers. For lattice momentum $k = a(p + q_\pi)$ with

$$q_\pi \in \{\, (0,0,0,0), (\pi/a, 0, 0, 0), (0, \pi/a, 0, 0), \ldots, (\pi/a, \pi/a, \pi/a, \pi/a) \,\}, \tag{3.29}$$

the total mass is given by $m + 2n_\pi/a$, where n_π is the number of components of q_π equal to π/a. This shows that in the continuum limit $a \to 0$, the doublers, *i.e.*, the states with $n_\pi \neq 0$, become infinitely heavy and decouple from the theory [80, 83].

The Wilson Dirac operator in position space can be obtained from eq. (3.28) by Fourier transformation,

$$\begin{aligned}
Q^f_{W,\text{free}}(x,y) &= \frac{1}{N_s^3 N_\tau} \sum_{p\in\tilde{\Lambda}} \tilde{Q}^f_{W,\text{free}}(p) e^{ip\cdot(x-y)} \\
&= m^f \delta_{x,y} + \frac{1}{aN_s^3 N_\tau} \sum_{p\in\tilde{\Lambda}} \sum_\mu (1 - \cos(p_\mu a) + i\gamma_\mu \sin(p_\mu a)) e^{ip\cdot(x-y)} \\
&= m^f \delta_{x,y} + \frac{1}{aN_s^3 N_\tau} \sum_{p\in\tilde{\Lambda}} \sum_\mu \left[1 - \frac{1}{2}(e^{ip_\mu a} + e^{-ip_\mu a}) + \frac{1}{2}\gamma_\mu (e^{ip_\mu a} - e^{-ip_\mu a})\right] e^{ip\cdot(x-y)} \\
&= m^f \delta_{x,y} + \frac{1}{aN_s^3 N_\tau} \sum_{p\in\tilde{\Lambda}} \sum_\mu \Bigg(e^{ip\cdot(x-y)} - \frac{1}{2} e^{ip\cdot(x-y+a\hat{\mu})} - \frac{1}{2} e^{ip\cdot(x-y-a\hat{\mu})} \\
&\qquad\qquad + \frac{1}{2}\gamma_\mu e^{ip\cdot(x-y+a\hat{\mu})} - \frac{1}{2}\gamma_\mu e^{ip\cdot(x-y-a\hat{\mu})} \Bigg) \\
&= \left(m^f + \frac{4}{a}\right) \delta_{x,y} - \frac{1}{2a} \sum_\mu [(\mathbb{1} - \gamma_\mu)\delta_{x+a\hat{\mu},y} + (\mathbb{1} + \gamma_\mu)\delta_{x-a\hat{\mu},y}]. \qquad (3.30)
\end{aligned}$$

This is the expression for the free theory. In order to obtain the corresponding Dirac operator for the interacting theory, one has to make eq. (3.30) gauge invariant by multiplying with the appropriate gauge links analogously to eq. (3.22) [83],

$$Q^f_W[U](x,y) = \left(m^f + \frac{4}{a}\right) \delta_{x,y} - \frac{1}{2a} \sum_\mu [(\mathbb{1} - \gamma_\mu) U_\mu(x) \delta_{x+a\hat{\mu},y} + (\mathbb{1} + \gamma_\mu) U_{-\mu}(x) \delta_{x-a\hat{\mu},y}]. \quad (3.31)$$

By comparing eqs. (3.22) and (3.31), one finds that the Wilson term in position space is given by [83]

$$\begin{aligned}
Q^f_W[U](x,y) - Q^f_{\text{naive}}[U](x,y) &= \frac{4}{a}\delta_{x,y} - \frac{1}{2a} \sum_\mu (U_\mu(x)\delta_{x+a\hat{\mu},y} + U_{-\mu}(x)\delta_{x-a\hat{\mu},y}) \\
&= -\frac{a}{2} \sum_\mu \frac{1}{a^2} (U_\mu(x)\delta_{x+a\hat{\mu},y} - 2\delta_{x,y} + U_{-\mu}(x)\delta_{x-a\hat{\mu},y}), \qquad (3.32)
\end{aligned}$$

i.e., by the gauge-invariant discretization of $-a/2\partial_\mu\partial_\mu$. The prefactor a means that this term vanishes in the classical continuum limit; its role is to suppress the doublers appearing in the quantum theory, as discussed above.

One particular downside of the introduction of the Wilson term is that it breaks chiral symmetry explicitly at any finite lattice spacing, even in the limit of massless quarks. This is due to the $4/a\delta_{x,y}$-term in eq. (3.31), which introduces a "mass-like" $\bar{\psi}(x)\psi(x)$ term into the action, which obviously breaks chiral symmetry, as discussed in section 2.2. More generally, the no-go theorem of Nielsen and Ninomiya [92–94] states that it is impossible to solve the fermion doubling problem in a way that preserves chiral symmetry in the strict continuum sense.

$\mathcal{O}(a)$-improved Wilson fermion action

As has been worked out with the classical small-a expansion of the naive fermion action in eq. (3.19), lattice artefacts to the fermion action start at $\mathcal{O}(a)$. The Wilson term in eq. (3.32) adds another, albeit necessary, $\mathcal{O}(a)$ contribution. In order to improve the approach to the continuum limit, one can implement the Symanzik improvement program [86, 87] which has already been lined out above for the gauge action. To recapitulate, we need to find all terms of

the next-higher dimension, *i.e.*, dimension 5, which have the same symmetries as the Wilson fermion action.

It has been shown in Refs. [95, 96] that, in the continuum, the most general dimension-5 contribution to the Lagrangian allowed by the pertinent symmetries (*i.e.*, invariance under gauge transformations, parity, $\pi/2$ rotations and charge conjugation) can (up to total derivatives) be written as a linear combination of the following terms,[5]

$$\mathcal{L}_1^{(5)} = \bar{\psi}(x)\sigma_{\mu\nu}F_{\mu\nu}(x)\psi(x), \tag{3.33}$$

$$\mathcal{L}_2^{(5)} = \bar{\psi}(x)\vec{D}_\mu(x)\vec{D}_\mu(x)\psi(x) + \bar{\psi}(x)\overleftarrow{D}_\mu(x)\overleftarrow{D}_\mu(x)\psi(x), \tag{3.34}$$

$$\mathcal{L}_3^{(5)} = m\,\mathrm{tr}[F_{\mu\nu}(x)F_{\mu\nu}(x)], \tag{3.35}$$

$$\mathcal{L}_4^{(5)} = m[\bar{\psi}(x)\gamma_\mu\vec{D}_\mu(x)\psi(x) - \bar{\psi}(x)\gamma_\mu\overleftarrow{D}_\mu(x)\psi(x)], \tag{3.36}$$

$$\mathcal{L}_5^{(5)} = m^2\bar{\psi}(x)\psi(x), \tag{3.37}$$

where the left action of the covariant derivative is defined as $\overleftarrow{D}_\mu(x) = \overleftarrow{\partial}_\mu + igA_\mu^a(x)T^a$, and $\sigma_{\mu\nu} = 1/2[\gamma_\mu, \gamma_\nu]$. Barring contact terms, the field equations of the continuum theory can be employed to derive two linear relations between the terms $\mathcal{L}_1^{(5)}$, $\mathcal{L}_2^{(5)}$, $\mathcal{L}_4^{(5)}$ and $\mathcal{L}_5^{(5)}$ [96]. These allow us to eliminate $\mathcal{L}_2^{(5)}$ and $\mathcal{L}_4^{(5)}$ from our list. We also note that $\mathcal{L}_2^{(5)}$ is the Lagrangian corresponding to the Wilson term of eq. (3.32), an interesting connection which has been explored in more detail in Ref. [95]. Furthermore, one notices that a lattice discretization of the term $\mathcal{L}_3^{(5)}$ is already part of the Wilson gauge action, while $\mathcal{L}_5^{(5)}$ is just the mass term appearing in the original fermion action. These two counterterms can thus be accounted for by a renormalization of the bare coupling g and the bare mass m [96]. We will come back to this issue in section 4.3.

The only term one is left with is hence $\mathcal{L}_1^{(5)}$, leading to an $\mathcal{O}(a)$-improved lattice action of the form [95, 96]

$$S_{F,SW}[\bar{\psi},\psi,U] = S_{F,\mathrm{Wilson}}[\bar{\psi},\psi,U] + \sum_{f=1}^{N_f} a^5 c_{SW}(g^2) \sum_{x\in\Lambda}\sum_{\mu<\nu} \bar{\psi}^f(x)\frac{1}{2}\sigma_{\mu\nu}\hat{F}_{\mu\nu}(x)\psi^f(x), \tag{3.38}$$

where $\hat{F}_{\mu\nu}$ is a lattice discretization of the field strength tensor $F_{\mu\nu}$. This is not unique, but conventionally chosen as [96]

$$\hat{F}_{\mu\nu}(x) = -\frac{1}{8a^2}(Q_{\mu\nu}(x) - Q_{\nu\mu}(x)), \tag{3.39}$$

$$Q_{\mu\nu}(x) = U_{\mu,\nu}(x) + U_{\nu,-\mu}(x) + U_{-\mu,-\nu}(x) + U_{-\nu,\mu}(x), \tag{3.40}$$

with the plaquettes defined as in eq. (3.4). The four terms in eq. (3.40) correspond to the four plaquettes shown in fig. 3.3, which give the figure the shape of a four-leaf clover. For this reason, the additional term in eq. (3.38) is known as the Sheikholeslami-Wohlert or Clover term [83]. The coefficient c_{SW} multiplying this term is a function of the bare coupling g. For on-shell improvement, it must be chosen so that the $\mathcal{O}(a)$ lattice artefacts in all on-shell, *i.e.*, spectral, quantities cancel. In Ref. [95] it has been shown by means of a classical small-a expansion that $c_{SW}(0) = 1$ to tree level in perturbation theory, while in Refs. [96, 97] a strategy for the non-perturbative determination has been developed.

[5] For convenience of notation, we have dropped the flavor index f here.

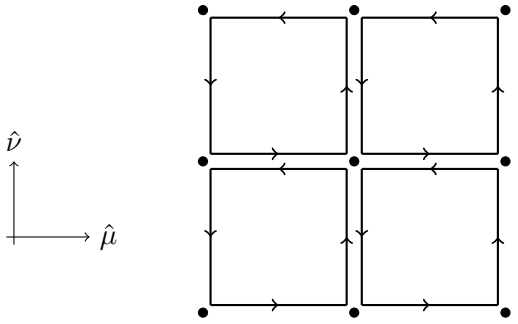


Figure 3.3: Graphical representation of the sum $Q_{\mu\nu}$ of plaquettes contributing to the lattice discretization of the field strength tensor eq. (3.39) [83, 96]

3.2 The path integral in Lattice QCD

With a suitably discretized version of the QCD action at hand, we can start thinking about the evaluation of observables in Lattice QCD. Expectation values are formally given in terms of the path integral [80]

$$\langle O \rangle = \frac{1}{Z_{\mathrm{LQCD}}} \int [d\bar{\psi}][d\psi][dU]\, e^{-S_{\mathrm{LQCD}}[\bar{\psi},\psi,U]} O[\bar{\psi},\psi,U], \tag{3.41}$$

$$Z_{\mathrm{LQCD}} = \int [d\bar{\psi}][d\psi][dU]\, e^{-S_{\mathrm{LQCD}}[\bar{\psi},\psi,U]}, \tag{3.42}$$

where, in our case, $S_{\mathrm{LQCD}} = S_{G,LW} + S_{F,SW}$ with the Lüscher-Weisz gauge action and the Wilson-Clover fermion action according to eqs. (3.12) and (3.38), respectively. The gauge links U are SU(3)-matrices, and their path-integral measure is defined as

$$\int [dU] = \prod_{x\in\Lambda} \prod_{\mu} \int dU_\mu(x), \tag{3.43}$$

where the individual measures for the integration over the link variables $dU_\mu(x)$ are given by the invariant group measure or Haar measure [83].

3.2.1 Pseudofermions

The fermion fields, on the other hand, are formulated in terms of anticommuting Grassmann numbers. A direct implementation is in principle possible, for example by using a so-called polymer representation [80]. Due to the Pauli principle, however, the corresponding contributions have oscillating signs, hindering the numerical evaluation of the path integral in eq. (3.41) by the methods to be developed in the subsequent sections.

Therefore, a more promising approach consists in evaluating the path integral over the fermion fields analytically and only numerically integrating the resulting effective bosonic theory. In (Lattice) QCD, this is possible because the fermion action is quadratic in the fields. For observables depending only on the gauge fields, the path integral for the fermions is thus

Gaussian [80],[6]

$$\begin{aligned}\langle O\rangle &= \frac{1}{Z_{\mathrm{LQCD}}}\int [dU]\,\exp(-S_{G,LW}[U])O[U] \\ &\qquad \prod_{f=1}^{N_f}\int [d\bar{\psi}^f][d\psi^f]\,\exp\left(-\sum_{x,y\in\Lambda}\bar{\psi}^f(x)Q^f_{SW}[U](x,y)\psi^f(y)\right) \\ &= \frac{1}{Z_{\mathrm{LQCD}}}\int [dU]\,\exp(-S_{G,LW}[U])O[U]\prod_{f=1}^{N_f}\det(Q^f_{SW}[U]) \\ &= \frac{1}{Z_{\mathrm{LQCD}}}\int [dU]\,\exp(-S_{\mathrm{eff}}[U])O[U]. \end{aligned} \tag{3.44}$$

Here, we have introduced the effective action $S_{\mathrm{eff}}[U] = S_{G,LW}[U] - \sum_{f=1}^{N_f}\ln\det Q^f_{SW}[U]$. Equation (3.44) shows how expectation values in a joint fermion-boson theory can be evaluated in terms of an effective, purely bosonic theory.

For a meaningful numerical treatment, the effective action needs to be real and positive. This is only guaranteed if the determinant of the Dirac operator is positive. The Wilson-Clover Dirac operator satisfies γ_5-hermiticity [80], so that its determinant is real,

$$\det(Q)^* = \det(Q^\dagger) = \det(\gamma_5 Q\gamma_5) = \det(Q). \tag{3.45}$$

Positivity, however, is not in general fulfilled. One possibility is to consider two exactly degenerate quark flavors. This is for example a good approximation for the two light quark flavors u and d. Then eq. (3.44) contains the product of two identical fermion determinants, which is obviously non-negative.

According to eq. (3.44), everything that needs to be computed for the fermions in order to evaluate expectation values of observables depending only on the gauge fields is, in principle, the determinant of the Dirac operator. On the lattice, this is a matrix with $12N_s^3N_\tau \times 12N_s^3N_\tau$ entries, a huge number even for lattices of a moderate size. Combined with the fact that its determinant is highly non-local [80], this renders an exact evaluation of the latter prohibitively expensive. A possible solution consists in the introduction of an auxiliary complex scalar, *i.e.*, bosonic, field $\phi_{\alpha k}(x)$ for each position x, Dirac index α and color index k [98]. Defining the path-integral measure for these so-called pseudofermions as

$$\int [d\phi] = \prod_{x,\alpha,k}\pi^{-1}\int_{-\infty}^{\infty} d\,\mathrm{Re}\,\phi_{\alpha k}(x)\int_{-\infty}^{\infty} d\,\mathrm{Im}\,\phi_{\alpha k}(x), \tag{3.46}$$

one arrives at the following expression for the product of the fermion determinants of two degenerate flavors [98],

$$\det(QQ^\dagger) = \int [d\phi]\,\exp[-\phi^\dagger(QQ^\dagger)^{-1}\phi]. \tag{3.47}$$

To summarize, for observables depending only on the gauge fields, we can replace the path integral over the fermions by an integral over effective bosonic degrees of freedom [98],

$$\langle O\rangle = \frac{1}{Z_{\mathrm{LQCD}}}\int [d\bar{\psi}][d\psi][dU]\,e^{-S_G[U]-S_F[\bar{\psi},\psi,U]}O[U] = \frac{1}{Z_{\mathrm{LQCD}}}\int [dU][d\phi]\,e^{-S_G[U]-S_\phi[\phi,U]}O[U], \tag{3.48}$$

[6] Note that we have rescaled the fermion fields by a factor of a^2 in order to remove the a^4 prefactor in the exponent originating from the sum over the lattice. This does not affect expectation values if it is done consistently in both the numerator and the denominator (the partition function).

where $S_\phi = \phi^\dagger (QQ^\dagger)^{-1}\phi$. Since the pseudofermion fields appear by construction only in S_ϕ, our strategy will be to evaluate the path integral over the pseudofermions (*i.e.*, the fermion determinant) for fixed gauge fields, and then to integrate over the gauge fields using the appropriate effective action $S_{\text{eff}}[U] = S_G[U] - \ln \int [d\phi] \exp(-S_\phi[\phi, U])$.

For observables depending also on the fermion fields, one can apply Wick's theorem (*cf.*, eq. (2.39)) to express their expectation values in terms of propagators, *i.e.*, inverses of the Dirac operator [80]. Therefore, it is also in this more general case possible to arrive at an expression involving a path integral only over the gauge links (*cf.*, eq. (5.31)). Methods for the efficient estimation of the inverse of the Dirac operator will be covered later in the context of calculations of nucleon matrix elements (*cf.*, section 5.2).

3.2.2 The Monte-Carlo method

The path integral in eq. (3.44) is, on the lattice, a finite-dimensional integral, making it in principle amenable to a numerical evaluation. However, the number of dimensions is extremely large, even for moderately large lattices. The only way to evaluate such an expression is by Monte-Carlo integration. Here, one replaces the integral by an average of the observable evaluated on N sample configurations of the gauge fields [83],

$$\langle O \rangle \approx \hat{O} = \frac{1}{N} \sum_{n=1}^{N} O[U_n], \tag{3.49}$$

which is an estimator for the true mean value. Already for more than three dimensions, this method becomes more efficient than the usual numerical quadrature algorithms [83]. Moreover, it can be shown that the error to eq. (3.49) behaves like $\mathcal{O}(N^{-1/2})$ [83], so that for *e.g.*, halving the error, the statistics needs to be quadrupled. This statement is of course only strictly true if the samples $[U_n]$ are uncorrelated, which they are in practice not, as we shall see in the next subsection.

The simplest approach to evaluating eq. (3.49) would be to generate N configurations $[U_n]$ randomly in the space of gauge field variables. This would, however, be very inefficient because the path integral in eq. (3.44) is, for sufficiently large lattices, sharply peaked at some specific configurations, namely those which minimize the action [80]. A more efficient Monte-Carlo integration thus samples configurations according to their weight in the integrand $\exp(-S_{\text{eff}}[U])$. In order to apply this so-called importance sampling, it is useful to introduce the concept of an ensemble in analogy to statistical mechanics: an ensemble consists of a (in theory infinite, in practice finite) number of field configurations at the same macroscopic physics parameters, following a density $P[U] \propto \exp(-S_{\text{eff}}[U])$. Note that the interpretation of $\exp(-S_{\text{eff}}[U])$ as a probability density is only possible if the effective action is real and positive, as already mentioned in the previous subsection.

Using importance sampling ensures that no computing time is wasted on configurations which have only a negligible contribution, and that a reasonably small number of samples is sufficient for a reliable estimation of the path integral. Far away from phase transitions, usually several hundred to a few thousand configurations are employed (*cf.*, section 4.5).

3.2.3 Markov chains

The remaining problem consists in finding N configurations $[U_n]$ which follow the probability distribution

$$dP[U] = \frac{1}{Z_{\text{LQCD}}} e^{-S_{\text{eff}}[U]} [dU]. \tag{3.50}$$

The idea is to start from some arbitrary configuration and then to construct a stochastic sequence of configurations which eventually follows the desired equilibrium distribution [83]. For this purpose, an updating algorithm is needed which generates from one configuration $[U]$ the next one $[U']$ according to a transition probability $T([U'] \leftarrow [U])$. The update step thus changes the probability distribution from $W[U]$ to

$$W'[U'] = \sum_{[U]} T([U'] \leftarrow [U]) W[U]. \tag{3.51}$$

The transition probability needs to fulfill the normalization condition

$$\sum_{[U']} T([U'] \leftarrow [U]) = 1 \tag{3.52}$$

as well as $T([U'] \leftarrow [U]) > 0 \ \forall [U], [U']$. The latter is called strong ergodicity because it means that every configuration can be reached with a finite probability from any other one. A stochastic process satisfying these two conditions is called a Markov process, and the corresponding sequence of configurations a Markov chain [80].

One more important property that our Markov chains will need to fulfill is global balance. It means that, once the equilibrium distribution from eq. (3.50) is reached, the probability to step into some configuration $[U']$ must be the same as to step out of it [83],

$$\sum_{[U]} T([U'] \leftarrow [U]) P[U] = \sum_{[U]} T([U] \leftarrow [U']) P[U']. \tag{3.53}$$

Using the normalization condition eq. (3.52), the right-hand side of eq. (3.53) simply becomes $P[U']$, so that global balance is equivalent to demanding that the equilibrium distribution is a fixed point of the Markov process. This also implies that if one starts with an arbitrary probability distribution which has a non-zero overlap with the equilibrium distribution, the Markov process will converge to the equilibrium distribution [80].

Markov chains provide us with a convenient way to sample configurations from in principle arbitrarily complicated probability distributions. The downside of this method is, however, that subsequent configurations are not statistically independent. The updating process introduces a finite autocorrelation among the sequence of configurations it generates, which must be taken into account in the error estimation (*cf.*, section 6.1).

3.2.4 The Metropolis algorithm

Finding a transition probability which both fulfills global balance and leads to an efficient sampling of the configuration space can be a challenging task. One therefore commonly applies the Metropolis algorithm [99]. The general idea is to divide the update step into two parts [100]. In the first part, one chooses a candidate configuration according to some approximate, in principle arbitrary, *a priori* selection probability $T_0([U'] \leftarrow [U])$. Afterwards, in order to ensure global balance, one accepts the candidate configuration $[U']$ as the next configuration in the Markov chain with the acceptance probability

$$T_A([U'] \leftarrow [U]) = \min\left(1, \frac{T_0([U] \leftarrow [U']) P[U']}{T_0([U'] \leftarrow [U]) P[U]}\right). \tag{3.54}$$

The total transition probability is thus given by

$$T([U'] \leftarrow [U]) = T_0([U'] \leftarrow [U]) T_A([U'] \leftarrow [U]). \tag{3.55}$$

If a suggested update is rejected, the unchanged configuration $[U]$ is used again in the Markov chain. Not to do this would unjustifiably reduce the number of configurations in state $[U]$ relative to $[U']$ [99].

In order to prove that this algorithm satisfies the global balance condition eq. (3.53), we compute

$$\begin{aligned} T_0([U'] \leftarrow [U]) P[U] T_A([U'] \leftarrow [U]) &= \min(T_0([U'] \leftarrow [U]) P[U], T_0([U] \leftarrow [U']) P[U']) \\ &= \min(T_0([U] \leftarrow [U']) P[U'], T_0([U'] \leftarrow [U]) P[U]) \\ &= T_0([U] \leftarrow [U']) P[U'] T_A([U] \leftarrow [U']). \end{aligned} \tag{3.56}$$

Consequently, the Metropolis algorithm fulfills not only global balance, but also detailed balance [80],

$$T([U'] \leftarrow [U]) P[U] = T([U] \leftarrow [U']) P[U']. \tag{3.57}$$

The other important property which the transition probability needs to satisfy is ergodicity. In practice, strong ergodicity $T([U'] \leftarrow [U]) > 0 \; \forall [U], [U']$ can be difficult to achieve. In pure gauge theory, for instance, it is common to restrict single updates to local changes of a single gauge link. Since the gauge action is local, this makes computing its change and thus the factor $P[U']/P[U]$ in eq. (3.54) very cheap. Even these local updates are usually further restricted by not considering arbitrary changes, but only a subset of the possible changes which can for example be generated by choosing a random SU(3) group element and multiplying it to the original gauge link. In order to make the algorithm as a whole still ergodic, one applies a large number of such individually non-ergodic steps and views their sequence as one big ergodic step [80].

3.3 The Hybrid Monte Carlo algorithm

In a theory including fermions, computing the change of the (effective) action is very expensive even if the proposed update is local because the fermion determinant is a highly non-local object. One may thus come up with the idea of using global updates in order to save on the number of required steps. Performing a global update by changing all gauge links at once in a naive manner, however, would lead to very small acceptance probabilities because it would be very likely that the new value of the action is much larger than the old one. Therefore, a more sophisticated way to propose candidates for non-local updates needs to be found, which is the topic of the present section.

In section 3.3.1, we will focus our presentation on a theory with two mass-degenerate flavors, so that the fermion determinant is guaranteed to be positive. Afterwards, in section 3.3.2, we will introduce a method to include single flavors in the simulation.

3.3.1 Molecular dynamics

A widely used approach to generating gauge configurations in a theory including fermions is provided by the Hybrid Monte Carlo algorithm [101–103]. The basic idea is to introduce an artificial computer time $\hat{t}$ parametrizing the evolution of the fields along the Markov chain, and a Hamiltonian dynamics describing this evolution. To that end, we write the gauge links as $U_\mu(x;\hat{t}) = \exp[i q_\mu^a(x;\hat{t}) T^a]$ with real coefficients $q_\mu^a(x;\hat{t})$ and introduce conjugate momenta $P_\mu(x;\hat{t}) = P_\mu^a(x;\hat{t}) T^a$. Then we can define the Hamiltonian [83]

$$H[q,P] = \frac{1}{2} \sum_{x,\mu,a} (P_\mu^a(x))^2 + S_G[U[q]] + \phi^\dagger (Q[U[q]] Q^\dagger[U[q]])^{-1} \phi \tag{3.58}$$

with the pseudofermion field ϕ from section 3.2.1. The evolution of q and P in the computer time $\hat{t}$ is thus governed by the following Hamilton equations of motion [83],

$$\frac{d}{d\hat{t}}P^a_\mu(x;\hat{t}) = -\frac{\partial H[q,P]}{\partial q^a_\mu(x;\hat{t})} = -F^a_\mu[U[q(x;\hat{t})],\phi], \tag{3.59}$$

$$\frac{d}{d\hat{t}}q^a_\mu(x;\hat{t}) = \frac{\partial H[q,P]}{\partial P^a_\mu(x;\hat{t})} = P^a_\mu(x;\hat{t}), \tag{3.60}$$

where

$$F^a_\mu[U[q(x;\hat{t})],\phi] = \frac{\partial}{\partial q^a_\mu(x;\hat{t})}(S_G[U[q]] + \phi^\dagger(Q[U[q]]Q^\dagger[U[q]])^{-1}\phi) \tag{3.61}$$

is the force. The Hamiltonian itself is a constant of the motion determined by the eqs. (3.59) and (3.60) which are known as molecular dynamics equations because they describe the time evolution of a classical system of particles.

As suggested in Ref. [104], the pseudofermion fields can be held fixed during the molecular dynamics steps and are updated in between. For this purpose, one generates a complex vector χ of random numbers following a probability distribution $\propto \exp(-\chi^\dagger\chi)$ and sets $\phi = Q\chi$, so that the probability distribution for ϕ is proportional to $\exp[-\phi^\dagger(QQ^\dagger)^{-1}\phi]$. Each molecular dynamics step then commences by randomly generating some initial momenta $P^a_\mu(x;\hat{t}_0)$ according to a Gaussian distribution $\propto \exp[-1/2\sum_{x,\mu,a}(P^a_\mu(x))^2]$ [83]. Afterwards, one integrates the eqs. (3.59) and (3.60) numerically over a distance $\Delta\hat{t}$, using the variables $q^a_\mu(x;\hat{t}_0)$ corresponding to the gauge configuration $U_\mu(x;\hat{t}_0)$ and the previously generated $P^a_\mu(x;\hat{t}_0)$ as initial values.

This integration introduces numerical errors which lead to a violation of the exact conservation of the Hamiltonian. These must be corrected for by following the procedure described above with a Metropolis step with the acceptance probability (*cf.*, eq. (3.54)) [103],

$$T_A([q',P'] \leftarrow [q,P]) = \min(1,\exp(-\Delta H)), \quad \Delta H = H[q',P'] - H[q,P]. \tag{3.62}$$

The total transition probability for the q-field (and thus the gauge field U) alone is given by [103]

$$T([q'] \leftarrow [q]) = \int [dP][dP'] \exp\left[-\frac{1}{2}\sum_{x,\mu,a}(P^a_\mu(x))^2\right] T_{\mathrm{md}}([q',P'] \leftarrow [q,P])T_A([q',P'] \leftarrow [q,P]), \tag{3.63}$$

where $T_{\mathrm{md}}([q',P'] \leftarrow [q,P])$ is the probability of obtaining the configuration $[q',P'] = [q(\hat{t}_0 + \Delta\hat{t}), P(\hat{t}_0 + \Delta\hat{t})]$ from $[q,P] = [q(\hat{t}_0), P(\hat{t}_0)]$ via the molecular dynamics evolution.[7] In order to prove that the transition probability in eq. (3.63) satisfies the detailed balance condition eq. (3.57), we need to demand that the dynamics is reversible, *i.e.*,

$$T_{\mathrm{md}}([q',P'] \leftarrow [q,P]) = T_{\mathrm{md}}([q,-P] \leftarrow [q',-P']), \tag{3.64}$$

and that the integration measure $[dq][dP]$ is area preserving. Observing that H is invariant under a sign flip of the conjugate momenta, the proof then proceeds by simple arithmetics [103].

Since the usual Runge-Kutta schemes for numerical integration do not fulfill reversibility, they cannot be employed for the integration of the molecular dynamics equations. The simplest alternative is to use the leapfrog algorithm, which ensures reversibility, at least under the

[7] In practice, T_{md} is given by a δ-function because the dynamics is completely deterministic. This is, however, not necessary for the proof of detailed balance as long as eq. (3.64) is fulfilled [103].

assumption of perfect arithmetic precision. To integrate from $\hat{t}_0$ to $\hat{t}_0 + \Delta\hat{t}$, one starts with an initial half-step in the conjugate momenta [83],

$$P_\mu^a\left(x;\hat{t}_0+\frac{\epsilon}{2}\right) = P_\mu^a(x;\hat{t}_0) - \frac{\epsilon}{2}F_\mu^a[U(x;\hat{t}_0),\phi], \tag{3.65}$$

followed by alternating $n = \Delta\hat{t}/\epsilon$ steps in U and $n-1$ steps in P of the form

$$U_\mu(x;\hat{t}_0+k\epsilon) = \exp\left[i\epsilon P_\mu^a\left(x;\hat{t}_0+\left(k-\frac{1}{2}\right)\epsilon\right)T^a\right]U_\mu(x;\hat{t}_0+(k-1)\epsilon), \tag{3.66}$$
$$k = 1,\ldots,n,$$

$$P_\mu^a\left(x;\hat{t}_0+\left(k+\frac{1}{2}\right)\epsilon\right) = P_\mu^a\left(x;\hat{t}_0+\left(k-\frac{1}{2}\right)\epsilon\right) - \epsilon F_\mu^a[U(x;\hat{t}_0+k\epsilon),\phi], \tag{3.67}$$
$$k = 1,\ldots n-1,$$

and a final half-step in the conjugate momenta,

$$P_\mu^a(x;\hat{t}_0+\Delta\hat{t}) = P_\mu^a\left(x;\hat{t}_0+\left(n-\frac{1}{2}\right)\epsilon\right) - \frac{\epsilon}{2}F_\mu^a[U(x;\hat{t}_0+\Delta\hat{t}),\phi]. \tag{3.68}$$

A simple proof that this integration scheme is area preserving and reversible can be found in Ref. [83]. Note that in each step, the force term needs to be evaluated, which involves, according to eq. (3.61), the inversion of the Dirac operator. This is the most expensive part of the algorithm [83].

For the leapfrog algorithm, the discretization error of the half-steps is $\mathcal{O}(\epsilon^2)$, while that of the intermediate, full steps is $\mathcal{O}(\epsilon^3)$ [103]. Due to the Metropolis accept-reject step following the integration, the Hybrid Monte Carlo algorithm is exact, regardless of the step size ϵ. In principle, one can thus use a large step size in order to save computing time. However, with too large step sizes, the acceptance rate in the Metropolis step will become low, making the algorithm inefficient again. Therefore, the value of the step size needs to be carefully optimized, which is a non-trivial task.

3.3.2 The Rational HMC algorithm

Our presentation of the HMC algorithm thus far has been limited to the case of two mass-degenerate flavors in order to ensure positivity of the fermion determinant. If one is dealing with a single flavor f of which it is *a priori* known that $\det(Q^f)$ is positive, a simple rooting trick allows one to circumvent the restriction to an even number of flavors: one approximates $(Q^f)^{-1}$ by an operator $TT^\dagger$, where the matrix inverse square root $T \approx (Q^f)^{-1/2}$ is approximated by a rational function. This is the basic idea of the Rational Hybrid Monte Carlo (RHMC) algorithm [105, 106].

Because of the high accuracy of the rational approximations, it is not necessary to correct for them in the Metropolis step [106]; it is sufficient to apply a reweighting procedure. Since the gauge field configurations have been generated with an action S'_{eff} which is only approximately equal to the true effective action S_{eff}, this difference (which amounts, in this case, to the difference between $\det Q^f$ and $\det(TT^\dagger)^{-1}$) must be accounted for in the measurement of observables [107, 108],

$$\langle O\rangle = \frac{1}{Z_{\text{LQCD}}}\int [dU]\, e^{-S_{\text{eff}}[U]}O[U] = \frac{1}{Z_{\text{LQCD}}}\int [dU]\, e^{-S'_{\text{eff}}[U]}e^{-(S_{\text{eff}}[U]-S'_{\text{eff}}[U])}O[U], \tag{3.69}$$

$$Z_{\text{LQCD}} = \int [dU]\, e^{-S_{\text{eff}}[U]} = \int [dU] e^{-S'_{\text{eff}}[U]}e^{-(S_{\text{eff}}[U]-S'_{\text{eff}}[U])}. \tag{3.70}$$

Introducing the reweighting factor $W_1[U] = \exp[-(S_{\text{eff}}[U] - S'_{\text{eff}}[U])]$ and the expectation value in the theory with the approximate action S'_{eff}, $\langle\cdot\rangle_r = \int [dU] \exp(-S'_{\text{eff}}[U])\cdot$, it is thus sufficient to compute W_1 on all generated gauge configurations and use

$$\langle O \rangle = \frac{\langle O W_1 \rangle_r}{\langle W_1 \rangle_r}. \tag{3.71}$$

It should be noted that the reweighting method generally only works well if there is a significant overlap between the distributions determined by S_{eff} and S'_{eff}, respectively. Otherwise, the HMC simulation using the modified action is unable to efficiently capture the accordingly long tail of the distribution of the reweighting factor. In the case of RHMC, the overlap requirement is fulfilled provided that the rational approximation to the inverse square root is sufficiently accurate.

As mentioned above, the RHMC algorithm may, strictly speaking, only be applied if the fermion determinant of the single flavor is known to be positive. For the Wilson fermion action and its improved modifications it is known that this is not true since these discretizations break chiral symmetry explicitly (*cf.*, section 3.1.2). Applying the RHMC algorithm nevertheless then amounts to replacing $\det(Q^f)$ by $|\det(Q^f)|$. One must therefore include an additional reweighting factor for the sign of the fermion determinant, $W_s = \det(Q^f)/|\det(Q^f)|$ [109]. If the Dirac operator is γ_5-hermitian (which is the case for the Wilson-Clover Dirac operator), its determinant is real (*cf.*, eq. (3.45)), so that all eigenvalues are either real or appear in complex conjugate pairs. The determinant being the product of all eigenvalues, it can hence only be negative if there is an odd number of negative real eigenvalues. This means that $W_s = (-1)^{n_{\text{neg}}}$, where n_{neg} is the number of negative real eigenvalues of the Dirac operator. A method for the practical determination of n_{neg} has been suggested in Ref. [109].

3.3.3 Thermalization

For the HMC algorithm, an initial configuration of gauge fields $[U_0]$ is required from which the evolution along the molecular dynamics trajectory starts. As mentioned in section 3.2.3, this is in principle arbitrary. Possible choices include setting all gauge links to $\mathbb{1}$, picking a random configuration or reusing a configuration generated in a previous simulation for example at different but similar physical parameters [83]. In any case, the initial configuration is generally not particularly likely under the target probability distribution eq. (3.50). Therefore, the system needs to be evolved along the Markov chain by performing a certain number of HMC steps until the equilibrium distribution is reached, which then corresponds to a fixed point of the Markov process. This procedure is known as thermalization. In order to determine when thermalization is complete and the actual measurement of observables in the equilibrium ensemble can begin, one monitors the Monte-Carlo history of certain observables. One waits until these stabilize within fluctuations, *i.e.*, a plateau is reached, and excludes the preceding configurations from the Monte-Carlo estimate in eq. (3.49).

3.4 Operators in Lattice QCD

So far, we have only dealt with the discretization of the QCD action and with algorithms to generate configurations following the distribution determined by it. In order to measure physically interesting observables on these configurations, the relevant operators also need to be discretized. In the following, we discuss three selected topics of particular relevance to the present thesis: the construction of a suitable interpolating (lattice) operator at the examples of the pion and the nucleon, the $\mathcal{O}(a)$-improvement of the vector current and the construction of a vector current which is conserved also on the lattice.

3.4.1 Interpolating operators

For the lattice determination of properties of hadrons, interpolators $O, \bar{O}$ with definite quantum numbers need to be found so that the corresponding Hilbert space operators $\hat{O}, \hat{O}^\dagger$ annihilate and create the correct particle states, respectively. Here, an interpolator is a functional of the lattice fields $U_\mu(x)$, $\psi^f(x)$ and $\bar{\psi}^f(x)$ satisfying the aforementioned requirements [83].

We start the discussion with mesons, as these only have two valence quarks and are hence conceptually simpler than baryons with three valence quarks. The pions π^+, π^0, π^- form an isospin triplet with $I = 1$, have zero spin $J = 0$ and negative parity $P = -1$. The valence quark content of the π^+ is thus $\bar{d}u$, while that of the π^- is $\bar{u}d$. Assuming exact isospin symmetry, as it is done in the present thesis, these two states are exactly mass-degenerate. A possible interpolating operator for the π^+ with the correct spin ($J = 0$) and isospin ($I = 1$, $I_z = +1$) is [83]

$$\Pi^+(x) = \bar{\psi}^d(x)\gamma_5\psi^u(x). \tag{3.72}$$

As a typical fermion bilinear, this is obviously gauge invariant and should transform as a pseudoscalar under Lorentz transformations. To make the latter more explicit, we define the action of the parity operator $\mathcal{P}$ on the lattice fermion fields as [83]

$$\mathcal{P}\psi^f(x_0, \mathbf{x}) = \gamma_0\psi^f(x_0, -\mathbf{x}), \quad \mathcal{P}\bar{\psi}^f(x_0, \mathbf{x}) = \bar{\psi}^f(x_0, -\mathbf{x})\gamma_0. \tag{3.73}$$

We can hence calculate

$$\mathcal{P}\Pi^+(x_0, \mathbf{x}) = \bar{\psi}^d(x_0, -\mathbf{x})\gamma_0\gamma_5\gamma_0\psi^u(x_0, -\mathbf{x}) = -\bar{\psi}^d(x_0, -\mathbf{x})\gamma_5\psi^u(x_0, -\mathbf{x}) = -\Pi^+(x_0, -\mathbf{x}). \tag{3.74}$$

After momentum projection, the pion interpolator Π^+ consequently has parity $P = -1$, as required.

The interpolator corresponding to the creation operator of a π^+ can be found by conjugating eq. (3.72) [83],

$$\bar{\Pi}^+(x) = (\Pi^+(x))^\dagger = -(\psi^u(x))^\dagger\gamma_5^\dagger(\bar{\psi}^d(x))^\dagger = -\bar{\psi}^u(x)\gamma_0\gamma_5^\dagger\gamma_0\psi^d(x) = \bar{\psi}^u(x)\gamma_5\psi^d(x) = \Pi^-(x). \tag{3.75}$$

Here, the minus sign in the first step arises from the interchange of the two Grassmann variables. Since the π^- is the antiparticle of the π^+, the creation interpolator of the π^+ is the same as the annihilation interpolator of the π^-.

As the aim of this thesis is to compute structural properties of the nucleon, we also need to construct a nucleon interpolator. The proton and the neutron are the $I_z = +1/2$ and $I_z = -1/2$ components of an isospin doublet with total isospin $I = 1/2$. While the proton has valence-quark content uud, the neutron is a ddu-like state. Under the assumption of exact isospin symmetry, the proton and neutron are exactly mass-degenerate, and are thus collectively referred to as nucleons. The simplest uud-type interpolator for a nucleon state is given by [83]

$$N(x) = \epsilon_{jkl}\psi_j^u(x)[(\psi_k^u(x))^T C\gamma_5\psi_l^d(x)], \tag{3.76}$$

where j, k, l are color indices in the fundamental representation, while the Dirac indices have been suppressed in favor of a matrix/vector notation, as usual. C is the charge conjugation matrix satisfying $C\gamma_\mu C^{-1} = -\gamma_\mu^T$. Contracting the color indices with the ϵ-tensor makes the interpolator a color singlet and gauge invariant [83]. Moreover, it is obvious that eq. (3.76) has isospin $I = 1/2$, $I_z = +1/2$ and spin $J = 1/2$, as needed. To determine the transformation of the nucleon interpolator under parity, we compute with the help of eq. (3.73),

$$\begin{aligned}\mathcal{P}N(x_0, \mathbf{x}) &= \epsilon_{jkl}\gamma_0\psi_j^u(x_0, -\mathbf{x})[(\psi_k^u(x_0, -\mathbf{x}))^T\gamma_0^T C\gamma_5\gamma_0\psi_l^d(x_0, -\mathbf{x})] \\ &= \epsilon_{jkl}\gamma_0\psi_j^u(x_0, -\mathbf{x})[(\psi_k^u(x_0, -\mathbf{x}))^T C\gamma_5\psi_l^d(x_0, -\mathbf{x})] = \gamma_0 N(x_0, -\mathbf{x}),\end{aligned} \tag{3.77}$$

where from the first to the second line, we have used that $\gamma_0^T C\gamma_5\gamma_0 = -\gamma_0^T C\gamma_0\gamma_5 = C\gamma_0\gamma_0\gamma_5 = C\gamma_5$. Consequently, the parity-projected interpolator [83]

$$N_\pm(x) = \Gamma_\pm^p N(x) = \frac{1}{2}(1 \pm \gamma_0)N(x) \tag{3.78}$$

has, after momentum projection, definite parity $P = \pm 1$. While N_+ describes the nucleon, which has positive parity, N_- couples to its negative-parity partner $N(1535)$ [83]. The interpolator for the corresponding creation operators can be constructed analogously to the case of the pion (*cf.*, eq. (3.75)). One finds

$$\bar{N}(x) = N^\dagger(x)\gamma_0 = \epsilon_{jkl}[\bar{\psi}_j^d(x)\gamma_5 C^\dagger(\bar{\psi}_k^u(x))^T]\bar{\psi}_l^u(x), \tag{3.79}$$

and thus for the corresponding parity-projected interpolators,[8]

$$\bar{N}_\pm(x) = \epsilon_{jkl}[\bar{\psi}_j^d(x)\gamma_5 C^\dagger(\bar{\psi}_k^u(x))^T]\bar{\psi}_l^u(x)\Gamma_\pm^p. \tag{3.80}$$

As a general remark, we note that any interpolator will always have a finite overlap with all states with the same quantum numbers. This is because the interpolators generally do not create eigenstates of the Hamiltonian [83]. In practice, this means that it is impossible to create, say, just a ground-state nucleon on the lattice. There will always be some contamination by excited states, which needs to be appropriately removed in the data analysis.

3.4.2 $\mathcal{O}(a)$-improvement of the vector current

In section 3.1, we have discussed the $\mathcal{O}(a)$-improvement of the action. With the chosen improvement condition, this is sufficient for the improvement of on-shell spectral quantities like hadron masses. As we shall see in chapter 5, in order to study the electromagnetic form factors of nucleons, it is necessary to additionally put the vector current

$$V_\mu^{ff'}(x) = \bar{\psi}^f(x)\gamma_\mu\psi^{f'}(x) \tag{3.81}$$

on the lattice. Correlation functions involving currents like this are not automatically improved. By contrast, one needs to explicitly improve the current by adding an appropriate linear combination of all dimension-4 operators with the same symmetries as the original current which has dimension 3. For the vector current in eq. (3.81), there is only one such dimension-4 term. In the continuum, it is given by the derivative of the tensor current, so that we can write for the improved vector current [110, 111]

$$(V_\mu^I)^{ff'}(x) = V_\mu^{ff'}(x) + c_V a\partial_\nu T_{\mu\nu}^{ff'}(x). \tag{3.82}$$

Here, we have defined the tensor current as

$$T_{\mu\nu}^{ff'}(x) = -\bar{\psi}^f(x)\sigma_{\mu\nu}\psi^{f'}(x), \quad \sigma_{\mu\nu} = \frac{1}{2}[\gamma_\mu, \gamma_\nu], \tag{3.83}$$

and introduced the improvement coefficient c_V. In this thesis, we employ the non-perturbative determination of c_V from Ref. [112]. Instead of the partial derivative ∂_ν, one uses the symmetric lattice derivative which introduces only $\mathcal{O}(a^2)$ lattice artefacts (*cf.*, eq. (3.15)) [112].

[8] This expression reduces to the one quoted in Ref. [83] if one assumes their specific convention for the Dirac matrices.

3.4.3 Conserved vector current

Due to the explicit breaking of chiral symmetry by the Wilson-Clover fermion action, the local discretization of the vector current in eq. (3.81) is not conserved on the lattice, *i.e.*, $\partial_\mu V_\mu \neq 0$ [112]. Using a vector variation of the action, one can derive a point-split vector current [84, 112, 113],

$$(V_\mu^c)^{ff'}(x) = \frac{1}{2}\left[\bar{\psi}^f(x+a\hat{\mu})(1+\gamma_\mu)U_\mu^\dagger(x)\psi^{f'}(x) - \bar{\psi}^f(x)(1-\gamma_\mu)U_\mu(x)\psi^{f'}(x+a\hat{\mu})\right], \quad (3.84)$$

which fulfills $\partial_\mu^*(V_\mu^c)^{ff'}(x) = 0$ as an exact lattice identity, where $\partial_\mu^*\psi(x) = (\psi(x) - \psi(x - a\hat{\mu}))/a$ is the backward lattice derivative. It follows that this conserved vector current is protected against renormalization, so that its multiplicative renormalization factor Z_V is 1 and all $\mathcal{O}(a)$ mass-dependent coefficients, which arise from the improvement at finite quark masses [111], vanish [112, 113].

In practice, we employ the symmetrized version

$$(V_\mu^{cs})^{ff'}(x) = \frac{1}{2}\left[(V_\mu^c)^{ff'}(x) + (V_\mu^c)^{ff'}(x - a\hat{\mu})\right], \quad (3.85)$$

which has the same behavior under space-time reflections as the local vector current and thus ensures that the same counterterms as in eq. (3.82) are present to remove $\mathcal{O}(a)$ lattice artefacts [113]. The coefficient c_V, however, differs between the two discretizations of the vector current [112].

4 CLS simulations for $N_f = 2 + 1$ Lattice QCD

After having discussed the lattice discretization of QCD in general, we now turn to a specific set of simulations undertaken by the Coordinated Lattice Simulations (CLS) initiative using the `openQCD` package [114]. The CLS ensembles employed in this thesis have $2 + 1$ flavors of dynamical quarks, meaning that the up- and down-quarks are treated as mass-degenerate, and the strange quark as a single, heavier flavor. In the following, we will guide the reader through the chosen action and boundary conditions, the quark mass trajectory, the algorithmic setup of the simulations and the scale setting. All of these topics have been described extensively in Refs. [90, 115], on which the present chapter draws heavily and to which we refer the interested reader for further details and references. We will conclude the chapter with an overview of the specific CLS ensembles used in this thesis.

4.1 Action and boundary conditions

The $N_f = 2 + 1$ CLS gauge ensembles are generated using the Lüscher-Weisz gauge action eq. (3.12) with the tree-level coefficients $c_0 = 5/3$, $c_1 = -1/12$ and $c_2 = c_3 = 0$, as discussed at the end of section 3.1.1. For the dynamical quarks, the Wilson-Clover action eq. (3.38) with the non-perturbatively determined Sheikholeslami-Wohlert coefficient c_{SW} from Ref. [116] is employed. In principle, as already mentioned in chapter 3, the boundary conditions should be chosen periodic in space and (anti-)periodic in time for fermions and bosons, respectively. In practice, however, some of the CLS ensembles have been generated with open boundary conditions in time, breaking translational invariance in the time direction. In the following, we will explain why this can nevertheless constitute a useful modification of the theory.

The continuum limit of a Lattice Quantum Field Theory is always defined in terms of a second-order phase transition, *i.e.*, a critical point. This means that the correlation length ξ (in lattice units) diverges as one approaches the continuum limit [80]. Intuitively, this behavior can be understood as follows: Consider a physical object with an intrinsic length scale l in physical units. The fields from which this object is built need to stay correlated at this same length scale. On a lattice with finite lattice spacing a, the object must thus be described by a system with a correlation length in lattice units of $\xi \sim l/a$. Consequently, in order for the physical length scale l to remain constant, ξ needs to diverge in the continuum limit $a \to 0$.

For Markov Chain Monte Carlo (MCMC) simulations, one expects that the autocorrelation time[1] of an observable O along the Markov chain scales as $\hat{t}_{\text{auto}}(O) \propto \xi^{z(O)}$, where z is the dynamical critical exponent, which depends on the details of the updating process and on the observable. The fact that MCMC simulations become less and less efficient as one approaches a critical point (*e.g.*, the continuum limit) is referred to as critical slowing down [80].

One observable which evolves for practically all algorithms particularly slowly in computer time is the global topological charge which is an integer number on periodic lattices. In Ref. [117], for instance, a dynamical critical exponent of $z \approx 5$ has been reported for the squared topological charge in simulations of pure gauge theory. For smaller lattice spacings, the algorithm fails to tunnel between sectors with different values of the topological charge because such transitions can be seen as non-perturbative lattice artefacts [118]. Consequently, the

[1] In section 6.1, a concrete measure for the autocorrelation time will be introduced.

topological charge is effectively frozen on the corresponding ensembles. This calls the ergodicity of the whole algorithm into question, even if one is not interested in the topological charge: the slowly moving modes of the Markov process which heavily affect the topological charge might also couple to the observable of interest, leading to potentially biased results [117].

A possible solution to this problem consists in imposing open boundary conditions in the time direction. For the gauge fields $A_\mu(x)$, one imposes the following conditions,

$$F_{0k}(x)|_{x_0=0} = F_{0k}(x)|_{x_0=a(N_\tau-1)} = 0, \quad k = 1, 2, 3, \tag{4.1}$$

which preserve gauge invariance [118]. In terms of gauge links, one sets all those links to zero where either endpoint is not contained in the lattice eq. (3.1) [114]. For the quark and antiquark fields $\psi(x)$ and $\bar{\psi}(x)$, one demands that [118]

$$\Gamma_+^p \psi(x)|_{x_0=0} = \Gamma_-^p \psi(x)|_{x_0=a(N_\tau-1)} = 0, \quad \bar{\psi}(x)\Gamma_-^p|_{x_0=0} = \bar{\psi}(x)\Gamma_+^p|_{x_0=a(N_\tau-1)} = 0, \tag{4.2}$$

where $\Gamma_\pm^p$ are the parity projectors from eq. (3.78). These boundary conditions preserve parity and time reflection symmetry [118]. Note furthermore that imposing the boundary conditions eqs. (4.1) and (4.2) effectively reduces the physical time extent of the lattice by one unit to $T = a(N_\tau - 1)$ [90].

In general, open boundary conditions may spoil the $\mathcal{O}(a)$-improvement of the action. It can be shown, however, that all bulk $\mathcal{O}(a)$-counterterms and their coefficients remain unchanged compared to the theory with periodic boundary conditions, which we have discussed so far [118]. For the improvement of functions involving fields close to or at the boundary, further $\mathcal{O}(a)$ boundary counterterms are required. In the case of the gauge action, it is sufficient for tree-level improvement to change the weight of the space-like loops C on the boundaries in eq. (3.12) from 1 to 1/2 [119]. Moreover, the sum has to be restricted to those loops which are fully contained in the lattice, *i.e.*, of which all corners are in the time interval $[0, N_\tau - 1]$ [118]. This is obviously equivalent to the condition quoted above in terms of the gauge links. Regarding the fermion action, the sum over the lattice sites has to be restricted to the time interval $[1, N_\tau - 2]$. Since the resulting action does not depend on the quark fields at the boundaries $x_0 = 0$ and $x_0 = a(N_\tau - 1)$, one may simply set all components of the quark fields at the boundaries to zero [118]. $\mathcal{O}(a)$-improvement at the boundaries here generally requires further boundary terms; their coefficients, however, vanish at tree level of perturbation theory [114].

The field space of a theory satisfying the open boundary conditions eqs. (4.1) and (4.2) is connected in a topological sense, so that there are no disconnected topological sectors as in a continuum theory with periodic boundary conditions. This means that the topological charge is not quantized and can assume arbitrary real values. In an MCMC simulation, it can thus change smoothly along a trajectory by flowing in and out of the lattice via its boundaries [118].

It can be shown that open boundary conditions in QCD are stable under quantum fluctuations, and that the space of physical states is independent of the boundary conditions [118]. This implies that many physical quantities like hadron masses can in principle be extracted from correlation functions on open boundary lattices in the same way as on lattices with periodic boundary conditions in time [118]. The only difference is, of course, that the translation invariance in time is broken and only approximately restored in the bulk of the lattice, where the effects of the boundaries are exponentially suppressed [119]. The state representing the boundary itself has the same quantum numbers as the vacuum state [118].

4.2 Algorithmic setup

In order to increase the performance and reliability of the HMC algorithm, several improvements are applied to the basic procedure described in section 3.3. We will summarize these in the

following, starting with the treatment of the light-quark determinant, continuing with details about the RHMC algorithm employed for the strange-quark determinant and finally addressing the integration of the molecular dynamics equations and the inversion of the Dirac operator.

4.2.1 Even-odd preconditioning for light quarks

In this subsection and the following two, we shall be concerned with the determinant of the two degenerate light quarks. Instead of the corresponding Wilson-Clover Dirac operator, which we denote here simply by Q and which is γ_5-hermitian, one considers the Hermitian operator $\tilde{Q} = \gamma_5 Q$ [90]. Because it involves at any given lattice site only interactions with this same site and its nearest-neighboring sites (*cf.*, eqs. (3.31) and (3.38)), it is advantageous to decompose the lattice into an even-odd checkerboard pattern, so that the nearest neighbors of a site labelled by one color are all of the other color [120]. On a thus decomposed space, the matrix $\tilde{Q}$ has the block structure [121]

$$\tilde{Q} = \begin{pmatrix} \tilde{Q}_{ee} & \tilde{Q}_{eo} \\ \tilde{Q}_{oe} & \tilde{Q}_{oo} \end{pmatrix}, \tag{4.3}$$

where e refers to even and o to odd sites. $\tilde{Q}_{ee}$ and $\tilde{Q}_{oo}$ are site-diagonal, only mixing color and spin components on each site, since there are only nearest-neighbor interactions. Therefore, these two matrices are easily inverted [121]. To facilitate the computation of the determinant of the whole matrix $\tilde{Q}$, we consider its Schur decomposition [121],

$$\tilde{Q} = UAL = \begin{pmatrix} 1 & \tilde{Q}_{eo}\tilde{Q}_{oo}^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{Q} & 0 \\ 0 & \tilde{Q}_{oo} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \tilde{Q}_{oo}^{-1}\tilde{Q}_{oe} & 1 \end{pmatrix}, \tag{4.4}$$

where $\hat{Q}$ is the even-odd Schur complement of $\tilde{Q}$,

$$\hat{Q} = \tilde{Q}_{ee} - \tilde{Q}_{eo}\tilde{Q}_{oo}^{-1}\tilde{Q}_{oe}. \tag{4.5}$$

As $\det L = \det U = 1$, we obtain for the relevant determinant,

$$\det(QQ^\dagger) = \det(\gamma_5\gamma_5 QQ^\dagger) = \det(\gamma_5 Q(\gamma_5 Q)^\dagger) = \det(\tilde{Q}^2) = \det(\tilde{Q}_{oo}^2)\det(\hat{Q}^2). \tag{4.6}$$

In effect, we have reduced the size of the matrices the determinants of which need to be calculated by a factor of two.

4.2.2 Twisted-mass reweighting for light quarks

The Wilson Dirac operator breaks chiral symmetry explicitly and is thus not protected against having arbitrarily low eigenvalues. Even if one considers its square, so that the eigenvalues are guaranteed to be non-negative, the spectral gap (the smallest eigenvalue in absolute size) can be arbitrarily small. It can be shown that the average spectral gap is proportional to the quark mass. For light quarks, not (much) heavier than the physical u- and d-quarks, the probability for finding very small eigenvalues, close to zero, on a given gauge configuration is hence significantly increased [122]. This can cause a number of numerical instabilities in the simulation, from the reversibility of the integration of the molecular dynamics equations to the ergodicity and sampling efficiency of the HMC algorithm [122].

In order to avoid such instabilities, it has been suggested in Ref. [123] to introduce a small twisted-mass term into the action and employ reweighting to correct for it. In this way, the contribution of the low modes is shifted from the HMC to the reweighting factor, thus stabilizing

the former. For the generation of the CLS ensembles, the twisted-mass modification is applied to the even-odd Schur complement $\hat{Q}$ from eq. (4.5) [90],

$$\det(QQ^\dagger) = \det(\tilde{Q}^2) = \det(\tilde{Q}_{oo}^2)\det(\hat{Q}^2) \to \det(\tilde{Q}_{oo}^2)\det\frac{\hat{Q}^2+\mu_0^2}{\hat{Q}^2+2\mu_0^2}\det(\hat{Q}^2+\mu_0^2). \tag{4.7}$$

The corresponding reweighting factor which needs to be included in the calculation of expectation values analogously to eq. (3.71) therefore reads [123],

$$W_0 = \det\frac{\hat{Q}^2(\hat{Q}^2+2\mu_0^2)}{(\hat{Q}^2+\mu_0^2)^2}. \tag{4.8}$$

The choice of the twisted-mass parameter μ_0 requires some care. In principal, larger values lead to a stronger stabilization of the HMC because the modified operator from eq. (4.7) has a spectral gap of $\mathcal{O}(\mu_0)$ [123]. However, the fluctuations of the reweighting factor also increase with μ_0 [90]. The latter can be very problematic for observables with a strong (anti-)correlation with W_0 (*i.e.*, with the low modes of the Dirac operator) because the reweighting then enlarges their statistical errors significantly.

Given a chosen value of μ_0, a precise estimation of the reweighting factor is paramount to mitigate this issue to the maximum extent possible. For this reason, we employ the reweighting factors computed with exact low-mode deflation as detailed in Ref. [124] on all ensembles except E300, where only the stochastic evaluation according to Ref. [90] is available.

4.2.3 Hasenbusch factorization of the light quark determinant

To further improve the efficiency of the calculation, the mass splitting introduced by Hasenbusch [125] is applied to the last term in eq. (4.7) [90],

$$\det(\hat{Q}^2+\mu_0^2) = \det(\hat{Q}^2+\mu_{N_{HB}}^2)\prod_{i=1}^{N_{HB}}\det\frac{\hat{Q}^2+\mu_{i-1}^2}{\hat{Q}^2+\mu_i^2}, \tag{4.9}$$

where $\mu_0 < \mu_1 < \cdots < \mu_{N_{HB}}$. This factorization leads to a reduced condition number for each of the factors compared to the original matrix. Therefore, one expects the fermionic contribution to the force in eq. (3.61) to be smaller, so that the step size for the integration of the molecular dynamics equations can be chosen larger [125].

In practice, each of the $N_{HB}+1$ determinants in eq. (4.9) is computed with a separate set of pseudofermions [125]. While the diagonal determinant $\det\tilde{Q}_{oo}$ is combined with the one with the largest twisted mass (*i.e.*, with the first factor in eq. (4.9)), another set of pseudofermions is introduced for the middle factor in eq. (4.7). Hence, the total number of pseudofermion fields required for the light quarks is $N_{HB}+2$ [90].

We note that the introduction of further pseudofermions via the mass splitting implies a slight overhead in the computation of the fermion forces because some additional inversions become necessary. If the μ_i parameters are chosen adequately, this is, however, more than compensated by the speed-up achieved thanks to the larger step size [125]. For details on the tuning of N_{HB} and the μ_i parameters in the context of CLS, we refer to Ref. [90].

4.2.4 RHMC for the strange quark determinant

So far, we have discussed the efficient implementation of the HMC algorithm for the two degenerate light quarks. For the strange quark, the RHMC algorithm explained in section 3.3.2 is employed. This can again be combined with the even-odd preconditioning from section 4.2.1.

Accordingly, the modulus of the strange-quark determinant is approximated by a rational function of order $[N_p, N_p]$ [90],

$$|\det(Q^s)| = \det(Q^s_{oo})\det(\hat{Q}^s) = \det(Q^s_{oo})\det\left(\sqrt{(\hat{Q}^s)^2}\right) \approx \det(Q^s_{oo})\det\left(A^{-1}\prod_{i=1}^{N_p}\frac{\hat{Q}^2+\bar{\mu}_i^2}{\hat{Q}^2+\bar{\nu}_i^2}\right). \tag{4.10}$$

The parameters A^{-1} and $\bar{\mu}_i$, $\bar{\nu}_i$ are known and can be determined from Zolotarev's optimal approximation of the inverse square root [126] in the interval $[r_a, r_b]$ which is given by the smallest and the largest eigenvalue of $(\hat{Q}^s)^2$ [90]. The approximation in eq. (4.10) needs to be corrected for by introducing a corresponding reweighting factor W_1, while another reweighting factor W_s becomes necessary due to the neglect of a possible sign of $\det(Q^s)$ (*cf.*, section 3.3.2). W_1 can again be computed very precisely using exact low-mode deflation [124].

In the simulation, the N_p' factors with the smallest $\bar{\mu}_i$ are evaluated with individual pseudofermion fields, while the remaining $N_p - N_p'$ factors together with $\det(Q^s_{oo})$ are expressed as a single pseudofermion integral. The total number of pseudofermion fields employed for the strange quark is thus $N_p' + 1$. This decomposition allows one to optimize the inversion of the corresponding matrices as well as the step size used for the integration of the molecular dynamics equations separately for each determinant [90].

4.2.5 Integration of the molecular dynamics

For the integration of the molecular dynamics eqs. (3.59) and (3.60) only the simplest method, the second-order leapfrog algorithm, has been presented in section 3.3.1. This can be improved upon in several ways.

First of all, one notices that the gauge contribution to the force in eq. (3.61) is much cheaper to compute than the fermionic contribution. Moreover, we have split the latter into $N_{HB} + N_p' + 3$ different components, each of which is represented by its own set of pseudofermions, and some of which are considerably smaller and/or less fluctuating than others [90]. Sexton and Weingarten have suggested in Ref. [127] that in such a situation, a more efficient integration scheme can be achieved by splitting the integration into several levels. The gauge force is integrated on an inner level with a smaller step size, and the fermion force on an outer level (or possibly several outer levels) with a larger step size. Since the time-consuming part of the calculation is the evaluation of the fermion force, reducing the step size for the gauge force barely increases the overall cost, but significantly increases the precision [127]. A more precise integration of the molecular dynamics equations leads to a lower degree of energy violation and thus a larger acceptance rate in the Metropolis step, making the HMC algorithm overall more efficient.

In the setup chosen by CLS [90], three different levels are employed: an inner level for the gauge force, an intermediate level for most of the fermion forces and an outer level for those fermion forces which contribute particularly little to the energy violation. Integrating these less precisely saves a substantial amount of computing time for the integration, while barely affecting the acceptance rate.

A further improvement of the integration can be achieved by using a more optimized, higher-level integrator than the simple leapfrog algorithm. As higher-level integrators generally require more computing time than lower-level ones, this choice must always be weighed against the achievable improvement in the acceptance rate. Therefore, for the inner two levels, the fourth-order integrator suggested by Omelyan, Mryglod and Folk [128] is employed (yet with different step sizes, as discussed above), while for the outermost level, their second-order integrator is used.

The thus defined integration algorithm leaves as tunable parameters the trajectory length

$\Delta\hat{t}$ in computer time as well as the number of fermion terms which are integrated on the outermost level. The latter are given by $N_{HB,2}$ Hasenbusch factors and $N'_{p,2}$ poles from the rational approximation of the strange-quark determinant. While longer trajectories lead to shorter autocorrelation times [117], they might, at some point, also give rise to instabilities in the integrator [90]. As a compromise, $\Delta\hat{t} = 2$ has been used on all ensembles except J303, where $\Delta\hat{t} = 4$. The chosen values for $N_{HB,2}$ and $N'_{p,2}$ can be found in Ref. [90].

4.2.6 Solver

In order to compute the fermion forces, the inverse of a matrix M related to the lattice Dirac operator (for example via eq. (4.7), eq. (4.9) or eq. (4.10)) needs to be computed. As all that is needed is the inverse of M applied to the pseudofermion field ϕ (*cf.*, eq. (3.61)), this is equivalent to solving the matrix-vector equation $M\psi = \phi$ for ψ. The Dirac operator and hence M is a very large matrix, so that a direct solution is impossible. A suitable alternative is provided by iterative methods. With decreasing quark masses, however, the probability for finding small eigenvalues of the Dirac operator is enhanced, as discussed in section 4.2.2. Consequently, the condition number of M rises, requiring larger and larger iteration counts.

A significant improvement can be achieved by applying low-mode deflation [129]. Here, one separates the low modes of M and computes the solution on them exactly. The iterative algorithm thus only needs to be applied to the deflated equation, where the contribution of the deflation subspace, *i.e.*, of the low modes, has been projected out. A practical problem which arises with the deflation technique is that the number of eigenvalues below a certain threshold grows proportionally to the lattice volume V. The computational effort required to solve the eigenvalue problem and construct the deflation subspace from the eigenvectors is then at least proportional to V^2 [129]. This renders the direct application of low-mode deflation on large volumes infeasible.

To overcome this burden, one employs domain-decomposed deflation [129, 130], which does not need the deflation subspace to be spanned by exact eigenvectors of M. Rather, the lattice is decomposed into non-overlapping rectangular blocks of fixed size, and the eigenvectors are computed on these instead of on the global lattice. Local coherence of the low modes of the Dirac operator [129] then guarantees that the local deflation subspaces still have a good overlap with the (globally defined) low modes.

The remaining deflated equation is solved by an iterative method which can be further improved by applying a preconditioner. In the case of CLS, a preconditioner based on the Schwarz alternating procedure (SAP) [131, 132] is used. The SAP was originally devised to solve elliptic partial differential equations. In that context it is based on a division of the domain on which the solution is sought into overlapping subdomains [133]. If it is only used as a preconditioner, the subdomains can also be chosen to be non-overlapping, which reduces the number of required subdomains [132].[2]

One hence proceeds by partitioning the lattice into a grid of non-overlapping rectangular blocks so that they can be alternately colored in a checkerboard pattern. Note the difference to the even-odd preconditioning in section 4.2.1, where the coloring was imposed on the level of individual sites, while for the SAP, it is imposed on the block level. It has been shown in Ref. [132] that all blocks of the same color can be treated simultaneously, so that the algorithm operates on the level of two domains Ω and Ω^* containing the union of all blocks of one color each. In each iteration cycle, the current approximate solution ψ is first updated on Ω while keeping it fixed on Ω^*, and then the other way around [132].

[2] In the following, we will therefore use the term SAP, somewhat abusively, also for the method with non-overlapping subdomains.

Taken on its own, the SAP is not sufficiently efficient as a solver. It can, however, play the role of a preconditioner providing an approximate solution and thus reducing the required iteration count of the actual solver [129, 132]. For the latter, the generalized conjugate residual (GCR) algorithm [134] is employed. It is a Krylov space solver which is mathematically equivalent to the generalized minimal residual (GMRES) algorithm [135], but is technically advantageous if an approximate preconditioner (like the SAP) is being used [132].

In summary, the deflation projection and the Schwarz preconditioning interact by reducing the low-mode and high-mode contributions to the residue in alternation, and in this way significantly speed up the GCR solver [129].

4.3 Quark mass trajectory

In order to match ensembles at different lattice spacings, one introduces the dimensionless parameters [90]

$$\phi_2 = 8t_0 M_\pi^2, \qquad \phi_4 = 8t_0 \left(M_K^2 + \frac{1}{2} M_\pi^2 \right), \tag{4.11}$$

where t_0 is the Wilson flow time parameter [136] which will be discussed in more detail in the next section. To leading order in chiral perturbation theory (χPT), the parameters in eq. (4.11) are proportional to the sum of the quark masses, $\phi_2 \propto (m_u + m_d)$ and $\phi_4 \propto (m_u + m_d + m_s)$ [90].

Since the CLS ensembles are not simulating the full Standard Model, but only QCD with three quark flavors and exact isospin symmetry, the physical values of the meson masses need to be defined in this limit. The analysis in Ref. [137] has shown that in the isospin limit of pure QCD, the mass of the charged pion and the root-mean-square mass of the K^+ and K^0 are, respectively, given by

$$M_{\pi,\mathrm{phys}} = 134.8(3)\ \mathrm{MeV}, \qquad M_{K,\mathrm{phys}} = 494.2(4)\ \mathrm{MeV}. \tag{4.12}$$

Because t_0 is not an experimentally measurable observable, its value needs to be determined from lattice simulations. We use the FLAG estimate [138] of

$$\sqrt{t_{0,\mathrm{phys}}} = 0.14464(87)\ \mathrm{fm} \tag{4.13}$$

for the value of $\sqrt{t_0}$ at the physical point defined above in $N_f = 2 + 1$ QCD, and hence obtain for the physical values of the parameters in eq. (4.11),

$$\phi_{2,\mathrm{phys}} = 0.0781(10), \qquad \phi_{4,\mathrm{phys}} = 1.089(13), \tag{4.14}$$

where we have added the errors originating from eqs. (4.12) and (4.13) in quadrature.

When we were discussing the $\mathcal{O}(a)$-improvement of the Wilson fermion action towards the end of section 3.1.2, we have found two counterterms (eqs. (3.35) and (3.37)) which were already part of the original action, and have absorbed them in a redefinition of the bare coupling g and the bare mass m. If a renormalization scheme is chosen in which all renormalization conditions are imposed at the same point in bare parameter space, such rescalings of the bare parameters do not alter the renormalized amplitudes [96]. However, one conventionally uses a mass-independent renormalization scheme, where the renormalization conditions are imposed at zero quark masses. In this case, reparametrizations of the bare theory must be taken care of because the link between the massless and the massive theory is established via the bare parameters [96]. In particular, one introduces a modified bare coupling and modified bare quark

masses according to [96, 111]

$$\tilde{g}^2 = g^2 \left[1 + \frac{b_g(g^2)}{N_f} a \operatorname{tr} M\right], \tag{4.15}$$

$$\begin{aligned}\tilde{m}^f_{\text{sub}} = m^f_{\text{sub}} + \frac{r_m(g^2) - 1}{N_f} \operatorname{tr} M + a \Big[& b_m(g^2)(m^f_{\text{sub}})^2 + \bar{b}_m(g^2) m^f_{\text{sub}} \operatorname{tr} M \\ & + \frac{r_m(g^2) d_m(g^2) - b_m(g^2)}{N_f} \operatorname{tr}(M^2) \\ & + \frac{r_m(g^2) \bar{d}_m(g^2) - \bar{b}_m(g^2)}{N_f} (\operatorname{tr} M)^2 \Big]\end{aligned} \tag{4.16}$$

to account for the additional $\mathcal{O}(a)$-improvement terms from eqs. (3.35) and (3.37). Here, we have introduced the subtracted bare quark masses $m^f_{\text{sub}} = m^f - m_c$, where m_c is the critical quark mass. The latter is the value of the bare quark mass for which, given completely degenerate quarks, the renormalized quark mass vanishes [80, 111]. Moreover, we have defined the matrix M containing the subtracted bare quark masses. For $N_f = 2 + 1$, for example, it reads $M = \operatorname{diag}(m^l, m^l, m^s) - m_c \mathbb{1}_3$ with $m^u = m^d = m^l$.

For many years, in particular the coefficient b_g necessary for the $\mathcal{O}(a)$-improvement of the bare coupling constant has only been known perturbatively [139]. While a first non-perturbative calculation has meanwhile appeared [140], this has not been the case at the point in time when the generation of the $N_f = 2 + 1$ CLS ensembles has been started [90]. The preferred strategy therefore consists in defining a quark mass trajectory along which $\operatorname{tr} M$, *i.e.*, the sum over the subtracted bare quark masses, is kept constant [90]. This is obviously equivalent to keeping the sum over the bare quark masses themselves constant. Along such a trajectory, it is, according to eq. (4.15), sufficient to keep g fixed in order to keep $\tilde{g}$ and thus the lattice spacing fixed, so that knowledge of b_g is not required [90]. We note that according to eq. (4.16), even if $\operatorname{tr} M$ is constant, the sum of the renormalized quark masses $m^f_R = Z_m \tilde{m}^f_{\text{sub}}$ is not, but differs by $\mathcal{O}(am)$-effects [111, 115].

For the $N_f = 2 + 1$ CLS simulations, the different lattice spacings have been matched at the SU(3)-flavor-symmetric point, where $M_\pi = M_K \approx 420$ MeV [90]. This means that the quark mass trajectory is characterized by $2m^l + m^s = 3m_{\text{sym}} = \text{const}$, where m_{sym} is the bare quark mass at the aforementioned symmetric point [115].

4.4 Scale setting

Because computer simulations can only deal with dimensionless variables, all observables are primarily expressed in units of the lattice spacing a. In order to make contact with physical quantities defined in the continuum, the value of a on the generated ensembles needs to be related to an experimentally determined observable. This observable is thus treated as input to the simulations from which subsequently other dimensionful quantities can be predicted.

To set the scale for the CLS ensembles, the Wilson flow time parameter t_0 [136] is used as an intermediate theory scale. It has the advantage that it is a purely gluonic quantity which can be measured very precisely on the lattice. On the other hand, it is not directly accessible in experiments, so that its physical value needs to be determined from the lattice simulations themselves as part of the scale setting procedure.

One therefore proceeds by measuring the dimensionless ratio t_0/a^2 and a hadronic quantity which is accessible in experiment on a number of ensembles at different values of g (or, equivalently, $\beta = 6/g^2$) and ϕ_2. In the case of the CLS scale setting [115], a linear combination

of decay constants of the pion and kaon,

$$f_{\pi K} = \frac{2}{3}\left(f_K + \frac{1}{2}f_\pi\right), \tag{4.17}$$

is employed as the hadronic quantity required to make contact to experiment. This is convenient because χPT predicts it to be constant up to small corrections along the trajectory from section 4.3 [141]. For the purpose of the scale setting analysis, ϕ_4 is used as a proxy for the sum of the quark masses, because keeping ϕ_4 fixed leads to an $\mathcal{O}(a)$-improved trajectory in contrast to the sum of the bare quark masses (*cf.*, the discussion in section 4.3). This might introduce a variation of the improved coupling constant $\tilde{g}$ due to higher-order effects in χPT, which have, however, been found to be negligibly small [115].

The scale setting analysis therefore proceeds in the following steps: First, using some initial value of t_0 in physical units, the target point in the (ϕ_2, ϕ_4)-plane is defined. The observables $\sqrt{t_0}/a$, $af_{\pi K}$ and their product are then shifted to the target ϕ_4. Using different fit *ansätze* [115], the thus shifted values of $\sqrt{t_0}f_{\pi K}$ are subsequently extrapolated to the target ϕ_2 and the continuum limit. Dividing by the experimental value of $f_{\pi K}$ in physical units, one hence arrives at a value of $\sqrt{t_0}$ in fm. The procedure is then iterated until a fixed point is reached, *i.e.*, the output value for $\sqrt{t_0}$ in fm agrees with the input value. This fixed point defines the physical value $\sqrt{t_{0,\mathrm{phys}}}$ in the continuum and thus the actual target ϕ_4.

On a trajectory defined by constant ϕ_4, χPT calculations show that $\sqrt{t_0}/a$ is (approximately) independent of ϕ_2 [142]. This has been confirmed to a large degree on the CLS ensembles [143]. Therefore, each value of β uniquely corresponds to a value of $\sqrt{t_0}/a$. For the determination of the latter, the extrapolation step can hence be avoided: one only needs to take the values for this observable on one ensemble at the symmetric point for each value of β and shift them to the final target value of ϕ_4 [115]. We will call the thus defined quantity in the following $\sqrt{t_0^{\mathrm{sym}}}/a$. For further details on the CLS scale setting analysis, we refer to Ref. [115].

A major advantage of using an intermediate theory scale is that the calibration, *i.e.*, the value of $\sqrt{t_{0,\mathrm{phys}}}$ in fm, can be replaced easily once updated calculations become available. In this thesis, we employ the FLAG estimate for $N_f = 2 + 1$ [138] quoted in eq. (4.13) above. It represents an average over the results of CLS [115], RBC/UKQCD [144] and BMW [145]. For our analysis, the utilization of an intermediate theory scale implies that we first need to express all dimensionful quantities in units of t_0 using the t_0^{sym}/a^2 values from Ref. [115], which are listed in table 4.1. Only our final results are converted to physical units by means of the calibration, *i.e.*, $\sqrt{t_{0,\mathrm{phys}}}$ in fm. This procedure ensures that the error of the calibration is treated independently of that of the (more precise) pure lattice measurement of t_0^{sym}/a^2. In particular, it avoids an artificial increase of errors on intermediate results, so that fit qualities remain meaningful.

4.5 Ensembles

Due to the small, but finite mistuning of ϕ_4 in the simulations, the quark mass trajectory defined in section 4.3 does in reality not exactly hit the physical point. In order to enable a correction for this effect in extra- or interpolation fits to the physical point, further trajectories, for example at constant (renormalized) strange quark mass [146], have been generated by CLS.

However, the mistuning is expected to only have a negligibly small effect on the observables which shall be investigated in this thesis because they only depend very weakly on the strange quark mass. This is due to the fact that we will restrict our studies to the nucleon which only has light quarks as valence quarks. The strange quark only enters via suppressed sea-quark contributions. Even though the latter are not entirely negligible, a small difference in the

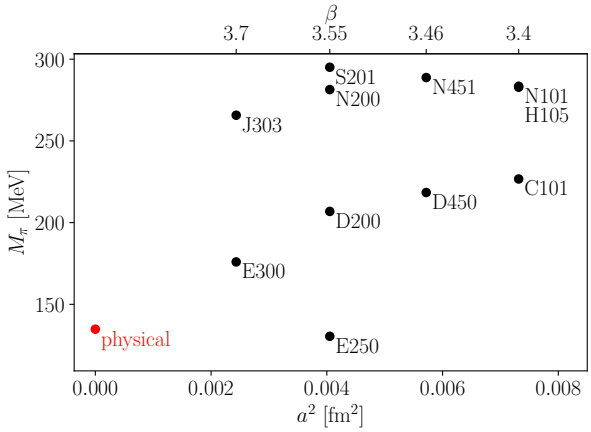


Figure 4.1: Location of the ensembles used in this thesis in the (M_π, a^2)-plane

strange quark mass at the target point from its physical value will have an even smaller impact on the total (valence + sea) result for the nucleon. Therefore, it is, within our current statistical precision, sufficient to only employ the $\mathrm{tr}\, M = \text{const.}$ ensembles and extra- or interpolate to the correct ϕ_2, *i.e.*, pion mass.

The pion mass, on the other hand, has a very strong influence on our observables. Ensembles on which the pion is much heavier than ≈ 300 MeV can hence be considered to be too remote from the physical point to be of significant importance to the extrapolation to the latter. Consequently, we restrict our analysis to the ensembles shown in fig. 4.1 and table 4.1. They cover four lattice spacings in the range from 0.049 fm to 0.086 fm, which gives us a good control over the approach to the continuum limit. The pion masses, on the other hand, go down to the slightly-below-physical value of 130 MeV (E250), so that we can actually perform an interpolation in this variable rather than an extrapolation. The values of t_0^{sym}/a^2 listed in table 4.1 have been obtained in Ref. [115] as described in section 4.4. These are converted to a in fm using eq. (4.13)[3]. The quoted numbers for the pion and nucleon masses correspond to the ones listed in table 6.1 in units of $\sqrt{t_0}$, but converted to physical units using eq. (4.13) and rounded to full MeV. In the actual analysis, the exact values determined in section 6.4 are employed in units of $\sqrt{t_0}$. Similarly, a in fm is never employed in the actual analysis and merely quoted here for ease of reference.

The naming scheme for the ensemble IDs has been chosen as follows: The first character is a single letter encoding the size of the box in lattice units ($N_s^3 \times N_\tau$). It is followed by three digits, the first of which corresponds to the value of β, *i.e.*, the lattice spacing. The second digit can be either 0 or 5, where the former stands for open boundary conditions in time and the latter for (anti-)periodic boundary conditions. The last digit simply enumerates ensembles with otherwise identical IDs, for example at different pion masses or even on different quark mass trajectories.

[3] and therefore slightly differ from the ones quoted in Ref. [115].

Table 4.1: Overview of the ensembles used in this thesis. $N_{\text{cfg}}^{\text{conn,disc}}$ refer to the number of configurations used for the computation of the quark-connected and -disconnected contributions, respectively (*cf.*, section 5.4). The remaining columns are explained in the text.

ID	β	t_0^{sym}/a^2	a [fm]	N_τ	N_s	M_π [MeV]	m_N [MeV]	$N_{\text{cfg}}^{\text{conn}}$	$N_{\text{cfg}}^{\text{disc}}$
C101	3.40	2.860(11)	0.08553(54)	96	48	227	977	1988	994
N101	3.40	2.860(11)	0.08553(54)	128	48	283	1031	1588	1588
H105	3.40	2.860(11)	0.08553(54)	96	32	283	1051	1024	1024
D450	3.46	3.659(16)	0.07561(48)	128	64	218	977	498	498
N451	3.46	3.659(16)	0.07561(48)	128	48	289	1054	1010	1010
E250	3.55	5.164(18)	0.06365(40)	192	96	130	939	398	796
D200	3.55	5.164(18)	0.06365(40)	128	64	207	991	1996	998
N200	3.55	5.164(18)	0.06365(40)	128	48	281	1064	1708	1708
S201	3.55	5.164(18)	0.06365(40)	128	32	295	1136	2092	2092
E300	3.70	8.595(29)	0.04934(31)	192	96	176	976	569	569
J303	3.70	8.595(29)	0.04934(31)	192	64	266	1046	1073	1073

5 Nucleon matrix elements from Lattice QCD

Matrix elements of current operators between nucleon states can be used to probe the internal structure of nucleons. Since the dynamics at the typical hadronic scales is entirely dominated by the strong nuclear force, the structure of hadrons can be studied to a good approximation in pure QCD. The coupling of QCD at these scales is large, as can be seen from fig. 2.1. This implies that perturbation theory cannot be applied; a non-perturbative treatment is absolutely required. Lattice QCD provides a versatile, well-tested and systematically improvable tool for such non-perturbative studies.

The calculation of physically interesting quantities from the lattice proceeds in three essential steps: Firstly, gauge configurations following the distribution eq. (3.50) are generated. This is the topic which we have discussed in the previous chapter. Secondly, one measures observables, typically n-point correlation functions involving the particles of interest, on the previously generated background gauge fields. In the present chapter, we shall be concerned with this step. Lastly, the data for the correlation functions need to analyzed in order to extract the physically relevant properties which will enable a comparison to experiments. Accordingly, the next chapter will be dedicated to our data analysis.

A systematic framework for the computation of nucleon matrix elements on the lattice already exists. In the following, we will focus on the electromagnetic structure of the nucleon as parametrized by the corresponding form factors, and present the relevant observables as well as the strategies used for their calculation ("measurement"). This chapter is partly inspired by the PhD theses [67, 84].

5.1 Electromagnetic form factors and radii

In order to study the internal structure of nucleons experimentally, one can scatter low-energy leptons off them. The latter have the advantage of being elementary particles with no complicated internal structure themselves, which significantly simplifies the description of the scattering process. For the scattering amplitude to be sensitive to the electromagnetic properties of the nucleon, the leptonic probe needs to be charged. The simplest charged lepton is the electron which is therefore widely used as a precision probe of the electromagnetic structure of the nucleon.

The single-photon-exchange amplitude contributing to the corresponding elastic scattering cross section can be parametrized by matrix elements of the electromagnetic current between final and initial nucleon states [147]. In Minkowski space, we define these matrix elements as

$$\left\langle N(\mathbf{p}'^{(M)}, s') \,\middle|\, V_{\mathrm{em}}^{(M)\mu}(x^{(M)}) \,\middle|\, N(\mathbf{p}^{(M)}, s) \right\rangle = e^{iq^{(M)}\cdot x^{(M)}} \bar{u}^{s'}(\mathbf{p}'^{(M)}) \mathcal{V}_{\mathrm{em}}^{(M)\mu}(q^{(M)}) u^{s}(\mathbf{p}^{(M)}), \quad (5.1)$$

where $N(\mathbf{p}^{(M)}, s)$ denotes a nucleon state with Minkowski three-momentum $\mathbf{p}^{(M)}$ and spin s, $u^s(\mathbf{p}^{(M)})$ the corresponding Dirac spinor, and $q^{(M)\mu} = p'^{(M)\mu} - p^{(M)\mu}$ the Minkowski four-momentum transfer. The electromagnetic vector current of quarks is given by the sum over the flavor-diagonal currents [64],

$$V_{\mathrm{em}}^{(M)\mu}(x^{(M)}) = \sum_f Q^f V^{(M)\mu f}(x^{(M)}) = \sum_f Q^f \bar{\psi}^f(x^{(M)}) \gamma^{(M)\mu} \psi^f(x^{(M)}), \quad (5.2)$$

where Q^f is the charge (in units of the elementary charge e) of quark flavor f. Thus, we have

$$V_{\text{em}}^{(M)\mu} = \frac{2}{3}V^{(M)\mu u} - \frac{1}{3}V^{(M)\mu d} - \frac{1}{3}V^{(M)\mu s} + \cdots = \frac{1}{6}\left(V^{(M)\mu S} + 3V^{(M)\mu V}\right) + \cdots, \tag{5.3}$$

where the ellipsis denotes the terms for the charm and heavier quarks, and we have defined for future convenience the isovector and isoscalar combinations $V = u - d$ and $S = u + d - 2s$, respectively.

In eq. (5.1), we have furthermore introduced the photon-nucleon vertex function $\mathcal{V}_{\text{em}}^{(M)\mu}(q^{(M)})$. Lorentz invariance, parity conservation and the conservation of the vector current (in the continuum) can be used to constrain the most general form which $\mathcal{V}_{\text{em}}^{(M)\mu}$ can take. Further employing the Gordon identity, one can write the vertex function as [64]

$$\mathcal{V}_{\text{em}}^{(M)\mu}(q^{(M)}) = \gamma^{(M)\mu}F_1^N(q^2) + i\frac{\sigma^{(M)\mu\nu}q_\nu^{(M)}}{2m_N}F_2^N(q^2), \tag{5.4}$$

where m_N is the mass of the nucleon, $\sigma^{(M)\mu\nu} = i/2[\gamma^{(M)\mu}, \gamma^{(M)\nu}]$, and we have introduced the nucleon's Dirac and Pauli form factors F_1^N and F_2^N, which can only depend on q^2. In an electron-nucleon scattering context, F_1^N represents the (electron) helicity-conserving and F_2^N the (electron) helicity-flip amplitude [147]. However, the electric and magnetic Sachs form factors [147]

$$G_E^N(Q^2) = F_1^N(Q^2) - \frac{Q^2}{4m_N^2}F_2^N(Q^2), \quad G_M^N(Q^2) = F_1^N(Q^2) + F_2^N(Q^2), \tag{5.5}$$

are the experimentally more relevant linear combinations, as we shall see below. Here, we have defined $Q^2 = -q^2$, so that $Q^2 > 0$ in the space-like region (which is the only region accessible on a Euclidean lattice). The form factors are thus the essential quantities necessary for understanding the behavior of nucleons in low-energy scattering experiments, *i.e.*, their response to external probes.

The differential cross section for elastic electron-proton scattering in the one-photon-exchange approximation is given by [147]

$$\frac{d\sigma}{d\Omega_e} = \frac{\alpha_{\text{em}}^2}{4E_e^2}\frac{\cos^2(\theta_e/2)}{\sin^4(\theta_e/2)}\frac{E_e'}{E_e}\frac{\epsilon(G_E^p(Q^2))^2 + \tau(G_M^p(Q^2))^2}{\epsilon(1+\tau)}, \tag{5.6}$$

where E_e and E_e' are the energies of the incident and scattered electron, respectively, θ_e is the electron scattering angle and α_{em} the electromagnetic fine-structure constant. Moreover, we have defined the kinematic parameters τ and ϵ,

$$\tau = \frac{Q^2}{4m_N^2}, \quad \epsilon = \left[1 + 2(1+\tau)\tan^2\left(\frac{\theta_e}{2}\right)\right]^{-1}. \tag{5.7}$$

A simple technique to disentangle the contributions of the electric and magnetic form factor to the cross section is to perform a Rosenbluth separation. Here, one measures the cross section for different beam energies and scattering angles. At constant $Q^2 = 4E_eE_e'\sin^2(\theta_e/2)$, *i.e.*, constant τ, one can then perform a linear fit to the reduced cross section,

$$\sigma_R = \frac{4E_e^2}{\alpha_{\text{em}}^2}\frac{\sin^4(\theta_e/2)}{\cos^2(\theta_e/2)}\frac{E_e}{E_e'}\epsilon(1+\tau)\frac{d\sigma}{d\Omega_e} = \epsilon(G_E^p(Q^2))^2 + \tau(G_M^p(Q^2))^2, \tag{5.8}$$

as a function of ϵ, and extract the slope $(G_E^p)^2$ and the intercept $\tau(G_M^p)^2$ [147]. For modern experiments with a large number of data points, one can alternatively directly fit the measured

cross sections to the ones obtained from different models for the form factors, an approach extensively studied in Ref. [14].

In the limit $Q^2 \to 0$, the virtual photon which mediates the electromagnetic interaction between the electron and the nucleon has a very long wavelength. Thus, it sees the nucleon as a point-like particle with a given electric charge z (+1 for the proton, 0 for the neutron) and magnetic moment μ_M. In this limit, the form factors therefore become $G_E(0) = z$ and $G_M(0) = \mu_M$ [147]. These relations can also be strictly derived by using the non-relativistic limit of the spinors in eq. (5.1) [64]. For small but finite Q^2, the electric form factor dominates the cross section eq. (5.6) which then takes the form of the cross section for an electron scattering from a static charge distribution. Hence, in this limit, the electric and magnetic form factors can be interpreted as the Fourier transforms of the nucleon's charge and magnetization densities, respectively [147].[1]

Analogously, the radii of these distributions can be defined as the slope of the corresponding form factors at $Q^2 = 0$ [147],

$$\langle r^2_{E,M} \rangle = -\frac{6}{G_{E,M}(0)} \left. \frac{\partial G_{E,M}(Q^2)}{\partial Q^2} \right|_{Q^2=0} . \tag{5.9}$$

For the electric radius of the neutron, which is electrically neutral and thus has $G^n_E(0) = 0$, the normalization factor is omitted from the definition,

$$\langle r^2_E \rangle^n = -6 \left. \frac{\partial G^n_E(Q^2)}{\partial Q^2} \right|_{Q^2=0} . \tag{5.10}$$

Note that it will turn out that $\langle r^2_E \rangle^n < 0$. This clearly shows that while this definition of radius is useful to gain insights about the charge distribution inside the neutron, it can obviously not define its size [148].

A quantity which comes closer to the classical notion of size is the mechanical radius. It is defined in terms of the D-term of the nucleon,

$$\langle r^2_{\text{mech}} \rangle = \frac{6D(0)}{\int_0^\infty dQ^2 \, D(Q^2)} . \tag{5.11}$$

The D-term is one of the nucleon's gravitational form factors which are defined by the nucleon matrix elements of the energy-momentum tensor [149, 150]. The calculation of such observables is, however, outside the scope of this thesis which has a focus on the electromagnetic form factors and the corresponding radii given by eqs. (5.9) and (5.10).[2]

In Lattice QCD, we work in a Euclidean space-time, while the nucleon matrix elements we are interested in are defined in Minkowski space (*cf.*, eqs. (5.1) and (5.4)). Therefore, we need to translate these to the Euclidean notation. In the following, we will not translate the zero-component of the momentum, because it permits the interpretation as energy only in Minkowski space, $p^{(M)0} = E_{\mathbf{p}}$. We will hence work with the energies instead. For the Fourier phase factor in eq. (5.1), we obtain using eqs. (2.40) and (2.41),

$$e^{iq^{(M)} \cdot x^{(M)}} = e^{iq^{(M)}_0 x^{(M)0} + iq^{(M)}_k x^{(M)k}} = e^{(E_{\mathbf{p}'} - E_{\mathbf{p}}) x^{(E)}_0 - iq^{(E)}_k x^{(E)}_k} . \tag{5.12}$$

Regarding the photon-nucleon vertex function, we start by noting that if one uses the conserved vector current (*cf.*, eqs. (3.84) and (3.85)), the same arguments which lead to the continuum parametrization in eq. (5.4) are also valid on the lattice.

[1] Note that in a relativistic setting at arbitrary Q^2, the interpretation of the form factors is not as straightforward any more.

[2] For a recent lattice calculation of the proton's D-term and mechanical radius, we refer to Ref. [151].

For the translation of this equation to Euclidean space-time, we need to take the different convention used for the Dirac matrices into account. It is related to the Minkowski convention by eq. (2.45). First of all, we notice that we are interested in the matrix elements of the vector current which contains a factor of γ^μ. Consequently, its spatial components in Euclidean space differ by a factor of $(-i)$ from those in Minkowski space. In order to compute the correct Minkowski-space matrix elements, which we are ultimately interested in, we take this factor of $(-i)$ directly into account in our definition of $\mathcal{V}^{(E)}_{\mathrm{em},k}$. The term proportional to F_1^N in eq. (5.4) hence takes the same form in Minkowski and in Euclidean space. For the term proportional to F_2^N, we remind the reader that we have defined $\sigma^{(M)\mu\nu} = i/2[\gamma^{(M)\mu}, \gamma^{(M)\nu}]$ and $\sigma^{(E)}_{\mu\nu} = 1/2[\gamma^{(E)}_\mu, \gamma^{(E)}_\nu]$. Using eq. (2.45), we thus have that $\sigma^{(E)}_{0j} = -\sigma^{(M)0j}$ and $\sigma^{(E)}_{jk} = i\sigma^{(M)jk}$. With these conventions, we find for the photon-nucleon vertex function in Euclidean notation,

$$\mathcal{V}^{(E)}_{\mathrm{em},0} = \mathcal{V}^{(M)0}_{\mathrm{em}} = \gamma^{(M)0} F_1^N(Q^2) + i\frac{\sigma^{(M)0k} q_k^{(M)}}{2m_N} F_2^N(Q^2) = \gamma_0^{(E)} F_1^N(Q^2) + i\frac{\sigma^{(E)}_{0k} q_k^{(E)}}{2m_N} F_2^N(Q^2), \tag{5.13}$$

$$\begin{aligned}\mathcal{V}^{(E)}_{\mathrm{em},k} &= -i\mathcal{V}^{(M)k}_{\mathrm{em}} = -i\left[\gamma^{(M)k} F_1^N(Q^2) + \frac{i}{2m_N}(\sigma^{(M)k0} q_0^{(M)} + \sigma^{(M)kj} q_j^{(M)}) F_2^N(Q^2)\right] \\ &= \gamma_k^{(E)} F_1^N(Q^2) + i\frac{i\sigma^{(E)}_{k0}(E_{\mathbf{p}'} - E_{\mathbf{p}}) + \sigma^{(E)}_{kj} q_j^{(E)}}{2m_N} F_2^N(Q^2).\end{aligned} \tag{5.14}$$

In the following, we will drop the superscripts (E) again, so that all quantities are implicitly Euclidean, unless otherwise noted.

For the lattice computation, it is convenient to use the isospin basis, *i.e.*, the currents $V^{\mu S}$ and $V^{\mu V}$ instead of the electromagnetic combination in eq. (5.2). Moreover, we assume strong SU(2) isospin symmetry, so that we only need to calculate matrix elements with the proton interpolator from eq. (3.76). The matrix elements of the electromagnetic vector current between proton states can be reconstructed from the ones of the isovector and isoscalar vector currents simply by building the correct linear combination according to eq. (5.3),

$$\langle p|V^\mu_{\mathrm{em}}|p\rangle = \frac{1}{6}\left(\langle p|V^{\mu S}|p\rangle + 3\langle p|V^{\mu V}|p\rangle\right). \tag{5.15}$$

Here, we neglect the contributions of the charm and heavier quarks, which are expected to be much smaller than our statistical uncertainties.

Deriving the equation needed to infer the matrix elements of the neutron from those computed for the proton requires somewhat more algebra. Under the assumption of SU(2) isospin symmetry, it follows from an application of the Wigner-Eckart theorem in isospin space that [53]

$$\langle p|V^\mu_{\mathrm{em}}|p\rangle - \langle n|V^\mu_{\mathrm{em}}|n\rangle = \langle p|V^{\mu V}|p\rangle. \tag{5.16}$$

Therefore, we find using eq. (5.15),

$$\begin{aligned}\langle n|V^\mu_{\mathrm{em}}|n\rangle &= \langle p|V^\mu_{\mathrm{em}}|p\rangle - \langle p|V^{\mu V}|p\rangle = \frac{1}{6}\left(\langle p|V^{\mu S}|p\rangle + 3\langle p|V^{\mu V}|p\rangle\right) - \langle p|V^{\mu V}|p\rangle \\ &= \frac{1}{6}\left(\langle p|V^{\mu S}|p\rangle - 3\langle p|V^{\mu V}|p\rangle\right),\end{aligned} \tag{5.17}$$

i.e., the same form as for the proton, but with the opposite sign in front of the isovector contribution. This means that it is sufficient to compute the separate matrix elements of the isovector and isoscalar vector currents only between proton states in order to obtain both $\langle p|V^\mu_{\mathrm{em}}|p\rangle$ and $\langle n|V^\mu_{\mathrm{em}}|n\rangle$.

5.2 Propagator calculation

The lattice calculation of the desired nucleon matrix elements requires the evaluation of observables which depend on the fermion fields, both via the nucleon interpolators eqs. (3.76) and (3.79) and via the vector current. As noted at the end of section 3.2.1, such expectation values can be expressed in terms of propagators by applying Wick's theorem eq. (2.39). Therefore, we need an efficient way to evaluate propagators, *i.e.*, inverses of the Dirac operator, on a given background gauge field.

Even if the Dirac matrix Q itself is local, as is the case for Wilson-Clover fermions, its inverse Q^{-1} is not. Hence, the full propagator is a dense and far too large matrix to keep it in computer memory. However, its entries are highly correlated because they all describe propagation on the same background gauge field [83]. This reduces the information content of the matrix and thus allows one to calculate only parts of it.

In particular, instead of all-to-all propagators $Q^{-1}(y,x)_{\beta k,\alpha j}$ between two arbitrary lattice sites y and x with Dirac indices β and α and color indices k and j, respectively, we consider point-to-all propagators $Q^{-1}(y,x_0)_{\beta k,\alpha_0 j_0}$, where the coordinate of the source x_0 and its indices α_0, j_0 are fixed. For this purpose, we introduce point sources

$$S_0^{(x_0,\alpha_0,j_0)}(x)_{\alpha j} = \delta_{x x_0}\delta_{\alpha\alpha_0}\delta_{j j_0}, \tag{5.18}$$

so that we can write [83]

$$Q^{-1}(y,x_0)_{\beta k,\alpha_0 j_0} = \sum_{x,\alpha,j} Q^{-1}(y,x)_{\beta k,\alpha j} S_0^{(x_0,\alpha_0,j_0)}(x)_{\alpha j}. \tag{5.19}$$

The point sources are placed at the position x_0 of the nucleon source interpolator, while the sink is projected to a definite momentum $\mathbf{p}'$. Since this Fourier transformation involves a sum over the lattice, what we need to compute is exactly the propagator from x_0 to any lattice site y, *i.e.*, the point-to-all propagator in eq. (5.19) for all 12 combinations of Dirac and color indices (α_0, j_0) at x_0.

One can make use of the translation invariance on a lattice with periodic boundary conditions by calculating the observables for several choices of x_0 on the same gauge configuration and averaging the results. A slightly more advanced strategy to implement this idea in practice will be presented in section 5.3.4. Moreover, we smear the point sources over a finite spatial range in order to increase the overlap of our interpolator with the nucleon ground state, a topic which will be discussed in more detail in section 5.3.3. We will come back to all the concrete expressions we evaluate in the following two sections.

According to eq. (5.19), the remaining problem consists in computing the action of the inverse Dirac operator on a quark source, $G = Q^{-1}S$. This is equivalent to solving the linear system $QG = S$ for G, which can be implemented taking advantage of the solver optimizations which were developed for the calculation of the fermion forces in the HMC (*cf.*, section 4.2.6).

5.3 The nucleon two-point function

The first important observable which we need to measure on the lattice in order to gain insight about the internal structure of the nucleon is its two-point function. We will start by discussing these two-point functions and their computation in the present section, before moving on to the more complicated case of the three-point functions, from which we will obtain the matrix elements eq. (5.1), in the next section. The two-point function is needed because it provides us with information about the nucleon mass on the employed ensembles as well as on *a priori* unknown overlap factors which also enter in the three-point functions, as we shall see.

On a given gauge configuration, we define the momentum-projected nucleon two-point function as follows,

$$C_2(\mathbf{p}';x,y_0) = \sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})}\,\mathrm{tr}[\Gamma^p N(y)\bar{N}(x)], \tag{5.20}$$

where the trace is taken over Dirac space, and N is the nucleon interpolator from eq. (3.76). According to eq. (3.78), the projection matrix Γ^p needs to be proportional to $1/2(1+\gamma_0)$ to ensure positive parity. We can also multiply it by an additional factor of $(1+i\gamma_5\gamma_j)$ to polarize the spin of the nucleon in the j-direction. Depending on the exact point where the two-point function is needed in the analysis, we adopt different choices of Γ^p, which we will define accordingly in chapter 6.

On a lattice with periodic boundary conditions (or in the bulk of a lattice with open boundary conditions), the expectation value of C_2 can only depend on the source-sink separation $t_{\mathrm{sep}} = y_0 - x_0$ due to translation invariance. Analogously, it should not depend on the spatial source position $\mathbf{x}$. Therefore, we place the nucleon creation interpolator on different source positions x and average the corresponding results. On ensembles with open boundary conditions in time, we need to take care that the time component x_0 lies sufficiently far in the bulk of the lattice, a topic which will be discussed in detail in section 5.3.5.

5.3.1 Spectral representation of the nucleon two-point function

In order to clarify the physical significance of the expectation value of the two-point function, we will now derive its spectral representation. Without loss of generality, we can set, for the purpose of this calculation, the source position x to zero and omit the average over sources. We will start by inserting a complete set of states $|n,s\rangle$ which are eigenstates of the Hamiltonian H according to $H|n,s\rangle = E_n|n,s\rangle$, *i.e.*, n labels the energy and s the spin. The completeness relation reads in our convention,

$$\mathbb{1} = \sum_{n,s} \frac{1}{2E_n} |n,s\rangle\langle n,s|\,. \tag{5.21}$$

Furthermore using the Heisenberg evolution in (Euclidean) time of the Hilbert space operator $\hat{N}$ corresponding to the interpolator N, we obtain,

$$\begin{aligned}
\langle C_2\rangle(\mathbf{p}';y_0) &= \sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot\mathbf{y}} \sum_{n,s} \frac{1}{2E_n}\,\mathrm{tr}\left[\Gamma^p \langle 0|\hat{N}(y)|n,s\rangle\langle n,s|\hat{N}^\dagger(0)|0\rangle\right] \\
&= \sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot\mathbf{y}} \sum_{n,s} \frac{1}{2E_n}\,\mathrm{tr}\left[\Gamma^p \left\langle 0\middle| e^{Hy_0}\hat{N}(0,\mathbf{y})e^{-Hy_0}\middle|n,s\right\rangle\langle n,s|\hat{N}^\dagger(0)|0\rangle\right] \\
&= \sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot\mathbf{y}} \sum_{n,s} e^{-E_n y_0}\frac{1}{2E_n}\,\mathrm{tr}\left[\Gamma^p \langle 0|\hat{N}(0,\mathbf{y})|n,s\rangle\langle n,s|\hat{N}^\dagger(0)|0\rangle\right],
\end{aligned} \tag{5.22}$$

where we have set the vacuum energy to zero, *i.e.*, we have defined $H|0\rangle = 0$.

We assume that the overlaps of $\hat{N}$ with the $|n,s\rangle$ states satisfy

$$\langle 0|\hat{N}(0,\mathbf{y})|n,s\rangle = e^{i\mathbf{p}_n\cdot\mathbf{y}} c_n(\mathbf{p}_n) u^s(\mathbf{p}_n), \tag{5.23}$$

$$\langle n,s|\hat{N}^\dagger(0,\mathbf{y})|0\rangle = e^{-i\mathbf{p}_n\cdot\mathbf{y}} c_n^*(\mathbf{p}_n)\bar{u}^s(\mathbf{p}_n), \tag{5.24}$$

where $\mathbf{p}_n$ is the total three-momentum corresponding to the state $|n,s\rangle$, and the overlap functions $c_n(\mathbf{p}_n)$ are *a priori* unknown. Inserting this into eq. (5.22) yields

$$\langle C_2\rangle(\mathbf{p}';y_0) = \sum_n \sum_{\mathbf{y}} e^{-i(\mathbf{p}'-\mathbf{p}_n)\cdot\mathbf{y}} e^{-E_n y_0} \frac{|c_n(\mathbf{p}_n)|^2}{2E_n}\,\mathrm{tr}\left[\Gamma^p \sum_s u^s(\mathbf{p}_n)\bar{u}^s(\mathbf{p}_n)\right]$$

$$= \sum_n \delta_{\mathbf{p}',\mathbf{p}_n} e^{-E_n y_0} \frac{|c_n(\mathbf{p}')|^2}{2E_n} \operatorname{tr}\left[\Gamma^p \sum_s u^s(\mathbf{p}')\bar{u}^s(\mathbf{p}')\right]. \tag{5.25}$$

The Dirac spinors fulfill in Minkowski space the following spin sum rule [64],

$$\sum_s u^s(\mathbf{p}^{(M)})\bar{u}^s(\mathbf{p}^{(M)}) = p_\mu^{(M)}\gamma^{(M)\mu} + m_N = E_\mathbf{p}\gamma^{(M)0} + p_k^{(M)}\gamma^{(M)k} + m_N. \tag{5.26}$$

Using eqs. (2.41) and (2.45), we can easily translate this to the Euclidean notation,

$$\sum_s u^s(\mathbf{p}^{(E)})\bar{u}^s(\mathbf{p}^{(E)}) = E_\mathbf{p}\gamma_0^{(E)} - ip_k^{(E)}\gamma_k^{(E)} + m_N. \tag{5.27}$$

Therefore, eq. (5.25) becomes

$$\langle C_2\rangle(\mathbf{p}';y_0) = \sum_n \delta_{\mathbf{p}',\mathbf{p}_n} e^{-E_n y_0} \frac{|c_n(\mathbf{p}')|^2}{2E_n} \operatorname{tr}\left[\Gamma^p(E_{\mathbf{p}'}\gamma_0 - ip_k\gamma_k + m_N)\right]. \tag{5.28}$$

As stated above, Γ^p is either $\Gamma_0^p = 1/2(1+\gamma_0)$ or $\Gamma_j^p = 1/2(1+\gamma_0)(1+i\gamma_5\gamma_j)$. The trace can be evaluated using the anticommutation relations of the Dirac matrices, the definition of γ_5 in eq. (2.46), and their well-known trace relations. Since the term proportional to $\gamma_5\gamma_j$ does not contribute in the trace of eq. (5.28), the result is the same for all choices of Γ^p,

$$\langle C_2\rangle(\mathbf{p}';y_0) = \sum_n \delta_{\mathbf{p}',\mathbf{p}_n} e^{-E_n y_0} \frac{|c_n(\mathbf{p}')|^2}{E_n}(E_{\mathbf{p}'} + m_N) \xrightarrow{y_0\gg 0} e^{-E_{\mathbf{p}'}y_0}|c_N(\mathbf{p}')|^2\left(1 + \frac{m_N}{E_{\mathbf{p}'}}\right). \tag{5.29}$$

For $y_0 \to \infty$, only the state with the lowest energy among those which couple to $\hat{N}$ (by having the same quantum numbers) and which have the correct momentum $\mathbf{p}'$, *i.e.*, among those with non-zero overlap factors $c_n(\mathbf{p}')$, contributes. By construction, this is the nucleon state with momentum $\mathbf{p}'$ and energy $E_{\mathbf{p}'} = \sqrt{\mathbf{p}'^2 + m_N^2}$. For finite y_0, eq. (5.29) describes a tower of exponentials with increasing energy to which all the aforementioned states contribute.

5.3.2 Implementation of the nucleon two-point function

To compute the expectation value of a fermionic observable $O[U,\bar{\psi},\psi] = \prod_{f=1}^{N_f} O^f[U,\bar{\psi}^f,\psi^f]$ on the lattice, we need to evaluate the following functional integral,

$$\langle O\rangle_F[U] = \frac{1}{Z_F[U]} \prod_{f=1}^{N_f} \int [d\bar{\psi}^f][d\psi^f]\, O^f[U,\bar{\psi}^f,\psi^f] \exp\left(-\sum_{x,y\in\Lambda} \bar{\psi}^f(x)Q^f[U](x,y)\psi^f(y)\right), \tag{5.30}$$

where $Z_F[U] = \prod_{f=1}^{N_f}\det(Q^f[U])$. Comparing to eq. (3.44), it follows that the full Lattice-QCD expectation value of O is given by

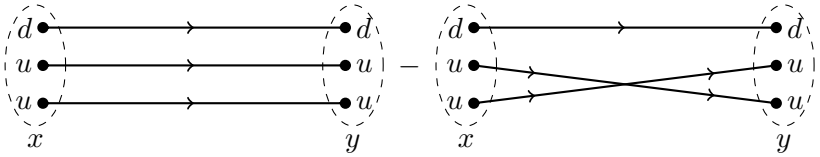
$$\begin{aligned}\langle O\rangle &= \frac{1}{Z_{\mathrm{LQCD}}}\int [dU]\, \exp(-S_G[U]) Z_F[U]\,\langle O\rangle_F[U] = \frac{1}{Z_{\mathrm{LQCD}}}\int [dU]\, \exp(-S_{\mathrm{eff}}[U])\,\langle O\rangle_F[U] \\ &= \langle\langle O\rangle_F\rangle_G.\end{aligned} \tag{5.31}$$

This can be computed from the usual gauge configurations which have been generated according to the distribution eq. (3.50) in the same way as for observables which only depend on the gauge field to start with.

In order to apply the above strategy to the nucleon two-point function, we plug eqs. (3.76) and (3.79) into eq. (5.20) and use Wick's theorem eq. (2.39) to evaluate the fermionic integral,

$$\begin{aligned}
\langle C_2 \rangle_F &= \left\langle \sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})} \operatorname{tr}[\Gamma^p N(y)\bar{N}(x)] \right\rangle_F \\
&= \Bigg\langle \sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})} \Gamma^p_{\beta\alpha} \epsilon_{jkl} \psi^u_{\alpha j}(y) \psi^u_{\gamma k}(y) C_{\gamma\delta}\gamma_{5,\delta\epsilon} \psi^d_{\epsilon l}(y) \\
&\qquad\qquad \epsilon_{mno} \bar{\psi}^d_{\rho m}(x) \gamma_{5,\rho\sigma} C^\dagger_{\sigma\mu} \bar{\psi}^u_{\mu n}(x) \bar{\psi}^u_{\beta o}(x) \Bigg\rangle_F \\
&= \sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})} \Gamma^p_{\beta\alpha} \epsilon_{jkl} \epsilon_{mno} C_{\gamma\delta}\gamma_{5,\delta\epsilon}\gamma_{5,\rho\sigma} C^\dagger_{\sigma\mu} S^d(y-x)_{\epsilon l,\rho m} \\
&\qquad [S^u(y-x)_{\alpha j,\beta o} S^u(y-x)_{\gamma k,\mu n} - S^u(y-x)_{\alpha j,\mu n} S^u(y-x)_{\gamma k,\beta o}]
\end{aligned} \tag{5.32}$$

$$= \quad - \tag{5.33}$$



We thus obtain two different Wick contractions corresponding to the two possible permutations of the u-quarks. Both diagrams are, however, of the same type and only differ by the indices of the u-quark propagators. For fixed source position x, we can hence express $\langle C_2 \rangle_F$ in terms of point-to-all propagators $S^f(y-x) \equiv (Q^f)^{-1}(y-x)$ which can be calculated with the point-source technique introduced in section 5.2.

As is evident from eq. (5.32), the translation invariance of the two-point function is already realized on the level of the fermionic expectation value. Therefore, we can average over different source positions x_j before taking the gauge average,

$$\langle C_2 \rangle (\mathbf{p}'; t_{\text{sep}}) = \langle \langle \langle C_2 \rangle_F \rangle_S \rangle_G (\mathbf{p}'; t_{\text{sep}}) = \left\langle \frac{1}{N_s} \sum_{j=1}^{N_s} \left\langle \sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x}_j)} \operatorname{tr}[\Gamma^p N(y)\bar{N}(x_j)] \right\rangle_F \right\rangle_G . \tag{5.34}$$

This means that we can even choose a different sef of source positions on each configuration, which slightly increases the effective statistics compared to the traditional setup of using a fixed set of source positions across all configurations. Irrespective of the boundary conditions, the sources can only be placed on edges of sub-blocks of the lattice because of the even-odd preconditioning (*cf.*, section 4.2.1) and SAP-based preconditioning (*cf.*, section 4.2.6) of the Dirac operator. In particular, the size of these sub-blocks needs to be compatible with the block size employed for the Schwarz preconditioning. On a given configuration, all sources have to be placed at the same "local" position within the sub-block, while in principle, this position is allowed to vary between configurations.

Equation (5.32) shows that the nucleon two-point function can be computed from the gauge average of a product of three fermion propagators. The variance of a Monte-Carlo estimate of such a correlation function is proportional to its fluctuation [152],

$$\sigma^2_{C_2} \propto \frac{1}{N} \left(\langle S^3 (S^\dagger)^3 \rangle - \langle S^3 \rangle^2 \right), \tag{5.35}$$

where N is the number of employed gauge configurations. Using a construction similar to the one in section 5.3.1, one can show that the first term in eq. (5.35) is, at large t_{sep}, dominated by an exponential falling off with the energy of the lowest state coupling to three quarks and

three antiquarks. As this is a state consisting of three pions rather than a nucleon-antinucleon pair, one expects [152]

$$\sigma_{C_2}(\mathbf{p}' = \mathbf{0}; t_{\text{sep}}) \propto \frac{1}{\sqrt{N}} \exp\left(-\frac{3}{2} M_\pi t_{\text{sep}}\right), \tag{5.36}$$

so that the signal-to-noise ratio is

$$\frac{\langle C_2 \rangle (\mathbf{p}' = \mathbf{0}; t_{\text{sep}})}{\sigma_{C_2}(\mathbf{p}' = \mathbf{0}; t_{\text{sep}})} \sim \sqrt{N} \exp\left[-\left(m_N - \frac{3}{2} M_\pi\right) t_{\text{sep}}\right]. \tag{5.37}$$

Since $m_N \gg {}^3\!/_2 M_\pi$ on all ensembles employed in this thesis (*cf.*, table 4.1), the signal-to-noise ratio falls off exponentially with increasing source-sink separation. This argument can easily be generalized to non-zero $\mathbf{p}'$, where the masses need to be replaced by the energies corresponding to $\mathbf{p}'$.

One hence faces a trade-off between reducing excited-state contamination and growing relative statistical noise, rendering the determination of appropriate fit ranges an involved task. Strategies to deal with this complication will be outlined in chapter 6; for the specific case of the two-point function and the determination of the nucleon mass from it see section 6.4.1.

5.3.3 Smearing

The aforementioned noise problem underlines the necessity to increase the overlap of the employed interpolators with the desired ground state, here the nucleon. This enables one to use smaller source-sink separations which have a better signal-to-noise ratio in the analysis. In the discussion so far, the source interpolator has been placed on a single space-time point x (*cf.*, eq. (5.18)). The nucleon is, however, not point-like, but its wave function has a finite spatial extent. A more realistic wave function can thus be obtained by smearing the nucleon source over a range of spatial positions. For this purpose, we define smeared quark fields as a linear combination of the original fields at several spatial positions in the vicinity of the point of interest x [83],

$$\tilde{\psi}^f(x_0, \mathbf{x})_{\alpha j} = \sum_{\mathbf{x}', \alpha', j'} \mathfrak{S}^{(x,\alpha,j)}(\mathbf{x}')_{\alpha' j'} \psi^f(x_0, \mathbf{x}')_{\alpha' j'}, \tag{5.38}$$

$$\tilde{\bar{\psi}}^f(x_0, \mathbf{x})_{\alpha j} = \sum_{\mathbf{x}', \alpha', j'} \bar{\mathfrak{S}}^{(x,\alpha,j)}(\mathbf{x}')_{\alpha' j'} \bar{\psi}^f(x_0, \mathbf{x}')_{\alpha' j'}. \tag{5.39}$$

To compute observables, the quark fields are integrated out according to eq. (5.30), yielding quark propagators instead. Hence, we need to derive the smeared quark propagator [83],

$$\begin{aligned}
(\tilde{Q}^f)^{-1}(y, x)_{\beta k, \alpha j} &= \left\langle \tilde{\psi}^f(y)_{\beta k} \tilde{\bar{\psi}}^f(x)_{\alpha j} \right\rangle_F \\
&= \left\langle \sum_{\substack{\mathbf{x}', \alpha', j' \\ \mathbf{y}', \beta', k'}} \mathfrak{S}^{(y,\beta,k)}(\mathbf{y}')_{\beta' k'} \psi^f(y_0, \mathbf{y}')_{\beta' k'} \bar{\mathfrak{S}}^{(x,\alpha,j)}(\mathbf{x}')_{\alpha' j'} \bar{\psi}^f(x_0, \mathbf{x}')_{\alpha' j'} \right\rangle_F \\
&= \sum_{\substack{\mathbf{x}', \alpha', j' \\ \mathbf{y}', \beta', k'}} \mathfrak{S}^{(y,\beta,k)}(\mathbf{y}')_{\beta' k'} (Q^f)^{-1}(y', x')_{\beta' k', \alpha' j'} \bar{\mathfrak{S}}^{(x,\alpha,j)}(\mathbf{x}')_{\alpha' j'}.
\end{aligned} \tag{5.40}$$

Here, we have defined $x'_0 = x_0$ and $y'_0 = y_0$.

In the numerical calculation, the construction of this quark propagator from a smeared source to a smeared sink can be implemented in three steps [153]. First, one applies the smearing

function $\bar{\mathfrak{S}}$ to a point source eq. (5.18) (which is nothing else than choosing a point x and indices α, j). Next, one solves the Dirac equation

$$\sum_{y,\beta,k} Q^f(x',y)_{\alpha' j',\beta k} (\underset{\sim}{Q}^f)^{-1}(y,x)_{\beta k,\alpha j} = \bar{\mathfrak{S}}^{(x,\alpha,j)}(\mathbf{x}')_{\alpha' j'} \tag{5.41}$$

to obtain the propagator from the smeared source to a point-like sink,

$$(\underset{\sim}{Q}^f)^{-1}(y,x)_{\beta k,\alpha j} = \sum_{\mathbf{x}',\alpha',j'} (Q^f)^{-1}(y,x')_{\beta k,\alpha' j'} \bar{\mathfrak{S}}^{(x,\alpha,j)}(\mathbf{x}')_{\alpha' j'}. \tag{5.42}$$

Lastly, the smearing function $\mathfrak{S}$ is applied to $(\underset{\sim}{Q}^f)^{-1}$,

$$\sum_{\mathbf{y}',\beta',k'} \mathfrak{S}^{(y,\beta,k)}(\mathbf{y}')_{\beta' k'} (\underset{\sim}{Q}^f)^{-1}(y',x)_{\beta' k',\alpha j} = (\tilde{Q}^f)^{-1}(y,x)_{\beta k,\alpha j}, \tag{5.43}$$

thus yielding eq. (5.40).

In this construction, the source and sink smearing functions $\bar{\mathfrak{S}}$ and $\mathfrak{S}$, respectively, are still arbitrary, allowing many possible smearing schemes. An intuitively well motivated shape for the smearing is Gaussian. Moreover, it is advantageous to define gauge-covariant smearing functions, because otherwise the gauge in the considered timeslice(s) would need to be fixed [83]. Therefore, it has been suggested in Ref. [153] to generate the smeared source iteratively,

$$\begin{aligned} \tilde{S}^{(x,\alpha,j)}(x')_{\alpha' j'} &= \sum_{x'',\alpha'',j''} \bar{\mathfrak{S}}^{(x',\alpha',j')}(\mathbf{x}'')_{\alpha'' j''} S_0^{(x,\alpha,j)}(x'')_{\alpha'' j''} \\ &= \sum_{x'',\alpha'',j''} \left[(1+\kappa_G H_G)^{N_G}\right](x',x'')_{\alpha' j',\alpha'' j''} S_0^{(x,\alpha,j)}(x'')_{\alpha'' j''}, \end{aligned} \tag{5.44}$$

i.e., the operator $(1+\kappa_G H_G)$ is applied N_G times, starting from a point source. Here, H_G is the spatial hopping matrix,

$$H_G(x,x')_{\alpha j,\alpha' j'} = \delta_{\alpha\alpha'} \sum_{\mu=1}^{3} \left[U_\mu(x)_{jj'} \delta_{x',x+a\hat{\mu}} + U_\mu^\dagger(x-a\hat{\mu})_{jj'} \delta_{x',x-a\hat{\mu}} \right]. \tag{5.45}$$

This means that the smearing only acts on the spatial and color indices, but leaves the Dirac components of the spinors invariant. Moreover, the different lattice sites are appropriately connected by gauge links, rendering the smearing functions gauge covariant. The sources resulting from this so-called Wuppertal smearing were found to be rotationally invariant and approximately Gaussian in shape [153].

The smearing prescription eq. (5.44) has two free parameters: κ_G and the number of iterations N_G. These can be tuned to achieve the desired profile. To quantify the spatial size of the smeared source, one introduces the smearing radius [154],

$$r_{\mathrm{sm}}^2(x) = \frac{\sum_{x'} |x'-x|^2 \|\tilde{S}^{(x)}(x')\|^2}{\sum_{x'} \|\tilde{S}^{(x)}(x')\|^2}, \tag{5.46}$$

where x is the location of the source. This definition depends on the background gauge field, because the latter enters H_G and thus the smeared source $\tilde{S}$. We choose κ_G and N_G such that a smearing radius of $r_{\mathrm{sm}} \approx 0.5$ fm is realized on average. Consequently, our sources have roughly the size of the proton, which leads to an improved overlap with the proton ground state, as illustrated in fig. 5.2 below. Of course, this argument is somewhat heuristic and many other smearing schemes and/or parameter choices are conceivable.

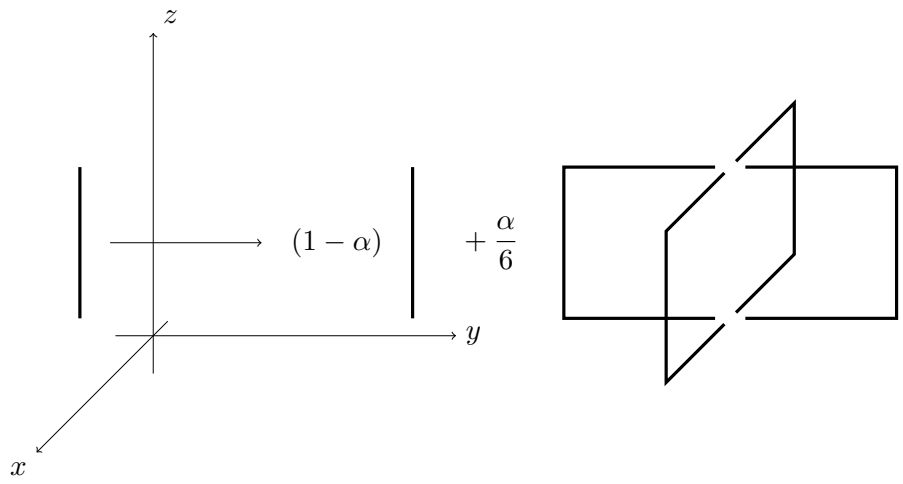


Figure 5.1: Illustration of the APE smearing procedure for a gauge link in z-direction (adapted from Ref. [155])

To further improve the signal of correlation functions, it is desirable to reduce the influence of the unphysical short-distance fluctuations of the gauge field on the interpolators. For this purpose, one can replace the gauge fields entering the (fermion) smearing operator via eq. (5.45) by smeared gauge fields [155]. The basic idea is to replace each gauge link entering eq. (5.45) by a local average over short paths connecting the link's endpoints. This is a gauge-covariant procedure, so that no gauge fixing is required [83]. Concretely, we use the version of Ref. [83] of the APE smearing originally proposed in Ref. [155]. However, we do not modify the time-like links and only average over purely space-like paths, as in Ref. [155]. Thus, the average is taken over the original link and the four neighboring space-like staples (plaquettes missing one link) connecting its endpoints,

$$U_j(x) \to P_{\mathrm{SU}(3)} \left[(1-\alpha) U_j(x) + \frac{\alpha}{6} \sum_{\eta=\pm} \sum_{k \neq j} C_{jk}^{\eta}(x) \right], \tag{5.47}$$

$$C_{jk}^{+}(x) = U_k(x) U_j(x + a\hat{k}) U_k^{\dagger}(x + a\hat{j}), \tag{5.48}$$

$$C_{jk}^{-}(x) = U_k^{\dagger}(x - a\hat{k}) U_j(x - a\hat{k}) U_k(x - a\hat{k} + a\hat{j}). \tag{5.49}$$

This is illustrated in fig. 5.1. The right-hand side of eq. (5.47) needs to be projected back to SU(3), for which an approximate algorithm is used in the numerical implementation.

The procedure can be iterated, and the number of iterations as well as the parameter α can be adjusted. We use $\alpha = 0.6$ and 15 iterations of eq. (5.47) for all values of the gauge coupling β. Therefore, the gauge-field smearing radius vanishes in the continuum limit, so that the continuum long-distance correlation functions are not altered [83].

The effect of the smearing on the two-point function is illustrated in fig. 5.2 for the ensemble N451. In the left panel, the expectation value $\langle C_2 \rangle (\mathbf{0}; t_{\mathrm{sep}})$ is shown, while in the right panel, one can see the effective energy

$$a E_{\mathrm{eff}}(\mathbf{p}'; t_{\mathrm{sep}}) = \ln \frac{\langle C_2 \rangle (\mathbf{p}'; t_{\mathrm{sep}})}{\langle C_2 \rangle (\mathbf{p}', t_{\mathrm{sep}} + a)}, \tag{5.50}$$

again at $\mathbf{p}' = \mathbf{0}$. According to eq. (5.29), the effective energy converges for $t_{\mathrm{sep}} \to \infty$ to the ground-state energy, here the nucleon mass. Therefore, we call $E_{\mathrm{eff}}(\mathbf{p}' = \mathbf{0})$ also effective

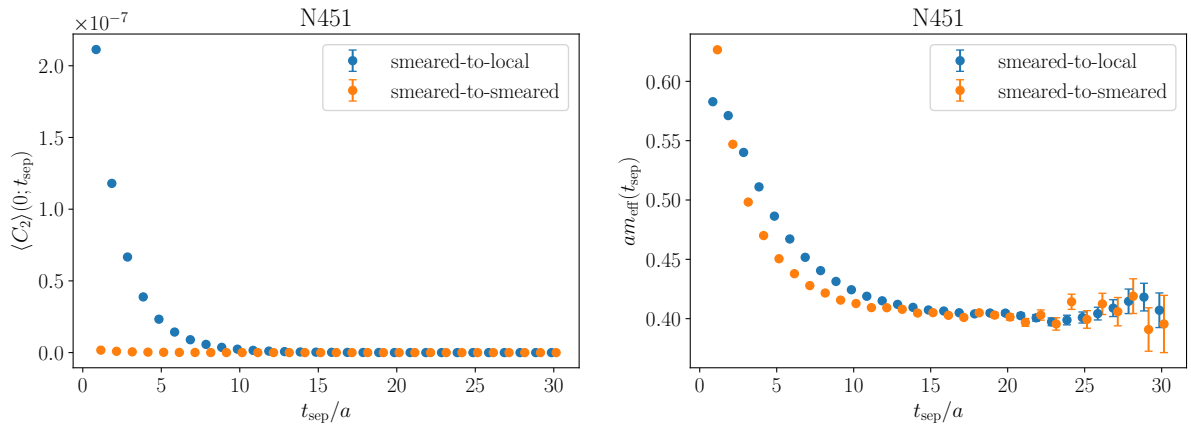


Figure 5.2: Comparison of the two-point function calculated from smeared-to-local and smeared-to-smeared propagators on the ensemble N451. The points are horizontally displaced for better visibility.

mass m_{eff}. In both panels of fig. 5.2, we compare the two-point functions calculated from smeared-to-local and smeared-to-smeared propagators, *i.e.*, the only difference between the two data sets is in the smearing of the sink. One can see that smearing the sink in addition to the source further helps in reducing the excited-state contamination at small source-sink separations, so that a plateau is reached earlier. For both data sets, the exponential error growth expected from eq. (5.37) and a corresponding increase in the statistical fluctuations is clearly visible in the effective mass at large source-sink separations (*cf.*, right panel of fig. 5.2).

Moreover, smearing both the source and the sink makes the procedure symmetric, thus ensuring that the overlap factors c_n in eqs. (5.23) and (5.24) are indeed identical (up to complex conjugation). This guarantees that the spectral representation of the two-point function is given by a sum of exponentials with positive coefficients (*cf.*, eq. (5.29)). It has been worked out in the appendix of Ref. [53] that applying a spatial smearing on a fixed timeslice (the one where the source is located) leads to spectral representations of smeared-to-smeared two-point as well as three-point functions which are formally identical to the non-smeared case (*cf.*, eqs. (5.28) and (5.54)), even if the objects u, $\bar{u}$ in eqs. (5.23) and (5.24) cannot strictly be interpreted as Dirac spinors any longer.

5.3.4 The truncated-solver method

The most expensive part of computing fermionic observables on the lattice is usually the inversion of the Dirac operator required to calculate propagators. A significant speedup (in terms of required computing time per achievable precision) can be reached by noticing that the Dirac equation does not need to be solved with high precision for all sources. This is the basic idea of the truncated-solver method [156–158]: One uses a relaxed stopping criterion for the majority of sources. To correct for the bias introduced in this way, one treats a small number of sources additionally with high-precision solves. Thus, we decompose the source average of an observable O as follows,[3]

$$\langle O \rangle_S = \frac{1}{N_{\text{LP}}} \sum_{j=1}^{N_{\text{LP}}} O^{\text{LP}}(x_j) + \frac{1}{N_{\text{HP}}} \sum_{j=1}^{N_{\text{HP}}} [O^{\text{HP}}(x_j) - O^{\text{LP}}(x_j)], \tag{5.51}$$

[3]Here, we assume that in O the fermion fields have already been integrated out.

where HP denotes high-precision and LP low-precision solves, and the x_j label the corresponding source positions. For the first N_{HP} sources, both low- and high-precision solves are performed, so that the bias introduced by the former can be estimated. This bias is then subtracted from the average over all sources with low-precision solves, yielding an unbiased and improved estimator for O. In practice, we observe that the bias term is almost negligibly small for our observables. Therefore, we can choose $N_{\mathrm{LP}} \gg N_{\mathrm{HP}}$ (*cf.*, table 5.1), so that the overall computing time is dominated by the low-precision solves.

5.3.5 Additional measurements on ensembles with open boundary conditions

On ensembles with periodic boundary conditions, the sources can be distributed randomly on the entire lattice volume, subject to the constraints pointed out below eq. (5.34). On ensembles with open boundary conditions in time, by contrast, the translation invariance is only (approximately) restored in the bulk of the lattice. Therefore, one needs to take care that boundary effects do not significantly influence the results computed on such ensembles. For this purpose, not only the temporal coordinate x_0 of the source itself needs to be placed sufficiently far away from the boundaries, but also that of the sink, which is located at $y_0 = x_0 + t_{\mathrm{sep}}$. The minimal distance which the source needs to keep from the boundaries thus depends on the source-sink separations which one aims to consider.

In the nucleon-structure efforts of the Mainz lattice group, a sufficient suppression of boundary effects has originally been ensured by placing all sources on a single timeslice for ensembles with open boundary conditions. However, the independent statistics one can achieve on a single timeslice is limited and has indeed quickly been exhausted on most ensembles. To overcome this limitation, it was suggested in Ref. [67] to put sources on additional timeslices, while closely monitoring the effect of the boundaries on the observables. The increased precision of the two-point function will become of particular importance for the computation of the disconnected part of the three-point function (*cf.*, section 5.4.4). We have continued the program of performing additional measurements of the two-point function, which was also started in Ref. [67], so that these now exist on all ensembles with open boundary conditions used in this thesis except S201. For the ensembles which we have treated from scratch (N101, E300 and J303), the timeslices are chosen randomly in a predefined bulk, and this is done independently for each configuration, as explained below eq. (5.34).

Since the boundary state has the quantum numbers of the vacuum [118], we expect that the observable which is most affected by boundary effects is the vacuum quark loop with a scalar insertion,

$$L_{\mathbb{1}}^{f}(\mathbf{q}) = \sum_{\mathbf{z}} e^{i\mathbf{q}\cdot\mathbf{z}} \operatorname{tr}\left[\langle \bar{\psi}^{f}(z)\psi^{f}(z)\rangle_{F}\right] = -\sum_{\mathbf{z}} e^{i\mathbf{q}\cdot\mathbf{z}} \operatorname{tr}[S^{f}(z,z)]. \tag{5.52}$$

We will explain in detail how such observables containing a propagator from a point to itself can be computed in section 5.4.4. If translation invariance holds, the quark propagator S^f only depends on the difference of its two arguments, so that the whole expression in eq. (5.52) is independent of z_0. In fig. 5.3, the gauge average of $L_{\mathbb{1}}^{\mathrm{light}}(\mathbf{0})$ is plotted on the open-boundary ensemble C101 as a function of the timeslice on which it is evaluated.[4] One can clearly observe the broken translation invariance in the vicinity of the temporal boundaries at least up to $t/a, (T-t)/a \approx 22$, which corresponds in physical units to about 1.9 fm.

Also the nucleon two-point function is influenced by the open boundary conditions. To facilitate the subsequent discussion, we start by remarking that in principle also the backward-propagating nucleon can be used in the analysis of the two-point function if the sign of γ_0 in the

[4] For the quark loop, in contrast to the nucleon two-point function, the u- and d-quarks do not need to be distinguished for the Wick contractions because only one flavor contributes to it.

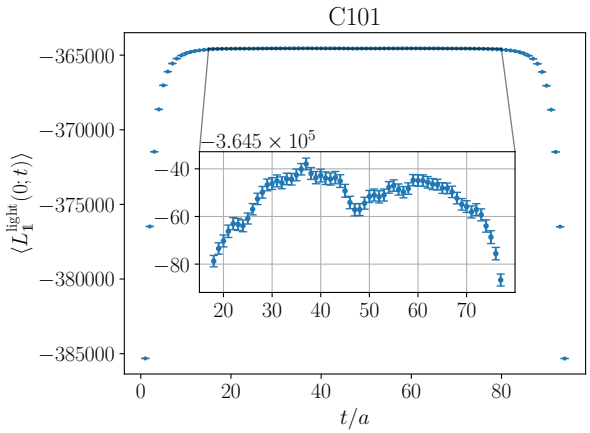


Figure 5.3: Gauge average of the scalar light-quark loop at zero momentum on the ensemble C101 as a function of the timeslice on which it is evaluated

parity-projection operator is flipped (*cf.*, eq. (3.78)). In the following, we will always use the forward- as well as the backward-propagating nucleon and average over the two corresponding results for all sources, unless they are excluded by one of the criteria which we will introduce below.

The influence of the boundary effects on the two-point function on C101 is illustrated in fig. 5.4, where we compare the results calculated from different subsets of sources according to their distance to the temporal boundaries. In particular, we show the results obtained imposing three different minimal distances which the source of the nucleon propagating towards the boundary needs to keep from the latter: 2.0 fm (which is fulfilled for all computed sources), 2.7 fm (which corresponds to omitting the sources placed on the largest (smallest) timeslice for the forward (backward) propagating nucleon) and 4.1 fm. If all computed sources are employed (blue points), the effective mass in the right panel of fig. 5.4 shows an artificial upward bend starting at $t_{\text{sep}} \approx 12a \approx 1.0$ fm. Such a statistically significant upward curvature should not exist according to eq. (5.29) which predicts the effective mass to be monotonically falling towards its asymptotic value, and can thus only be explained by boundary effects. Noting that for the most extreme sources entering this data set, the sink is only 1.0 fm away from the boundary for $t_{\text{sep}} = 1.0$ fm and propagation towards the boundary, this does not come as a surprise.

Omitting this most extreme case when determining the two-point function already removes most of this artefact in the effective mass, as can be seen from the orange set of points in the right panel of fig. 5.4. Taking a more detailed look at the two-point function itself in the left panel of fig. 5.4, one observes that for source-sink separations of about 2.0 fm or larger, the orange points start to deviate significantly from the green ones. The latter have been obtained with a much more conservative restriction on the distance to the boundary. One has to keep in mind, however, that imposing a stricter criterion for the distance to the boundary corresponds to employing fewer of the computed sources in the analysis, leading to larger errors. This can also be seen in the right panel of fig. 5.4.

Nevertheless, also prompted by the observations made for the scalar quark loop in fig. 5.3, we impose a minimal distance of 4.0 fm which the source of the nucleon propagating towards the boundary has to keep from the latter on ensembles with open boundary conditions.[5] Only

[5] On C101, the same set of sources actually fulfills a minimum distance of 4.1 fm, hence the slightly different number in the legend of fig. 5.4.

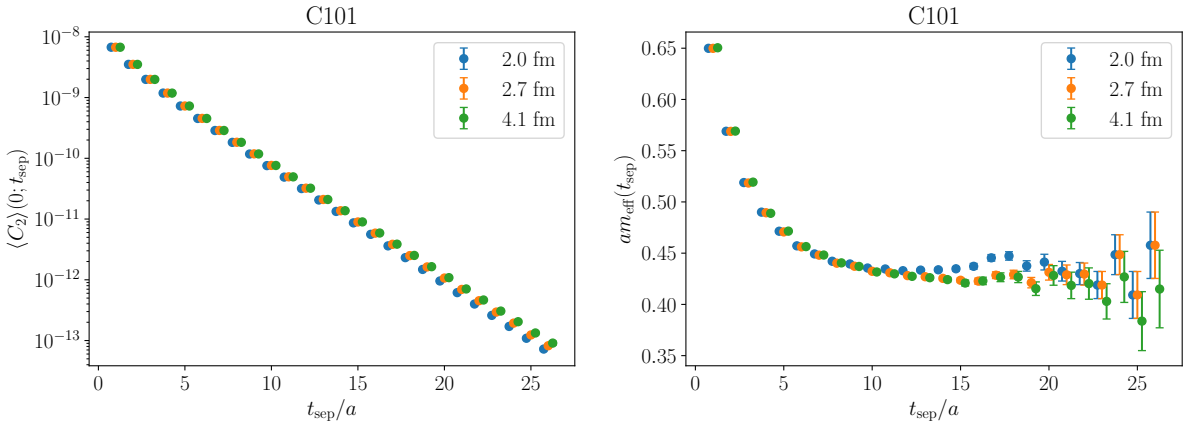


Figure 5.4: Comparison of the two-point function on the ensemble C101 calculated with different choices for the minimal distance which the source of the nucleon propagating towards the boundary needs to keep from the latter. The points are horizontally displaced for better visibility.

sources satisfying this condition are permitted for our determination of the nucleon mass (*cf.*, section 6.4.1), allowing us to safely consider source-sink separations of up to 2.0 fm for this purpose. On S201, the single timeslice on which all sources are placed is less than 4 fm from the lower temporal boundary of the lattice. Hence, we can only use the forward propagating nucleon on this ensemble.

The resulting total number of high- and low-precision measurements per ensemble are listed in table 5.1. Here, we have counted sources on which both propagation directions are employed twice, and such on which only one propagation direction is used only once. The fact that the aggregated number of measurements does not divide the number of configurations on some ensembles means that the number of measurements per configuration is not the same for all configurations on these ensembles. Duplicate source positions on individual configurations can arise due to the iterative setup of the runs and the random selection of the source positions. Such duplicates are not considered in the analysis, *i.e.*, each unique source position is only used once, leading to the aforementioned effect.

5.4 The nucleon three-point function

The nucleon two-point function introduced in the previous section can be used to determine the energy of the nucleon for a definite momentum from the lattice. In order to gain insight on its internal structure, a more advanced observable is required. For the determination of the electromagnetic form factors, matrix elements of the vector current between nucleon states are needed, as discussed in section 5.1. Hence, we insert a vector current between the two nucleon interpolators. This yields a three-point function, which we again project to a definite sink momentum and momentum transfer,

$$C_{3,V_\mu}(\mathbf{p}', \mathbf{q}; x, y_0, z_0) = \sum_{\mathbf{y},\mathbf{z}} e^{i\mathbf{q}\cdot\mathbf{z}} e^{-i\mathbf{p}'\cdot\mathbf{y}} e^{i(\mathbf{p}'-\mathbf{q})\cdot\mathbf{x}} \operatorname{tr}[\Gamma^p N(y) V_\mu(z) \bar{N}(x)]. \tag{5.53}$$

Similar to the case of C_2, translation invariance implies that the expectation value of C_{3,V_μ} only depends on the source-sink separation $t_{\mathrm{sep}} = y_0 - x_0$ and the operator insertion time relative to the source, $t = z_0 - x_0$, and not on the source position x.

Table 5.1: Number of gauge configurations and aggregated number of high-precision (HP) and low-precision (LP) measurements used for the determination of the nucleon mass

ID	N_{cfg}	$N_{\mathrm{meas,HP}}$	$N_{\mathrm{meas,LP}}$
C101	994	9939	301 581
N101	1588	5947	715 624
H105	1024	33 777	313 867
D450	498	7968	127 488
N451	1010	16 160	258 560
E250	796	12 736	407 552
D200	998	9980	303 194
N200	1708	15 372	460 672
S201	2092	4181	96 279
E300	569	1986	255 203
J303	1073	5729	212 829

5.4.1 Spectral representation of the nucleon three-point function

In order to quantify the connection of the nucleon three-point function in eq. (5.53) with the matrix element in eq. (5.1), we will now derive the spectral representation of the expectation value of the three-point function. We proceed analogously to section 5.3.1, *i.e.*, set the source position x to zero, insert complete sets of states $|n', s'\rangle$ and $|n, s\rangle$ according to eq. (5.21), and use the Heisenberg evolution in (Euclidean) time of the appearing operators. Moreover, we insert the overlaps of $\hat{N}$ with the $|n, s\rangle$ states (*cf.*, eqs. (5.23) and (5.24)) as well as the definition of the vector current matrix element (*cf.*, eq. (5.1)). Here, we need to use the Euclidean convention for the Fourier factor according to eq. (5.12). In the last step, we plug in the Euclidean spin sum rule eq. (5.27). Thus, we obtain,

$$
\begin{aligned}
\langle C_{3,V_\mu} \rangle (\mathbf{p}', \mathbf{q}; y_0, z_0) &= \sum_{\mathbf{y},\mathbf{z}} e^{i\mathbf{q}\cdot\mathbf{z}} e^{-i\mathbf{p}'\cdot\mathbf{y}} \sum_{n',s'} \sum_{n,s} \frac{1}{4E_{n'}E_n} \operatorname{tr}[\Gamma^p \langle 0|\hat{N}(y)|n', s'\rangle \langle n', s'|V_\mu(z)|n, s\rangle \\
&\qquad \langle n, s|\hat{N}^\dagger(0)|0\rangle] \\
&= \sum_{\mathbf{y},\mathbf{z}} e^{i\mathbf{q}\cdot\mathbf{z}} e^{-i\mathbf{p}'\cdot\mathbf{y}} \sum_{n',s'} \sum_{n,s} \frac{1}{4E_{n'}E_n} \operatorname{tr}\Big[\Gamma^p \Big\langle 0 \Big| e^{Hy_0} \hat{N}(0,\mathbf{y}) e^{-Hy_0} \Big| n', s' \Big\rangle \\
&\qquad \Big\langle n', s' \Big| e^{Hz_0} V_\mu(0,\mathbf{z}) e^{-Hz_0} \Big| n, s \Big\rangle \\
&\qquad \langle n, s|\hat{N}^\dagger(0)|0\rangle\Big] \\
&= \sum_{\mathbf{y},\mathbf{z}} e^{i\mathbf{q}\cdot\mathbf{z}} e^{-i\mathbf{p}'\cdot\mathbf{y}} \sum_{n',s'} \sum_{n,s} e^{-E_{n'}y_0} e^{-(E_n - E_{n'})z_0} \frac{1}{4E_{n'}E_n} \\
&\qquad \operatorname{tr}[\Gamma^p \langle 0|\hat{N}(0,\mathbf{y})|n', s'\rangle \langle n', s'|V_\mu(0,\mathbf{z})|n, s\rangle \\
&\qquad \langle n, s|\hat{N}^\dagger(0)|0\rangle] \\
&= \sum_{n',n} \sum_{\mathbf{y},\mathbf{z}} e^{i[\mathbf{q}-(\mathbf{p}_{n'}-\mathbf{p}_n)]\cdot\mathbf{z}} e^{-i(\mathbf{p}'-\mathbf{p}_{n'})\cdot\mathbf{y}} e^{-E_{n'}y_0} e^{-(E_n-E_{n'})z_0} \frac{c_{n'}(\mathbf{p}_{n'}) c_n^*(\mathbf{p}_n)}{4E_{n'}E_n} \\
&\qquad \operatorname{tr}\left[\Gamma^p \sum_{s',s} u^{s'}(\mathbf{p}_{n'}) \bar{u}^{s'}(\mathbf{p}_{n'}) \mathcal{V}_\mu(\mathbf{p}_{n'} - \mathbf{p}_n) u^s(\mathbf{p}_n) \bar{u}^s(\mathbf{p}_n)\right]
\end{aligned}
$$

$$= \sum_{n',n} \delta_{\mathbf{p}',\mathbf{p}_{n'}} \delta_{\mathbf{q},\mathbf{p}_{n'}-\mathbf{p}_n} e^{-E_{n'}y_0} e^{-(E_n - E_{n'})z_0} \frac{c_{n'}(\mathbf{p}')c_n^*(\mathbf{p}'-\mathbf{q})}{4E_{n'}E_n} \operatorname{tr}[\Gamma^p(E_{\mathbf{p}'}\gamma_0 - ip'_j\gamma_j + m_N)\mathcal{V}_\mu(\mathbf{q})(E_{\mathbf{p}'-\mathbf{q}}\gamma_0 - i(p'-q)_k\gamma_k + m_N)] \tag{5.54}$$

$$\xrightarrow{y_0,z_0\gg 0} e^{-E_{\mathbf{p}'}y_0} e^{-(E_{\mathbf{p}'-\mathbf{q}} - E_{\mathbf{p}'})z_0} \frac{c_N(\mathbf{p}')c_N^*(\mathbf{p}'-\mathbf{q})}{4E_{\mathbf{p}'}E_{\mathbf{p}'-\mathbf{q}}} \operatorname{tr}[\Gamma^p(E_{\mathbf{p}'}\gamma_0 - ip'_j\gamma_j + m_N)\mathcal{V}_\mu(\mathbf{q})(E_{\mathbf{p}'-\mathbf{q}}\gamma_0 - i(p'-q)_k\gamma_k + m_N)] \tag{5.55}$$

The two exponentials in eq. (5.55) can be written as $e^{-E_{\mathbf{p}'-\mathbf{q}}z_0} e^{-E_{\mathbf{p}'}(y_0-z_0)}$. As a consequence, the asymptotic limit is reached if the operator insertion (at time z_0) has a very large distance both from the source (at time 0) and from the sink (at time $y_0 > z_0$). In this limit, only the nucleon state with momentum $\mathbf{p}'$ and $\mathbf{p}' - \mathbf{q}$ contributes to the two spectral sums over n' and n, respectively.

For the remainder of this thesis, we will only consider three-point functions with the sink momentum $\mathbf{p}'$ set to zero, so that the source is implicitly projected to momentum $-\mathbf{q}$. This choice greatly simplifies the evaluation of the trace in eq. (5.54),

$$\mathcal{T}_{V_\mu}^{\Gamma^p}(\mathbf{p}' = \mathbf{0}, \mathbf{q}) = m_N \operatorname{tr}[\Gamma^p(1+\gamma_0)\mathcal{V}_\mu(\mathbf{q})(E_\mathbf{q}\gamma_0 + iq_k\gamma_k + m_N)]. \tag{5.56}$$

For this purpose, we need to plug in the photon-nucleon vertex function in Euclidean notation (*cf.*, eqs. (5.13) and (5.14)) as well as the definition of the electric and magnetic Sachs form factors (*cf.*, eq. (5.5)). Here, we note that for our kinematics

$$Q^2 = (q^{(E)})^2 = -(q^{(M)})^2 = -(m_N - E_\mathbf{q})^2 + \mathbf{q}^2 = -m_N^2 - E_\mathbf{q}^2 + 2m_N E_\mathbf{q} + \mathbf{q}^2 = 2m_N(E_\mathbf{q} - m_N). \tag{5.57}$$

If we choose the unpolarized projector $\Gamma_0^p = 1/2(1+\gamma_0)$, we find by using $(1+\gamma_0)^2 = 2(1+\gamma_0)$,

$$\begin{aligned}
\mathcal{T}_{V_0}^{\Gamma_0^p}(\mathbf{p}' = \mathbf{0}, \mathbf{q}) &= m_N \operatorname{tr}\left[(1+\gamma_0)\left(\gamma_0 F_1(Q^2) + i\gamma_0\gamma_k \frac{q_k}{2m_N} F_2(Q^2)\right)(E_\mathbf{q}\gamma_0 + iq_l\gamma_l + m_N)\right] \\
&= 4m_N\left[(m_N + E_\mathbf{q})F_1(Q^2) - \frac{\mathbf{q}^2}{2m_N}F_2(Q^2)\right] \\
&= 4m_N(m_N + E_\mathbf{q})\left[F_1(Q^2) - \frac{E_\mathbf{q} - m_N}{2m_N}F_2(Q^2)\right] = 4m_N(m_N + E_\mathbf{q})G_E(Q^2),
\end{aligned} \tag{5.58}$$

and for the spatial components,

$$\begin{aligned}
\mathcal{T}_{V_j}^{\Gamma_0^p}(\mathbf{p}' = \mathbf{0}, \mathbf{q}) &= m_N \operatorname{tr}\left\{(1+\gamma_0)\left[\gamma_j F_1(Q^2) + \left(\gamma_0\gamma_j(m_N - E_\mathbf{q}) + i\gamma_j\sum_{k\neq j}\gamma_k q_k\right)\frac{F_2(Q^2)}{2m_N}\right]\right. \\
&\qquad \left.(E_\mathbf{q}\gamma_0 + iq_l\gamma_l + m_N)\right\} \\
&= 4im_N q_j\left[F_1(Q^2) + \frac{m_N - E_\mathbf{q}}{2m_N}F_2(Q^2)\right] = 4im_N q_j G_E(Q^2).
\end{aligned} \tag{5.59}$$

Alternatively, we can choose a projector which polarizes the spin of the nucleon in the j-direction, *i.e.*, $\Gamma_j^p = 1/2(1+\gamma_0)(1+i\gamma_5\gamma_j)$. For the 0- and j-components of the vector current, we notice that the term proportional to $\gamma_5\gamma_j$ does not contribute to the trace. We therefore obtain the

same results as with Γ_0^p, *i.e.*, $\mathcal{T}_{V_0}^{\Gamma_j^p} = \mathcal{T}_{V_0}^{\Gamma_0^p}$ and $\mathcal{T}_{V_j}^{\Gamma_j^p} = \mathcal{T}_{V_j}^{\Gamma_0^p}$. For the spatial components of the vector current which are orthogonal to the polarization direction, the γ_5-term does contribute, and we find,

$$\begin{aligned}
\mathcal{T}_{V_{k\neq j}}^{\Gamma_j^p}(\mathbf{p}'=\mathbf{0},\mathbf{q}) &= m_N \operatorname{tr}\left\{(1+\gamma_0)(1+i\gamma_5\gamma_j)\left[\gamma_k F_1(Q^2) + \left(\gamma_0\gamma_k(m_N - E_\mathbf{q}) + i\gamma_k\sum_{l\neq k}\gamma_l p_l\right)\right.\right. \\
&\qquad \left.\left.\frac{F_2(Q^2)}{2m_N}\right](E_\mathbf{q}\gamma_0 + iq_m\gamma_m + m_N)\right\} \\
&= 4m_N\left[(iq_k + \epsilon_{jkl}q_l)F_1(Q^2) + (iq_k + \epsilon_{jkl}q_l)\frac{m_N - E_\mathbf{q}}{2m_N}F_2(Q^2)\right. \\
&\qquad \left. + \epsilon_{jkl}q_l\frac{E_\mathbf{q} + m_N}{2m_N}F_2(Q^2)\right] \\
&= 4m_N\left[iq_k\left(F_1(Q^2) + \frac{m_N - E_\mathbf{q}}{2m_N}F_2(Q^2)\right) + \epsilon_{jkl}q_l(F_1(Q^2) + F_2(Q^2))\right] \\
&= 4m_N[iq_k G_E(Q^2) + \epsilon_{jkl}q_l G_M(Q^2)]. \qquad (5.60)
\end{aligned}$$

These are therefore the only components from which the magnetic form factor can be extracted. For the extraction of the electric form factor, on the other hand, there are in principle several options. We will come back to this topic in section 6.5, where we will also discuss what the most advantageous choice is.

The expressions for the traces in the more complicated case $\mathbf{p}' \neq \mathbf{0}$ have been derived in Ref. [67]. Since they are not used in this thesis, we do not repeat them here.

5.4.2 Implementation of the nucleon three-point function

In order to implement the calculation of the nucleon three-point function, we have to evaluate its fermionic expectation value using Wick's theorem similar to section 5.3.2. For this purpose, the quark flavor of the inserted vector current needs to be specified. To obtain the matrix elements of the electromagnetic vector current with contributions up to the strange quark between proton and neutron states separately, the isovector $(u-d)$ and isoscalar $(u+d-2s)$ flavor combinations of the vector current are required (*cf.*, eqs. (5.15) and (5.17)). We therefore start by deriving the contractions for the u-, d- and s-quark currents separately and then build their aforementioned linear combinations. For the u-quark current, we obtain,

$$\begin{aligned}
\langle C_{3,V_\mu}^u\rangle_F &= \left\langle \sum_{\mathbf{y},\mathbf{z}} e^{i\mathbf{q}\cdot(\mathbf{z}-\mathbf{x})}e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})}\operatorname{tr}[\Gamma^p N(y)V_\mu^u(z)\bar{N}(x)]\right\rangle_F \\
&= \left\langle \sum_{\mathbf{y},\mathbf{z}} e^{i\mathbf{q}\cdot(\mathbf{z}-\mathbf{x})}e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})}\Gamma^p_{\beta\alpha}\epsilon_{jkl}\psi^u_{\alpha j}(y)\psi^u_{\gamma k}(y)C_{\gamma\delta}\gamma_{5,\delta\epsilon}\psi^d_{\epsilon l}(y)\bar{\psi}^u_{\eta m}(z)\gamma_{\mu,\eta\rho}\psi^u_{\rho m}(z)\right. \\
&\qquad \left.\epsilon_{nop}\bar{\psi}^d_{\sigma n}(x)\gamma_{5,\sigma\tau}C^\dagger_{\tau\varphi}\bar{\psi}^u_{\varphi o}(x)\bar{\psi}^u_{\beta p}(x)\right\rangle_F
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{\mathbf{y},\mathbf{z}} e^{i\mathbf{q}\cdot(\mathbf{z}-\mathbf{x})} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})} \Gamma^p_{\beta\alpha} \epsilon_{jkl}\epsilon_{nop} C_{\gamma\delta}\gamma_{5,\delta\epsilon}\gamma_{5,\sigma\tau} C^\dagger_{\tau\varphi}\gamma_{\mu,\eta\rho} S^d(y-x)_{\epsilon l,\sigma n} \qquad (5.61)\\
&\quad [S^u(y-z)_{\alpha j,\eta m} S^u(y-x)_{\gamma k,\varphi o} S^u(z-x)_{\rho m,\beta p}\\
&\quad + S^u(y-x)_{\alpha j,\beta p} S^u(y-z)_{\gamma k,\eta m} S^u(z-x)_{\rho m,\varphi o}\\
&\quad + S^u(y-x)_{\alpha j,\varphi o} S^u(y-x)_{\gamma k,\beta p} S^u(z-z)_{\rho m,\eta m}\\
&\quad - S^u(y-x)_{\alpha j,\varphi o} S^u(y-z)_{\gamma k,\eta m} S^u(z-x)_{\rho m,\beta p}\\
&\quad - S^u(y-z)_{\alpha j,\eta m} S^u(y-x)_{\gamma k,\beta p} S^u(z-x)_{\rho m,\varphi o}\\
&\quad - S^u(y-x)_{\alpha j,\beta p} S^u(y-x)_{\gamma k,\varphi o} S^u(z-z)_{\rho m,\eta m}]\\
&= \langle C^u_{3,V_\mu}\rangle^{\text{conn}}_F + \langle C^u_{3,V_\mu}\rangle^{\text{disc}}_F , \qquad (5.62)
\end{aligned}
$$

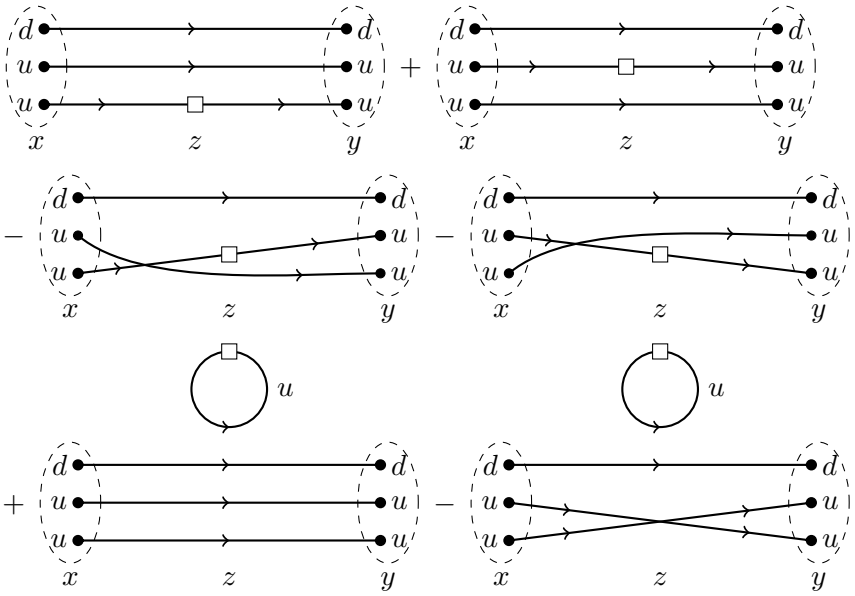
where $\langle C^u_{3,V_\mu}\rangle^{\text{disc}}_F = e^{-i\mathbf{q}\cdot\mathbf{x}} \langle C_2\rangle_F L^u_{V_\mu}$ and the vector-current vacuum quark loop is defined in analogy to eq. (5.52),

$$
L^f_{V_\mu}(\mathbf{q}) = \sum_{\mathbf{z}} e^{i\mathbf{q}\cdot\mathbf{z}} \operatorname{tr}\left[\langle \bar\psi^f(z)\gamma_\mu\psi^f(z)\rangle_F\right] = -\sum_{\mathbf{z}} e^{i\mathbf{q}\cdot\mathbf{z}} \operatorname{tr}[\gamma_\mu S^f(z,z)] \qquad (5.63)
$$

$$
= \qquad f \qquad (5.64)
$$

We have thus found two different types of Wick contractions, quark-connected and quark-disconnected ones. Diagrammatically, they can be represented as

$\langle C^u_{3,V_\mu}\rangle_F =$

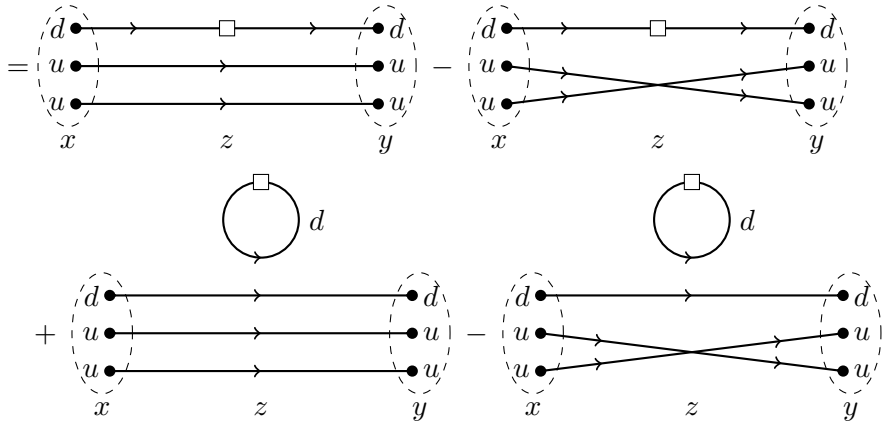


(5.65)

where the box indicates the position which is contracted with γ_μ.

For the d-quark vector current, the Wick contractions are

$$
\begin{aligned}
\langle C^d_{3,V_\mu}\rangle_F = \Big\langle \sum_{\mathbf{y},\mathbf{z}} & e^{i\mathbf{q}\cdot(\mathbf{z}-\mathbf{x})} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})} \Gamma^p_{\beta\alpha}\epsilon_{jkl}\psi^u_{\alpha j}(y)\psi^u_{\gamma k}(y) C_{\gamma\delta}\gamma_{5,\delta\epsilon}\psi^d_{\epsilon l}(y)\bar\psi^d_{\eta m}(z)\gamma_{\mu,\eta\rho}\psi^d_{\rho m}(z)\\
&\epsilon_{nop}\bar\psi^d_{\sigma n}(x)\gamma_{5,\sigma\tau} C^\dagger_{\tau\varphi}\bar\psi^u_{\varphi o}(x)\bar\psi^u_{\beta p}(x)\Big\rangle_F\\
= -\sum_{\mathbf{y},\mathbf{z}} & e^{i\mathbf{q}\cdot(\mathbf{z}-\mathbf{x})} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})} \Gamma^p_{\beta\alpha}\epsilon_{jkl}\epsilon_{nop} C_{\gamma\delta}\gamma_{5,\delta\epsilon}\gamma_{5,\sigma\tau} C^\dagger_{\tau\varphi}\gamma_{\mu,\eta\rho} \qquad (5.66)\\
&[S^d(y-z)_{\epsilon l,\eta m} S^d(z-x)_{\rho m,\sigma n} - S^d(y-x)_{\epsilon l,\sigma n} S^d(z-z)_{\rho m,\eta m}]\\
&[S^u(y-x)_{\alpha j,\varphi o} S^u(y-x)_{\gamma k,\beta p} - S^u(y-x)_{\alpha j,\beta p} S^u(y-x)_{\gamma k,\varphi o}]
\end{aligned}
$$

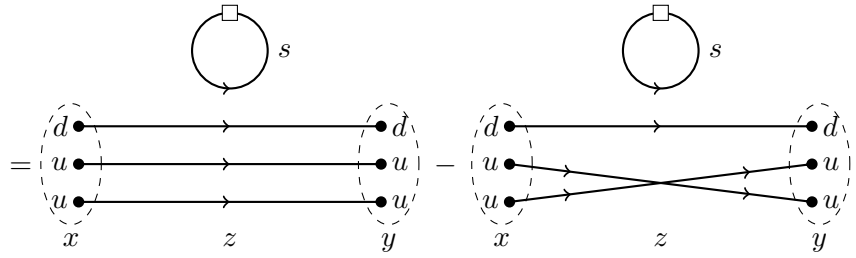


(5.67)

i.e., we again obtain quark-connected as well as quark-disconnected contractions. The disconnected contractions for V_μ^d are the same as for V_μ^u, but with the u-quark in the loop replaced by a d-quark. Since we assume SU(2) isospin symmetry (*i.e.*, $m_u = m_d$) in our calculations, the u- and d-quark propagators are the same, so that also the corresponding vacuum quark loops are identical. Consequently, the quark-disconnected diagrams cancel in the isovector $(u-d)$ combination. For the isoscalar combination, this is however not the case.

The Wick contractions for the s-quark vector current are purely quark-disconnected because our nucleon interpolator only contains u- and d-quark fields,

$$\begin{aligned}\langle C^s_{3,V_\mu}\rangle_F = \Bigg\langle \sum_{\mathbf{y},\mathbf{z}} & e^{i\mathbf{q}\cdot(\mathbf{z}-\mathbf{x})} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})} \Gamma^p_{\beta\alpha} \epsilon_{jkl} \psi^u_{\alpha j}(y) \psi^u_{\gamma k}(y) C_{\gamma\delta}\gamma_{5,\delta\epsilon} \psi^d_{\epsilon l}(y) \bar{\psi}^s_{\eta m}(z) \gamma_{\mu,\eta\rho} \psi^s_{\rho m}(z) \\ & \epsilon_{nop} \bar{\psi}^d_{\sigma n}(x) \gamma_{5,\sigma\tau} C^\dagger_{\tau\varphi} \bar{\psi}^u_{\varphi o}(x) \bar{\psi}^u_{\beta p}(x) \Bigg\rangle_F \\ = \sum_{\mathbf{y},\mathbf{z}} & e^{i\mathbf{q}\cdot(\mathbf{z}-\mathbf{x})} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})} \Gamma^p_{\beta\alpha} \epsilon_{jkl} \epsilon_{nop} C_{\gamma\delta}\gamma_{5,\delta\epsilon}\gamma_{5,\sigma\tau} C^\dagger_{\tau\varphi}\gamma_{\mu,\eta\rho} S^d(y-x)_{\epsilon l,\sigma n} \\ & S^s(z-z)_{\rho m,\eta m} [S^u(y-x)_{\alpha j,\varphi o} S^u(y-x)_{\gamma k,\beta p} - S^u(y-x)_{\alpha j,\beta p} S^u(y-x)_{\gamma k,\varphi o}]\end{aligned} \tag{5.68}$$



(5.69)

In the numerical implementation, we compute the isovector combination $\langle C^V_{3,V_\mu}\rangle_F = \langle C^u_{3,V_\mu}\rangle^{\text{conn}}_F - \langle C^d_{3,V_\mu}\rangle^{\text{conn}}_F$ and the connected part of the isoscalar combination $\langle C^S_{3,V_\mu}\rangle^{\text{conn}}_F = \langle C^u_{3,V_\mu}\rangle^{\text{conn}}_F + \langle C^d_{3,V_\mu}\rangle^{\text{conn}}_F$, as well as the light- and strange-quark loops. The full isoscalar combination can then be obtained as

$$\begin{aligned}\langle C^S_{3,V_\mu}\rangle_F &= \langle C^S_{3,V_\mu}\rangle^{\text{conn}}_F + \langle C^u_{3,V_\mu}\rangle^{\text{disc}}_F + \langle C^d_{3,V_\mu}\rangle^{\text{disc}}_F - 2\langle C^s_{3,V_\mu}\rangle^{\text{disc}}_F \\ &= \langle C^S_{3,V_\mu}\rangle^{\text{conn}}_F + 2e^{-i\mathbf{q}\cdot\mathbf{x}} \langle C_2\rangle_F (L^{\text{light}}_{V_\mu} - L^s_{V_\mu}).\end{aligned} \tag{5.70}$$

However, we build this linear combination of the connected and disconnected contributions only on the level of the effective form factors. This allows us to treat the two pieces slightly

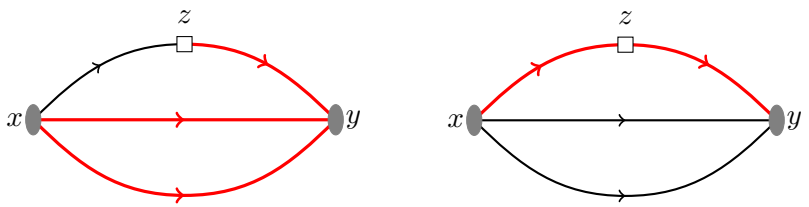


Figure 5.5: Diagrammatic illustration of the extended-propagator method with a fixed sink (left) or a fixed operator (right). The extended propagators are marked in red [159, 160].

differently for the extraction of the effective form factors, *e.g.*, by using different projectors. For further details, we refer to section 6.5 below.

In this thesis, we do not employ the local discretization of the vector current, which we have used for the derivation of the Wick contractions above in order to keep the required notation at a minimum. Instead, we employ the conserved vector current introduced in section 3.4.3, which comprises the same quark fields, but split over neighboring lattice points and with a slightly different Dirac structure. Therefore, the Wick contractions are exactly the same as for the local vector current; only the spatial indices of the propagators and the Dirac structure need to be modified according to eqs. (3.84) and (3.85). A simple example to demonstrate the effect of using a point-split current is the vacuum quark loop,

$$L^f_{\Gamma,+\hat{\mu}}(\mathbf{q}) = \sum_{\mathbf{z}} e^{i\mathbf{q}\cdot\mathbf{z}} \operatorname{tr}\left[\langle \bar{\psi}^f(z+a\hat{\mu})\Gamma U^\dagger_\mu(z)\psi^f(z)\rangle_F\right] = -\sum_{\mathbf{z}} e^{i\mathbf{q}\cdot\mathbf{z}} \operatorname{tr}[\Gamma U^\dagger_\mu(z) S^f(z, z+a\hat{\mu})], \tag{5.71}$$

$$L^f_{\Gamma,-\hat{\mu}}(\mathbf{q}) = \sum_{\mathbf{z}} e^{i\mathbf{q}\cdot\mathbf{z}} \operatorname{tr}\left[\langle \bar{\psi}^f(z)\Gamma U_\mu(z)\psi^f(z+a\hat{\mu})\rangle_F\right] = -\sum_{\mathbf{z}} e^{i\mathbf{q}\cdot\mathbf{z}} \operatorname{tr}[\Gamma U_\mu(z) S^f(z+a\hat{\mu}, z)]. \tag{5.72}$$

We obtain the same Wick contraction as for the local case (*cf.*, eq. (5.63)), but with one of the spatial indices of the propagator shifted by one lattice unit and an additional gauge link in the trace to maintain gauge covariance.

In the following two subsections, we will deal with the implementation details of the connected part and of the quark loops required for the disconnected part.

5.4.3 The connected part: extended-propagator method

The quark-connected Wick contractions contain a propagator $S(y-z)$ from the point of the operator insertion to the sink. This means that in principle both ends of the propagator are free (not fixed by the chosen source position). The computation of such an all-to-all propagator is a numerically very demanding task and should be avoided wherever possible. Therefore, it has been suggested to fix either the sink [159] or the operator insertion [160]. In both cases, one constructs a so-called generalized or extended propagator by solving the Dirac equation with (a combination of) the usual point-to-all propagators as a source. The extended propagators for the two methods are sketched in fig. 5.5.

If the sink is fixed, one has to choose a source-sink separation $t_{\text{sep}} = y_0 - x_0$, a projector Γ^p and a sink momentum $\mathbf{p}'$ before computing the extended propagator. Hence, an additional inversion is required for each choice of these three variables. On the other hand, any operator, operator insertion time $t = z_0 - x_0$ and momentum transfer $\mathbf{q}$ can be obtained from the same inversion. For the fixed-operator method, the situation is exactly the opposite. To generate the

data used in this thesis, the fixed-sink method has been employed because we are interested in many momentum transfers $\mathbf{q}$. Additionally, this has the advantage of enabling the computation of multiple operators without any significant extra cost, thus producing data for a number of physics projects simultaneously. Because the three-point functions obtained using different projectors are not independent, averaging over the polarization directions does typically not yield an increase in effective statistics which is commensurate with the extra cost it requires in the fixed-sink method. Therefore, it has been decided to employ only the z-polarized projector $\Gamma_3^p = 1/2(1+\gamma_0)(1+i\gamma_5\gamma_3)$ for the computation of the connected part. Moreover, we fix $\mathbf{p}' = \mathbf{0}$, as already mentioned in section 5.4.1.

For the fixed-sink method, the calculation of the three-point function proceeds in the following steps: First, for a given source position x, one computes the point-to-all propagator $S(y-x)$ to any lattice point y. This is the same type of propagator needed for the calculation of the two-point function. Next, one constructs the extended propagators. For a d-quark insertion, it is defined as [159]

$$\begin{aligned}\Sigma^d(\mathbf{p}';x,y_0,z)_{\sigma n,\eta m} = -\sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})}\Gamma^p_{\beta\alpha}\epsilon_{jkl}\epsilon_{nop}C_{\gamma\delta}\gamma_{5,\delta\epsilon}\gamma_{5,\sigma\tau}C^\dagger_{\tau\varphi}S^d(y-z)_{el,\eta m}\\ [S^u(y-x)_{\alpha j,\varphi o}S^u(y-x)_{\gamma k,\beta p} - S^u(y-x)_{\alpha j,\beta p}S^u(y-x)_{\gamma k,\varphi o}],\end{aligned} \tag{5.73}$$

and for a u-quark insertion as

$$\begin{aligned}\Sigma^u(\mathbf{p}';x,y_0,z)_{\sigma n,\eta m} = -\sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})}\Gamma^p_{\beta\alpha}\epsilon_{jkl}\epsilon_{nop}C_{\gamma\delta}\gamma_{5,\delta\epsilon}\gamma_{5,\chi\tau}S^d(y-x)_{el,\chi p}\\ (C^\dagger_{\tau\varphi}\delta_{\sigma\beta} + C^\dagger_{\tau\sigma}\delta_{\varphi\beta})[S^u(y-z)_{\alpha j,\eta m}S^u(y-x)_{\gamma k,\varphi o}\\ - S^u(y-x)_{\alpha j,\varphi o}S^u(y-z)_{\gamma k,\eta m}]\end{aligned} \tag{5.74}$$

where in the case of SU(2) isospin symmetry $S^u \equiv S^d$, so that the flavor index of the propagators is actually redundant. The extended propagators satisfy equations which can be solved with the same numerical techniques as those satisfied by the usual point-to-all propagators. For Σ^d, for example, one needs to solve [159]

$$\begin{aligned}\sum_z \Sigma^d(\mathbf{p}';x,y_0,z)_{\sigma n,\eta m}Q(z,b)_{\eta m,\xi q} = -e^{-i\mathbf{p}'\cdot(\mathbf{b}-\mathbf{x})}\Gamma^p_{\beta\alpha}\epsilon_{jkq}\epsilon_{nop}C_{\gamma\delta}\gamma_{5,\delta\xi}\gamma_{5,\sigma\tau}C^\dagger_{\tau\varphi}\delta_{y_0,b_0}\\ [S^u(b-x)_{\alpha j,\varphi o}S^u(b-x)_{\gamma k,\beta p}\\ - S^u(b-x)_{\alpha j,\beta p}S^u(b-x)_{\gamma k,\varphi o}]\end{aligned} \tag{5.75}$$

Once the point-to-all propagators are computed, the right-hand side of eq. (5.75) is known, and the equation can be solved with the methods introduced in section 5.2. For Σ^u, the calculation is analogous. In the last step, one contracts the extended propagator with the desired operator insertion (in our case γ_μ) and a (standard) propagator from the source to this insertion point. Thus, one obtains the connected part of the nucleon three-point function with a u- or d-quark vector current insertion [159],

$$\langle C^{u,d}_{3,V_\mu}(\mathbf{p}',\mathbf{q};x,y_0,z_0)\rangle^{\text{conn}}_F = \sum_{\mathbf{z}} e^{i\mathbf{q}\cdot(\mathbf{z}-\mathbf{x})}\,\mathrm{tr}[\Sigma^{u,d}(\mathbf{p}';x,y_0,z)\gamma_\mu S^{u,d}(z-x)]. \tag{5.76}$$

This shows again that while each combination of Γ^p, $\mathbf{p}'$ and y_0 requires a new inversion, the inserted Dirac operator, z_0 and $\mathbf{q}$ only need to be known for the final contraction in eq. (5.76), which is computationally cheap compared to the inversions.

Because of the exponentially dropping signal-to-noise ratio, it is desirable to perform more and more measurements for increasing source-sink separation. Hence, we employ iterative statistics for the different source-sink separations: with rising t_{sep}, the number of sources used for the computation of the connected part is increased. In practice, for every (or every second) step in t_{sep}, we employ all sources from the previous step plus the same number of newly selected sources, thus doubling the number of (low-precision) measurements. In this way, one can achieve that the effective statistics as a function of t_{sep} more closely resembles a constant instead of showing the aforementioned exponential decay.

The resulting number of high- and low-precision measurements per configuration are listed in table 5.2 together with the available values of t_{sep}/a. The issue of duplicate source positions which we encountered for the additional measurements of the two-point function did not occur here, so that the aggregated number of measurements is always the product of the number of configurations and the number of measurements per configuration, *i.e.*, $N_{\mathrm{meas}} = N_{\mathrm{cfg}} n_{\mathrm{meas}}$. The only exception is the last iteration on the ensemble N200, where 16 instead of 12 sources are used on 583 out of the total 1708 configurations. For illustration, we show the resulting aggregated numbers for the largest value of t_{sep} in the last two columns of table 5.2.

Table 5.2: Number of gauge configurations, and of high-precision (HP) and low-precision (LP) measurements per configuration used for the computation of the connected part. The superscripts label the source-sink separations t_{sep}/a to which the given numbers correspond. A " : 2" in the superscript signifies that only every second value in the given range of t_{sep}/a has been computed, *i.e.*, only the even ones. The last two columns show the aggregated number of measurements used for the largest value of t_{sep}.

ID	N_{cfg}	$n_{\mathrm{meas,HP}}$	$n_{\mathrm{meas,LP}}$	$N^{\mathrm{max}}_{\mathrm{meas,HP}}$	$N^{\mathrm{max}}_{\mathrm{meas,LP}}$
C101	1988	$1^{(4-17)}$	$2^{(4)}$, $4^{(5,6)}$, $8^{(7,8)}$, $16^{(9,10)}$, $32^{(11-17)}$	1988	63 616
N101	1588	$1^{(4-17)}$	$2^{(4)}$, $4^{(5,6)}$, $8^{(7,8)}$, $16^{(9,10)}$, $32^{(11-17)}$	1588	50 816
H105	1024	$1^{(4-10)}$, $4^{(11-17)}$	$2^{(4)}$, $4^{(5,6)}$, $8^{(7,8)}$, $16^{(9,10)}$, $48^{(11-17)}$	4096	49 152
D450	498	$1^{(4-14)}$, $2^{(15,16)}$, $4^{(17,18)}$, $8^{(19,20)}$	$0^{(4-6)}$, $2^{(7,8)}$, $4^{(9,10)}$, $8^{(11,12)}$, $16^{(13,14)}$, $32^{(15,16)}$, $64^{(17,18)}$, $128^{(19,20)}$	3984	63 744
N451	1010	$1^{(4-14:2)}$, $2^{(16)}$, $4^{(18)}$, $8^{(20)}$	$0^{(4,6)}$, $2^{(8)}$, $4^{(10)}$, $8^{(12)}$, $16^{(14)}$, $32^{(16)}$, $64^{(18)}$, $128^{(20)}$	8080	129 280
E250	398	$1^{(4-16:2)}$, $2^{(18)}$, $4^{(20)}$, $8^{(22)}$	$0^{(4,6)}$, $2^{(8)}$, $4^{(10)}$, $8^{(12)}$, $16^{(14)}$, $32^{(16)}$, $64^{(18)}$, $128^{(20)}$, $256^{(22)}$	3184	101 888
D200	1996	$1^{(4-22:2)}$	$0^{(4,6)}$, $2^{(8)}$, $4^{(10)}$, $8^{(12)}$, $16^{(14)}$, $32^{(16-22:2)}$	1996	63 872
N200	1708	$1^{(4-22:2)}$	$2^{(4,6)}$, $4^{(8,10)}$, $8^{(12,14)}$, $12-16^{(16-22:2)}$	1708	22 828
S201	2092	$1^{(4-22:2)}$	$4^{(4,6)}$, $8^{(8,10)}$, $16^{(12,14)}$, $32^{(16-22:2)}$	2092	66 944
E300	569	$1^{(4-28:2)}$	$2^{(4,6)}$, $4^{(8,10)}$, $8^{(12,14)}$, $16^{(16,18)}$, $32^{(20-28:2)}$	569	18 208
J303	1073	$1^{(4-28:2)}$	$0^{(4,6)}$, $2^{(8,10)}$, $4^{(12,14)}$, $8^{(16,18)}$, $16^{(20-28:2)}$	1073	17 168

5.4.4 The disconnected part

The quark-disconnected contribution to the nucleon three-point function is given by the product of the two-point function and the quark loop of a single flavor. As the latter contains a propagator from a point to itself (*cf.*, eq. (5.63)), the calculation of all-to-all propagators cannot be avoided in this case. Many strategies have been and are still being developed to tackle this numerically very challenging problem. In the following, we will give an overview over the methods which have been applied to the data used in this thesis.

Stocastic estimation

The full inverse of the Dirac operator is a dense and extremely large matrix, rendering it impossible to compute and store in memory all its entries. One therefore introduces random auxiliary fields (stochastic volume sources) $\eta_{j,\alpha a}(x)$ defined so that all their cumulants vanish except [161, 162]

$$\langle \eta^*_{j,\alpha a}(x)\eta_{k,\beta b}(y)\rangle_\eta = \delta_{jk}\delta_{\alpha\beta}\delta_{ab}\delta_{xy}, \tag{5.77}$$

where j, k label independent source fields, α, β are Dirac spinor indices as usual, and a, b are color indices (in the fundamental representation). We can construct a simple stochastic noise estimator for the quark loop in eq. (5.63) by averaging over N_s independent source fields [161, 162],

$$\hat{L}^f_{V_\mu}(\mathbf{q}) = -\frac{1}{N_s}\sum_{j=1}^{N_s}\sum_{\mathbf{z}}\sum_{x} e^{i\mathbf{q}\cdot\mathbf{z}}\eta^\dagger_j(z)\gamma_\mu S^f(z,x)\eta_j(x). \tag{5.78}$$

Given eq. (5.77), it is easy to see that $\langle\hat{L}\rangle_\eta = L$. In principle, different distributions of the noise sources can be used as long as eq. (5.77) is fulfilled. The perhaps most straightforward idea is to use a Gaussian distribution with mean 0 and a covariance matrix given by eq. (5.77) [161]. In Ref. [163], it has been suggested to use Z_2 noise to suppress the statistical deviation from unity of the diagonal elements of eq. (5.77) evaluated at finite N_s. We use U(1) noise whenever we need a stochastic estimator, which has also been found to yield smaller variances on the stochastically estimated quark loops than Gaussian noise [162].

Regardless of the exact distribution which the noise sources follow, the random-noise contribution to the error of the stochastic estimator in eq. (5.78) is expected to scale like $1/\sqrt{N_s}$ [162]. This provides a much too slow convergence rate in N_s to saturate gauge noise (*i.e.*, make the random-noise contribution to the error of eq. (5.78) subdominant compared to the gauge-noise contribution) at a feasible numerical cost [164]. For this reason, more advanced strategies are required in order to speed up the convergence of the stochastic estimators.

Split-even estimators or the one-end trick

If only the difference between the loops of two non-degenerate flavors is required, one can rewrite the corresponding trace [165],

$$\begin{aligned} L^f_{V_\mu}(\mathbf{q}) - L^{f'}_{V_\mu}(\mathbf{q}) &= -\sum_{\mathbf{z}} e^{i\mathbf{q}\cdot\mathbf{z}}\,\mathrm{tr}\{\gamma_\mu[S^f(z,z) - S^{f'}(z,z)]\} \\ &= -\sum_{\mathbf{z}}\sum_{x,y} e^{i\mathbf{q}\cdot\mathbf{z}}\,\mathrm{tr}\{\gamma_\mu S^f(z,x)[Q^{f'}(x,y) - Q^f(x,y)]S^{f'}(y,z)\} \\ &= -(m^{f'} - m^f)\sum_{\mathbf{z}}\sum_{x} e^{i\mathbf{q}\cdot\mathbf{z}}\,\mathrm{tr}[\gamma_\mu S^f(z,x)S^{f'}(x,z)], \end{aligned} \tag{5.79}$$

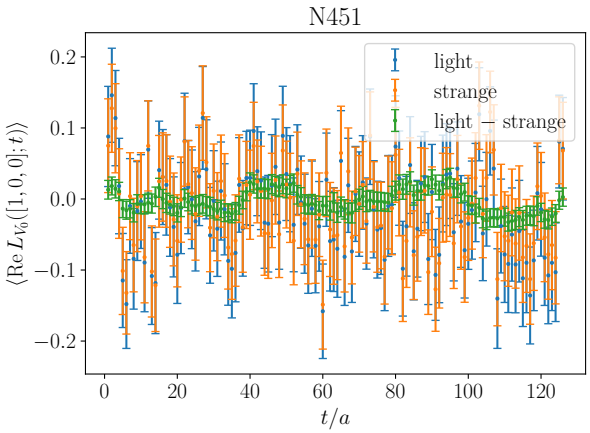


Figure 5.6: Gauge average of the real part of the 0-component of the vector quark loops at the first non-vanishing momentum in x-direction on the ensemble N451 as a function of the timeslice on which it is evaluated. Shown are the light- and strange-quark loops individually as well as their correlated difference which enters the isoscalar combination according to eq. (5.70).

where we have used that the Dirac operators of two flavors only differ in the mass term which is diagonal in all indices (*cf.*, eq. (3.31)). A naive stochastic estimator of eq. (5.79) would thus read

$$-(m^{f'}-m^{f})\frac{1}{N_s}\sum_{j=1}^{N_s}\sum_{\mathbf{z}}\sum_{x,y} e^{i\mathbf{q}\cdot\mathbf{z}}\eta_j^\dagger(z)\gamma_\mu S^f(z,x)S^{f'}(x,y)\eta_j(y). \tag{5.80}$$

A significantly improved estimator can be obtained by making use of the cyclic property of the trace to insert both noise sources at the "one end" of the product in the trace, where the identity (one) matrix in Dirac space is inserted. This is the so-called one-end trick [165, 166], and the corresponding estimator (which is called split-even estimator in Ref. [162]) reads

$$-(m^{f'}-m^{f})\frac{1}{N_s}\sum_{j=1}^{N_s}\sum_{\mathbf{z}}\sum_{x,y} e^{i\mathbf{q}\cdot\mathbf{z}}\eta_j^\dagger(x)S^{f'}(x,z)\gamma_\mu S^f(z,y)\eta_j(y). \tag{5.81}$$

It has been shown in Ref. [162] that the variance of this estimator is significantly smaller than that of eq. (5.80). This can be intuitively understood to be due to the fact that in eq. (5.81), both noise sources are ultraviolet filtered by a quark propagator [162].

Using the one-end trick, it is possible to saturate gauge noise for the vector-current loops with only $\mathcal{O}(100)$ noise sources [162, 164] (see below for the exact number of noise sources which we employ for the different flavors). For the isoscalar observables considered in this thesis, only the difference between the light- and strange-quark loops is needed according to eq. (5.70). Therefore, we can make use of the high-precision calculation based on the one-end trick, as well as of the partial cancellation of correlated (gauge) noise in the difference. The latter effect is illustrated in fig. 5.6, where we compare the gauge average of $L_{V_0}^{\mathrm{light}}(2\pi/L\hat{x})$, of $L_{V_0}^{s}(2\pi/L\hat{x})$ and of the difference of the two. It is clear to see that the fluctuations as well as the corresponding errors are much smaller in the correlated difference of the two flavors than in either of them individually.

The generalized hopping-parameter expansion

With the one-end trick, only the difference between the loops of two flavors can be computed. In order to calculate the loops of a single, individual quark flavor, an absolute scale needs to be introduced by computing one flavor explicitly with a different method. If this reference flavor is chosen to be very heavy, the hopping-parameter expansion provides an efficient way to reduce the random-noise contribution to the quark loops [156, 167]. Introducing the hopping parameter of quark flavor f,

$$\kappa^f = \frac{1}{2(am^f + 4)}, \tag{5.82}$$

one can rewrite the unimproved Wilson-Dirac operator eq. (3.31) as follows [83],

$$Q_W^f = \frac{1}{2a\kappa^f}(\mathbb{1} - \kappa^f H), \tag{5.83}$$

where the hopping matrix H collects the nearest-neighbor terms in the Dirac operator,

$$H(x,y) = \sum_\mu [(\mathbb{1} - \gamma_\mu)U_\mu(x)\delta_{x+a\hat{\mu},y} + (\mathbb{1} + \gamma_\mu)U_{-\mu}(x)\delta_{x-a\hat{\mu},y}]. \tag{5.84}$$

The form eq. (5.83) allows an expansion of the inverse Dirac operator, *i.e.*, the propagator, in a geometric series [156],

$$(Q_W^f)^{-1} = 2a\kappa^f \sum_{j=0}^{\infty}(\kappa^f H)^j = 2a\kappa^f \sum_{j=0}^{n-1}(\kappa^f H)^j + (\kappa^f H)^n (Q_W^f)^{-1}. \tag{5.85}$$

The benefit of this expansion lies in the fact that $(Q_W^f)^{-1}$ on the right-hand side is multiplied by $(\kappa^f)^n$, where $\kappa^f < 1$, thus suppressing the noise introduced by the stochastic estimation of the inverse [168].

When $\mathcal{O}(a)$-improvement is employed, the hopping-parameter expansion cannot be applied directly because the Sheikholeslami-Wohlert term is added to eq. (5.83) (*cf.*, eq. (3.38)). Therefore, the hopping-parameter expansion needs to be generalized [168]. In this context, it is advantageous to make use of the even-odd decomposition of the Wilson-Clover Dirac operator in analogy to eq. (4.3) [162]. For this purpose, we introduce the (2×2)-matrices

$$\mathbb{1}_{ee} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathbb{1}_{eo} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathbb{1}_{oe} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \mathbb{1}_{oo} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \tag{5.86}$$

and write the Dirac matrix as

$$Q = \mathbb{1}_{ee} \otimes Q_{ee} + \mathbb{1}_{eo} \otimes Q_{eo} + \mathbb{1}_{oe} \otimes Q_{oe} + \mathbb{1}_{oo} \otimes Q_{oo}, \tag{5.87}$$

where $\otimes$ denotes the Kronecker product (flattened tensor product). Here and in the following, we omit the flavor index f for notational convenience. The decomposition in eq. (5.87) implies

$$\begin{aligned}
&(\mathbb{1} + \mathbb{1}_{eo} \otimes Q_{eo}Q_{oo}^{-1} + \mathbb{1}_{oe} \otimes Q_{oe}Q_{ee}^{-1})(\mathbb{1}_{ee} \otimes Q_{ee} + \mathbb{1}_{oo} \otimes Q_{oo}) \\
&\quad = (\mathbb{1}_{ee} \otimes \mathbb{1} + \mathbb{1}_{oo} \otimes \mathbb{1} + \mathbb{1}_{eo} \otimes Q_{eo}Q_{oo}^{-1} + \mathbb{1}_{oe} \otimes Q_{oe}Q_{ee}^{-1})(\mathbb{1}_{ee} \otimes Q_{ee} + \mathbb{1}_{oo} \otimes Q_{oo}) \\
&\quad = \mathbb{1}_{ee} \otimes Q_{ee} + \mathbb{1}_{oo} \otimes Q_{oo} + \mathbb{1}_{eo} \otimes Q_{eo} + \mathbb{1}_{oe} \otimes Q_{oe} = Q,
\end{aligned} \tag{5.88}$$

where we have used the mixed-product property of the Kronecker and matrix products as well as the fact that $\mathbb{1}_{ee}^2 = \mathbb{1}_{ee}$, $\mathbb{1}_{oo}^2 = \mathbb{1}_{oo}$, $\mathbb{1}_{ee}\mathbb{1}_{oo} = \mathbb{1}_{oo}\mathbb{1}_{ee} = \mathbb{1}_{eo}\mathbb{1}_{ee} = \mathbb{1}_{oe}\mathbb{1}_{oo} = 0$, $\mathbb{1}_{eo}\mathbb{1}_{oo} = \mathbb{1}_{eo}$, and $\mathbb{1}_{oe}\mathbb{1}_{ee} = \mathbb{1}_{oe}$. We hence find for the inverse of the Dirac matrix,

$$Q^{-1} = (\mathbb{1}_{ee} \otimes Q_{ee} + \mathbb{1}_{oo} \otimes Q_{oo})^{-1}(\mathbb{1} + \mathbb{1}_{eo} \otimes Q_{eo}Q_{oo}^{-1} + \mathbb{1}_{oe} \otimes Q_{oe}Q_{ee}^{-1})^{-1}. \tag{5.89}$$

Defining the generalized hopping matrix

$$\tilde{H} = -(\mathbb{1}_{eo} \otimes Q_{eo} Q_{oo}^{-1} + \mathbb{1}_{oe} \otimes Q_{oe} Q_{ee}^{-1}), \tag{5.90}$$

which connects the even and odd lattice sites, we can thus again apply a geometric series expansion to the propagator [162],

$$Q^{-1} = (\mathbb{1}_{ee} \otimes Q_{ee} + \mathbb{1}_{oo} \otimes Q_{oo})^{-1} \sum_{j=0}^{2n-1} \tilde{H}^j + Q^{-1} \tilde{H}^{2n} = M_{2n} + R_{2n}. \tag{5.91}$$

An efficient stochastic estimator for the contribution to the quark loop involving the remainder term R_{2n} can be obtained by evenly splitting the $2n$ powers of $\tilde{H}$ in analogy to the one-end trick [162],

$$-\frac{1}{N_s} \sum_{j=1}^{N_s} \sum_{\mathbf{z}} \sum_{x,y,b} e^{i\mathbf{q}\cdot\mathbf{z}} \eta_j^\dagger(y) \tilde{H}^n(y,z) \gamma_\mu S(z,x) \tilde{H}^n(x,b) \eta_j(b). \tag{5.92}$$

Thus, both noise sources are ultraviolet filtered by $\tilde{H}^n$, which provides a significant variance reduction compared to the plain stochastic estimator in eq. (5.78) [162]. It has been found in Ref. [164] that using $n = 2$ for the generalized hopping-parameter expansion (*i.e.*, order $2n = 4$) combined with 4096 stochastic volume sources is sufficient to completely suppress the noise from the stochastic estimation of the remainder term.

Hierarchical probing

The remaining task consists in computing the first term of the generalized hopping-parameter expansion, *i.e.*, the (trace of the) matrix M_{2n} in eq. (5.91). As this is a sparse matrix, its diagonal elements can be obtained by probing [162]. If for a (sparse) matrix M there exist K probing vectors $v^0, \ldots, v^{K-1}$ which fulfill $\sum_{k=0}^{K-1} v_i^k v_j^k = \delta_{ij}$ for all i, j for which $M_{ij} \neq 0$, then the diagonal elements of M are given by [162],

$$M_{ii} = \sum_{k=0}^{K-1} \sum_j v_i^k M_{ij} v_j^k, \tag{5.93}$$

where we have not summed over i. If all K probing vectors are used, this is obviously an exact method. In Ref. [162], an efficient scheme has been suggested from which the required probing vectors can be obtained for local operators, *i.e.*, ones which only involve strictly diagonal elements of the propagator (and thus of M_{2n}). For the conserved vector current, however, we also need point-split operators involving elements of M_{2n} which are displaced from the diagonal by one site in the spatial indices (*cf.*, eqs. (5.71) and (5.72)).

For this purpose, a more general probing scheme is required [164]. We use hierarchical probing, which has been introduced in Ref. [169] to estimate the trace of the full inverse of the Dirac matrix, but is also applicable to the calculation of the trace of M_{2n}. The construction in Ref. [169] starts from the observation that for a general (sparse) matrix M, the appropriate probing vectors can be derived from a coloring of the graph of M. If the non-zero elements M_{ij} of M are those whose vertices i and j are not farther than k links apart in the graph of M, what is required is a distance-k coloring of this graph. This coloring is equivalent to the distance-1 coloring of the graph of M^k [169]. Computing M^k for large k can be, however, time and/or memory intensive, and the usual coloring algorithms also quickly become prohibitively expensive. Moreover, the colorings and the probing vectors derived from them are in general not related for different ks [169]. This means that in the process of tuning k, the work done for a previous, smaller value of k cannot be reused.

For these reasons, the concept of hierarchical coloring has been introduced in Ref. [169]. It relaxes the requirement of a full distance-k coloring, instead enforcing a nested coloring which can be obtained recursively: in each step, one computes independent distance-1 colorings of the subgraphs corresponding to each color from the previous coloring step. This hierarchical coloring produces more colors than the classic coloring algorithms, but is computationally tractable. Furthermore, the additional probing vectors associated with the extra colors improve the variance reduction achieved by the method [169].

To facilitate the process of finding appropriate probing vectors from the hierarchical coloring, we define the Hadamard matrices. An $N \times N$ matrix H is a Hadamard matrix if it has entries $H_{ij} = \pm 1$ and $HH^T = N\mathbb{1}_N$. N must be 1, 2, or a multiple of 4. Hadamard matrices whose dimension is a power of 2 can recursively be obtained as [169],

$$H_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad H_{2N} = H_2 \otimes H_N = \begin{pmatrix} H_N & H_N \\ H_N & -H_N \end{pmatrix}. \tag{5.94}$$

Hadamard vectors, *i.e.*, columns of Hadamard matrices, naturally project out the non-zero elements of banded matrices when used as probing vectors for the latter. For matrices with an arbitrary sparsity structure, one first performs the hierarchical coloring as explained above and then uses the thus obtained permutation of the matrix rows and columns to find the set of Hadamard vectors which spans the probing basis for this coloring. This hierarchical probing algorithm can be implemented very efficiently [169].

To remove the stochastic bias originating from deterministic probing vectors, one can generate a random vector and element-wise multiply it with the sequence of Hadamard vectors constructed as sketched above. The thus obtained vectors have the same properties as the original, deterministically determined probing vectors, but do not introduce a bias [169]. For the random vector, we use a spin- and color-diluted stochastic volume source. Here, dilution means that instead of sampling from one random distribution following eq. (5.77) for the full set of space-time, spin and color indices, one uses 12 independent distributions, one for each combination of spin and color indices [170]. This completely eliminates the variance due to the statistical deviation from zero of the spin-and-color-off-diagonal elements of eq. (5.77). It has been found in Ref. [164] that 512 Hadamard probing vectors are sufficient to saturate gauge noise for the point-split operators, while for the local operators, the thus defined method produces an exact result for the required trace of M_{2n}.

Frequency splitting

In Ref. [162], a very efficient strategy to compute the loops of a single flavor has been suggested based on the techniques described in the previous three subsections. The idea consists in calculating the one-end trick estimator eq. (5.81) for a chain of flavors $f_1, \ldots, f_{N_f}$ with $m^{f_1} < m^{f_2} < \cdots < m^{f_{N_f}}$ and evaluating the individual loops of the heaviest flavor f_{N_f} explicitly using the generalized hopping-parameter expansion. From these ingredients, one can recursively reconstruct the single-flavor loops of all other quark flavors. In this way, any single-flavor loop is split into several contributions arising from different frequencies, which allows for an individually optimized treatment of each of them [162].

Note that not all of the flavors employed in the aforementioned chain need to correspond to physical quarks which are present in the system under study; one is free to introduce intermediate flavors of arbitrary masses. We make use of this freedom to define an intermediate quark flavor i with a mass between the strange and charm quark, the hopping parameter κ_i of which we set according to [164],

$$\frac{1}{\kappa_i} = \frac{3}{4}\frac{1}{\kappa_s} + \frac{1}{4}\frac{1}{\kappa_c}. \tag{5.95}$$

The heaviest quark flavor we consider (to which we also apply the generalized hopping-parameter expansion) is the charm. This means that we employ in total four quark flavors: light, strange, intermediate and charm.

We use 512 stochastic volume sources to compute the difference of the light- and strange-quark loops with the one-end trick, and double this number for each subsequent flavor, *i.e.*, 1024 noise sources for strange minus intermediate and 2048 for intermediate minus charm. For the calculation of the remainder term in the generalized hopping-parameter expansion of the charm-quark loops, we double this number again, leading to the 4096 stochastic volume sources mentioned below eq. (5.92) above [164].

In this thesis, we consider isoscalar observables with contributions up to the strange quark. According to eq. (5.70), the corresponding three-point functions only depend on the difference between the light- and strange-quark loops. This can be directly estimated from the one-end trick, obviating the calculation of any single-flavor loops. Nevertheless, the latter have been computed on a large set of CLS ensembles in the context of Ref. [164] for the construction of different observables than the ones studied in this thesis.

Setup of sources for the two-point function

Using the methods described above, the variance of the quark loops has been driven down to gauge noise. A further improvement of the signal for the quark-disconnected contribution to the nucleon three-point function at the same number of gauge configurations can hence only be achieved by increasing the statistics for the nucleon two-point function. Therefore, on ensembles with open boundary conditions, we make use of the additional measurements of the two-point function which we have already discussed in detail in section 5.3.5. For these, the additional quark sources have been placed on several timeslices away from the middle of the lattice, so that we again need to ensure that the results are not significantly impacted by boundary effects.

For the construction of the disconnected contribution to the three-point function, we use smaller source-sink separations than for the fits to the two-point functions from which we extract the nucleon mass: the largest source-sink separation considered for the three-point functions is of the order of 1.4 fm to 1.5 fm (*cf.*, table 5.2), while for the determination of the nucleon mass, we employ fit ranges up to 2.0 fm. Moreover, the vector current quark loops are much less sensitive to boundary effects than the scalar ones we have looked at in section 5.3.5. This is illustrated in fig. 5.7, where the gauge average of $L_{V_0}^{\text{light}}(2\pi/L\hat{x}) - L_{V_0}^{s}(2\pi/L\hat{x})$ is shown on the open-boundary ensemble C101, and which is to be compared to the scalar loop in fig. 5.3. Furthermore, the disconnected contribution to the three-point functions is much more noisy than the two-point functions themselves, emphasizing the need to use as many sources as possible without introducing a systematic bias. Restricting the source positions to a small region in the bulk of the lattice, however, severely limits the achievable precision, as already noted in section 5.3.5.

For these reasons, the tradeoff between gaining statistics and ensuring that there is no significant contamination by boundary effects is different for the construction of the disconnected contribution to the three-point function than for the determination of the nucleon mass. Accordingly, we impose here a relaxed criterion of 2.5 fm which the source of the nucleon propagating towards the boundary has to keep from the latter. We remark that this is the same criterion as the one employed in Ref. [67]. On C101, for instance, this corresponds to the orange set of points in fig. 5.4. As we have already observed, for $t_{\text{sep}} < 2.0$ fm these points are very well compatible with the green ones which are based on a stricter criterion and are the ones used for the determination of the nucleon mass. This, together with the absence of strong boundary effects in fig. 5.7, already indicates that for the range of source-sink separations we

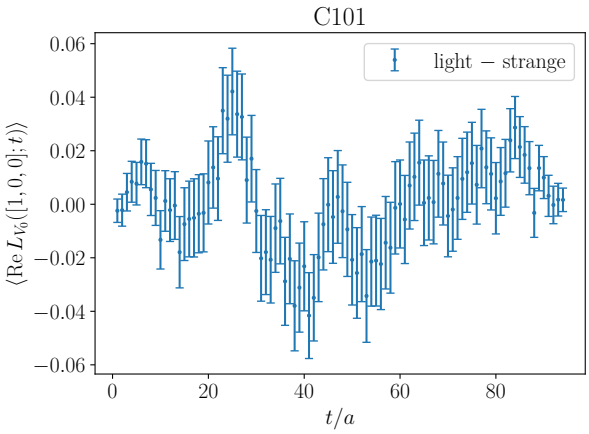


Figure 5.7: Gauge average of the real part of the 0-component of the vector quark loop at the first non-vanishing momentum in x-direction on the ensemble C101 as a function of the timeslice on which it is evaluated. Shown is the correlated difference between the light- and strange-quark loop.

intend to use for the three-point functions, the relaxed criterion introduced above should not lead to a significant contamination by boundary effects. For more extensive crosschecks of the form factors obtained with different criteria for the distance to the boundaries, we refer to the discussion in section 6.5 below.

The resulting total number of high- and low-precision measurements per ensemble are listed in table 5.3. In analogy to table 5.1, we have counted sources on which both propagation directions are employed twice, and such on which only one propagation direction is used only once.

Table 5.3: Number of gauge configurations and aggregated number of high-precision (HP) and low-precision (LP) measurements of the two-point function used for the computation of the disconnected part

ID	N_{cfg}	$N_{\text{meas,HP}}$	$N_{\text{meas,LP}}$
C101	994	15 902	475 930
N101	1588	6352	813 036
H105	1024	51 170	496 662
D450	498	7968	127 488
N451	1010	16 160	258 560
E250	796	12 736	407 552
D200	998	17 964	542 516
N200	1708	27 328	812 032
S201	2092	8362	192 558
E300	569	2276	327 744
J303	1073	6438	291 744

5.5 The pion two-point function

Even though this thesis is focussed on the structure of the nucleon, we also need to compute the two-point function of the pion in order to determine the pion mass on our ensembles from it. This is because the pion mass is used as a proxy for the light-quark mass which is varied between ensembles (*cf.*, the discussion in section 4.3). As we will see in section 6.7, the pion mass exerts a strong influence on the electromagnetic form factors and radii of the nucleon. Therefore, it is desirable to extract the pion masses on our ensembles with a good precision and appropriate error estimates, using methods which mirror the ones used for the nucleon mass as closely as possible. This enables us to perform a consistent and self-contained analysis.

Accordingly, we define the momentum-projected pion two-point function,

$$C_{2,\pi}(\mathbf{p}';x,y_0) = \sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})}\Pi^+(y)\bar{\Pi}^+(x), \tag{5.96}$$

where Π^+ and $\bar{\Pi}^+$ are the pion annihilation and creation operators from eqs. (3.72) and (3.75), respectively. Again, translation invariance implies that the expectation value of $C_{2,\pi}$ only depends on the source-sink separation $t_{\text{sep}} = y_0 - x_0$, and not on the source position x.

5.5.1 Spectral representation of the pion two-point function

In order to quantify the relation of the pion two-point function with the pion mass, we will now derive the spectral representation of $\langle C_{2,\pi}\rangle$. We will proceed in analogy to the case of the nucleon which we have discussed in section 5.3.1, *i.e.*, set the source position x to zero, insert a complete set of states and use the Heisenberg evolution in (Euclidean) time. The main difference to the nucleon is that the pion is a spin-zero particle, which actually simplifies the calculation. Thus, we define a complete set of spin-zero states $|n\rangle$ which are eigenstates of the Hamiltonian H according to $H|n\rangle = E_n|n\rangle$, and the completeness relation of which reads,

$$\mathbb{1} = \sum_n \frac{1}{2E_n}|n\rangle\langle n|. \tag{5.97}$$

For the nucleon, we have assumed that the overlaps of the states created by the interpolators with the eigenstates of the Hamiltonian are proportional to solutions of the free Dirac equation (*cf.*, eqs. (5.23) and (5.24)). Analogously, we will now assume for the pion that the overlaps of $\hat{\Pi}^+$ with the $|n\rangle$ states are proportional to solutions of the free Klein-Gordon equation, *i.e.*, plane waves,

$$\langle 0|\hat{\Pi}^+(0,\mathbf{y})|n\rangle = e^{i\mathbf{p}_n\cdot\mathbf{y}}c_n(\mathbf{p}_n), \tag{5.98}$$

$$\langle n|(\hat{\Pi}^+)^\dagger(0,\mathbf{y})|0\rangle = e^{-i\mathbf{p}_n\cdot\mathbf{y}}c_n^*(\mathbf{p}_n), \tag{5.99}$$

where $\mathbf{p}_n$ is the total three-momentum corresponding to the state $|n\rangle$.

Using these ingredients, we obtain for the expectation value of the pion two-point function,

$$\begin{aligned}
\langle C_{2,\pi}\rangle(\mathbf{p}';y_0) &= \sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot\mathbf{y}} \sum_n \frac{1}{2E_n}\langle 0|\hat{\Pi}^+(y)|n\rangle\langle n|(\hat{\Pi}^+)^\dagger(0)|0\rangle \\
&= \sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot\mathbf{y}} \sum_n \frac{1}{2E_n}\left\langle 0\middle| e^{Hy_0}\hat{\Pi}^+(0,\mathbf{y})e^{-Hy_0}\middle|n\right\rangle\langle n|(\hat{\Pi}^+)^\dagger(0)|0\rangle \\
&= \sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot\mathbf{y}} \sum_n e^{-E_n y_0}\frac{1}{2E_n}\langle 0|\hat{\Pi}^+(0,\mathbf{y})|n\rangle\langle n|(\hat{\Pi}^+)^\dagger(0)|0\rangle
\end{aligned}$$

$$= \sum_n \sum_{\mathbf{y}} e^{-i(\mathbf{p}'-\mathbf{p}_n)\cdot\mathbf{y}} e^{-E_n y_0} \frac{|c_n(\mathbf{p}_n)|^2}{2E_n} = \sum_n \delta_{\mathbf{p}',\mathbf{p}_n} e^{-E_n y_0} \frac{|c_n(\mathbf{p}')|^2}{2E_n} \tag{5.100}$$

$$\xrightarrow{y_0 \gg 0} e^{-E_{\pi,\mathbf{p}'} y_0} \frac{|c_\pi(\mathbf{p}')|^2}{2E_{\pi,\mathbf{p}'}}. \tag{5.101}$$

For $y_0 \to \infty$, only the pion state with momentum $\mathbf{p}'$ and energy $E_{\pi,\mathbf{p}'} = \sqrt{M_\pi^2 + \mathbf{p}'^2}$ contributes to the spectral sum.

On ensembles with (anti-)periodic boundary conditions in time, the contribution of the pion propagating backward in time, *i.e.*, in $T - y_0$ instead of in y_0, needs to be taken into account. Because $(\Pi^+)^\dagger = \Pi^-$ (*cf.*, eq. (3.75)) and the π^+ and π^- are degenerate in the case of exact SU(2) isospin symmetry, the energies and overlaps of the forward- and backward-propagating pion are equal [80]. Consequently, the ground-state contribution to the expectation value of the pion two-point function in a finite periodic box is given by

$$\begin{aligned}\langle C_{2,\pi} \rangle (\mathbf{p}'; y_0 \gg 0) &= \frac{|c_\pi(\mathbf{p}')|^2}{2E_{\pi,\mathbf{p}'}} \left(e^{-E_{\pi,\mathbf{p}'} y_0} + e^{-E_{\pi,\mathbf{p}'}(T-y_0)} \right) \\ &= \frac{|c_\pi(\mathbf{p}')|^2}{E_{\pi,\mathbf{p}'}} e^{-E_{\pi,\mathbf{p}'} T/2} \cosh\left[\left(\frac{T}{2} - y_0 \right) E_{\pi,\mathbf{p}'} \right]. \end{aligned} \tag{5.102}$$

We remark that an analogous contribution from backward propagation also exists for the nucleon. However, due to the nucleon's fermionic nature, this contribution describes not the nucleon itself, but its parity partner [80]. Therefore, if we want to consider the backward propagation also in this case, we need to flip the sign of the parity-projection operator in eq. (3.78), as already mentioned in section 5.3.5 above. One thus computes two different nucleon correlators, one with Γ_+^p and forward propagation and one with Γ_-^p and backward propagation. After averaging the two, one can apply the formula with only one exponential, *cf.*, eq. (5.29).

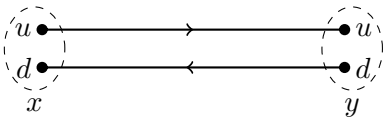
5.5.2 Implementation of the pion two-point function

In order to implement the calculation of the pion two-point function, we need to evaluate its fermionic expectation value using Wick's theorem. Plugging the pion interpolators eqs. (3.72) and (3.75) into eq. (5.96) and applying Wick's theorem eq. (2.39) to the fermionic expectation value, we find,

$$\begin{aligned}\langle C_{2,\pi} \rangle_F &= \left\langle \sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})} \Pi^+(y) \bar{\Pi}^+(x) \right\rangle_F \\ &= \left\langle \sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})} \bar{\psi}^d_{\alpha j}(y) \gamma_{5,\alpha\beta} \psi^u_{\beta j}(y) \bar{\psi}^u_{\gamma k}(x) \gamma_{5,\gamma\delta} \psi^d_{\delta k}(x) \right\rangle_F \\ &= -\sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})} \gamma_{5,\alpha\beta} \gamma_{5,\gamma\delta} S^d(x-y)_{\delta k,\alpha j} S^u(y-x)_{\beta j,\gamma k} \\ &= -\sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})} \operatorname{tr}[\gamma_5 S^d(x-y) \gamma_5 S^u(y-x)]. \end{aligned} \tag{5.103}$$

Since the Wilson-Clover Dirac operator is γ_5-hermitian, the same is valid for the propagators, *i.e.*, $\gamma_5 S^f(x-y)\gamma_5 = (S^f(y-x))^\dagger$. Using this relation, we can further simplify the trace in eq. (5.103),

$$\langle C_{2,\pi} \rangle_F = -\sum_{\mathbf{y}} e^{-i\mathbf{p}'\cdot(\mathbf{y}-\mathbf{x})} \operatorname{tr}\left[(S^d(y-x))^\dagger S^u(y-x) \right] =$$



(5.104)

We have thus obtained only one quark-connected contraction. Equation (5.104) furthermore shows that the pion two-point function can be computed from the same type of point-to-all propagators which we use for the nucleon two-point function.

The full expectation value of the pion two-point function is calculated by averaging eq. (5.104) over all N generated gauge configurations. The variance of such a Monte-Carlo estimate can be assessed analogously to eq. (5.35),

$$\sigma^2_{C_{2,\pi}} \propto \frac{1}{N}\left(\langle (S^\dagger S)^2\rangle - \langle S^\dagger S\rangle^2\right). \tag{5.105}$$

At large source-sink separations, the first term is dominated by an exponential falling off with the energy of the lowest state coupling to two quarks and two antiquarks, *i.e.*, two pions [152]. Therefore, both terms in eq. (5.105) are for $\mathbf{p}' = \mathbf{0}$ asymptotically proportional to $\exp(-2M_\pi t_{\text{sep}})$. From this, we expect that

$$\sigma_{C_{2,\pi}}(\mathbf{p}'; t_{\text{sep}}) \propto \frac{\langle C_{2,\pi}\rangle(\mathbf{p}'; t_{\text{sep}})}{\sqrt{N}}, \tag{5.106}$$

so that the signal-to-noise ratio is approximately independent of t_{sep} [152]. This situation is much more favorable than for the nucleon, where we have found an exponentially decaying signal-to-noise ratio in eq. (5.37).

The zero-momentum two-point function of the pion on the ensemble N451 is shown in the left panel of fig. 5.8. Because N451 has periodic boundary conditions, $\langle C_{2,\pi}\rangle$ is symmetric around $T/2$ (*cf.*, eq. (5.102)). In order to reduce the number of data points which we will eventually have to fit to determine the pion mass, we hence symmetrize the two-point function around $T/2$, *i.e.*, we define

$$\langle C_{2,\pi}\rangle^{\text{sym}}(\mathbf{p}'; t_{\text{sep}}) = \frac{1}{2}\left[\langle C_{2,\pi}\rangle(\mathbf{p}'; t_{\text{sep}}) + \langle C_{2,\pi}\rangle(\mathbf{p}'; T - t_{\text{sep}})\right], \quad 0 < t_{\text{sep}} < \frac{T}{2}. \tag{5.107}$$

Moreover, eq. (5.102) implies that the ratio $\langle C_{2,\pi}\rangle(\mathbf{p}'; t_{\text{sep}} + a)/\langle C_{2,\pi}\rangle(\mathbf{p}'; t_{\text{sep}})$ is not proportional to $\exp[-aE_{\pi,\text{eff}}(\mathbf{p}'; t_{\text{sep}})]$, so that we cannot use the simple eq. (5.50) to calculate the effective energy of the pion. Instead, the effective energy is given by the solution of [83]

$$\frac{\langle C_{2,\pi}\rangle^{\text{sym}}(\mathbf{p}'; t_{\text{sep}})}{\langle C_{2,\pi}\rangle^{\text{sym}}(\mathbf{p}'; t_{\text{sep}} + a)} = \frac{\cosh\left[E_{\pi,\text{eff}}(\mathbf{p}'; t_{\text{sep}})(T/2 - t_{\text{sep}})\right]}{\cosh\left[E_{\pi,\text{eff}}(\mathbf{p}'; t_{\text{sep}})(T/2 - t_{\text{sep}} - a)\right]} \tag{5.108}$$

at each value of t_{sep}. Although no closed analytic solution of eq. (5.108) for $E_{\pi,\text{eff}}$ is known, the solution can easily be found with numerical methods. The thus determined effective mass $m_{\pi,\text{eff}} \equiv E_{\pi,\text{eff}}(\mathbf{p}' = \mathbf{0})$ is plotted in the right panel of fig. 5.8. It can be seen that a good signal with almost constant errors is obtained even for very large source-sink separations, which is to be contrasted with the nucleon effective mass shown in the right panel of fig. 5.2.

This demonstrates that the achievable precision for the pion at a given number of sources is much better than for the nucleon. Moreover, the error of the pion mass is completely subdominant in our analysis compared to the one of the baryonic observables. Therefore, the pion two-point function is generally only measured on a subset of the sources employed for the nucleon. The total number of high- and low-precision measurements per ensemble which we use for the pion are listed in table 5.4. On the ensembles with periodic boundary conditions (D450, N451 and E250), we average over the forward- and backward propagating pion according to eq. (5.107). Accordingly, we have counted all source positions twice for these ensembles, so as to enable a fairer comparison with the numbers on the ensembles with open boundary conditions, where we only ever use the forward-propagating pion.

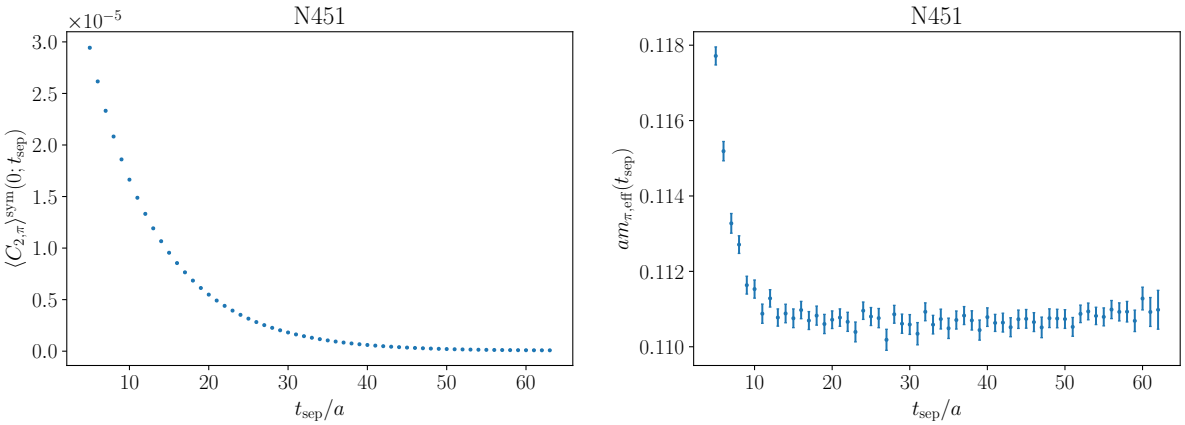


Figure 5.8: Zero-momentum pion two-point function and effective mass on the ensemble N451

Table 5.4: Number of gauge configurations and aggregated number of high-precision (HP) and low-precision (LP) measurements used for the determination of the pion mass

ID	N_{cfg}	$N_{\text{meas,HP}}$	$N_{\text{meas,LP}}$
C101	1988	3974	94 209
N101	1588	1588	50 816
H105	1024	1024	16 384
D450	498	7968	127 488
N451	1010	16 160	258 560
E250	398	6368	203 776
D200	1996	3991	95 354
N200	1708	3416	36 361
S201	2092	4181	96 279
E300	569	569	18 208
J303	1073	2146	25 696

6 Analysis and results

After having measured the relevant correlation functions on the lattice, the next step consists in analyzing the raw data in order to extract the observables of interest. In our case, these are the electromagnetic form factors of the proton and neutron. From the behavior of the form factors at $Q^2 = 0$, we will determine the electromagnetic radii (*cf.*, eq. (5.9)) and magnetic moments. Moreover, we will explore other definitions of radii which are relevant for the study of the proton's electromagnetic size in atomic spectroscopy experiments. For these so-called Zemach and Friar radii, integrals over the form factors need to be evaluated instead of derivatives.

We will start this chapter by reviewing some of the lattice-specific analysis techniques, before moving on to our particular analysis of hadron masses, form factors and radii. Lastly, we will introduce a method to systematically average the results of different analysis variations, and compare the thus obtained final results to other studies on and off the lattice.

6.1 Autocorrelation

In Lattice-QCD simulations, the gauge configurations on which the observables are measured are generated from a stochastic sequence, the Markov chain (*cf.*, section 3.2.3). As a consequence, successive configurations are not independent of each other, so that there is also a high chance that the observables measured on them are correlated. To make a quantitative statement about this correlation, one defines the autocorrelation function of the primary observable O [83],

$$\begin{aligned} C_O(\hat{t}, \hat{t} + \Delta\hat{t}) &= \left\langle [O(\hat{t}) - \langle O(\hat{t})\rangle][O(\hat{t} + \Delta\hat{t}) - \langle O(\hat{t} + \Delta\hat{t})\rangle] \right\rangle \\ &= \langle O(\hat{t})O(\hat{t} + \Delta\hat{t})\rangle - \langle O(\hat{t})\rangle \langle O(\hat{t} + \Delta\hat{t})\rangle . \end{aligned} \tag{6.1}$$

Here, $\hat{t}$ is the computer time introduced in section 3.3.1, which labels the consecutive configurations. The angle brackets refer to the actual expectation values under the equilibrium distribution, which are defined at an infinite number of configurations N. For a Markov chain in equilibrium, which is reached for $\hat{t} \to \infty$, the expectation values become independent of $\hat{t}$, so that one has $\langle O(\hat{t})\rangle = \langle O(\hat{t} + \Delta\hat{t})\rangle$. Thus, the autocorrelation function depends in this case only on the computer time separation $\Delta\hat{t}$ [80], and we write,

$$C_O(\Delta\hat{t}) = C_O(\hat{t}, \hat{t} + \Delta\hat{t}). \tag{6.2}$$

The normalized autocorrelation function is then defined as

$$\Gamma_O(\Delta\hat{t}) = \frac{C_O(\Delta\hat{t})}{C_O(0)}, \tag{6.3}$$

where $C_O(0) = \sigma_O^2$, the variance of the observable O.

For correlated random variables $O(n)$, the variance of the Monte-Carlo estimator of the mean value (*cf.*, eq. (3.49)) is [83]

$$\sigma^2_{\hat{O}} = \left\langle (\hat{O} - \langle O\rangle)^2 \right\rangle = \left\langle \left(\frac{1}{N} \sum_{n=1}^{N} [O(n) - \langle O\rangle] \right)^2 \right\rangle = \frac{1}{N^2} \left\langle \sum_{n,m=1}^{N} [O(n) - \langle O\rangle][O(m) - \langle O\rangle] \right\rangle$$

$$= \frac{1}{N^2} \sum_{n,m=1}^{N} C_O(|n-m|) = \frac{1}{N^2} \sum_{t=-N}^{N} (N-|t|) C_O(|t|) = \frac{C_O(0)}{N} \sum_{t=-N}^{N} \Gamma_O(|t|) \left(1 - \frac{|t|}{N}\right)$$
$$\approx \frac{\sigma_O^2}{N} 2 \left(\frac{1}{2} + \sum_{t=1}^{N} \Gamma_O(t)\right) = \frac{\sigma_O^2}{N} 2\tau_{O,\mathrm{int}}(N), \tag{6.4}$$

where the integrated autocorrelation time

$$\tau_{O,\mathrm{int}}(N) = \frac{1}{2} + \sum_{t=1}^{N} \Gamma_O(t) = \frac{1}{2} \sum_{t=-N}^{N} \Gamma_O(|t|) \tag{6.5}$$

has been introduced. From the second to the last line in eq. (6.4), the factor $1 - |t|/N$ has been neglected. This is justified for $N \to \infty$, as $\Gamma_O(t)$ is exponentially suppressed for large t [83]. For the practical determination of the autocorrelation time this implies that one may truncate the sum in eq. (6.5) at relatively small values of t. One can thus avoid using large values of t, where the estimate of $\Gamma_O(t)$ might become unreliable. Comparing the results for different truncations then allows an accurate measurement of the autocorrelation time (*cf.*, section 6.1.1).

For completely uncorrelated observables, $\Gamma_O(t) = 0 \ \forall t > 0$, so that $\tau_{O,\mathrm{int}} = 1/2$ and $\sigma_{\hat{O}}^2 = \sigma_O^2/N$. A comparison of the full eq. (6.4) with the result for uncorrelated observables reveals the significance of the integrated autocorrelation time: the effective number of independent measurements is $N/2\tau_{O,\mathrm{int}}$. If the numerical measurement of O takes a substantial amount of CPU time (as is generally the case for hadronic observables which require inversions of the Dirac operator), it is usually advantageous to skip about $2\tau_{O,\mathrm{int}}$ configurations between the measurements. On the other hand, if the autocorrelation time is expected to be small for the observable of interest and only a limited number of configurations are available, it can make sense to perform measurements on each configuration.

The discussion so far, and in particular the calculation in eq. (6.4), is only valid for primary observables, *i.e.*, such which can be calculated directly on a single configuration and the expectation value of which can be estimated with eq. (3.49) [80]. In lattice studies, however, one is frequently dealing with secondary or derived observables which are determined by functions of the averages of primary observables, $\hat{F} = f(\hat{O}_\alpha)$ [171]. Here, α labels different primary observables, and $\hat{F}$ can depend on several $\hat{O}_\alpha$s. In particular, due to the reweighting, all observables considered in this thesis are secondary (*cf.*, eq. (3.71)).

The variance of $\hat{F}$ is given by $\sigma_{\hat{F}}^2 = \langle (\hat{F} - F)^2 \rangle$. In order to evaluate this expression, we Taylor expand the estimator $\hat{F}$ around the true expectation value F [171],

$$\hat{F} = F + \sum_{\alpha} f_\alpha \left(\hat{O}_\alpha - \langle O_\alpha \rangle\right) + \frac{1}{2} \sum_{\alpha,\beta} f_{\alpha\beta} \left(\hat{O}_\alpha - \langle O_\alpha \rangle\right) \left(\hat{O}_\beta - \langle O_\beta \rangle\right) + \mathcal{O}\left(\left(\hat{O}_\alpha - \langle O_\alpha \rangle\right)^3\right), \tag{6.6}$$

where the coefficients are given by derivatives of the function f evaluated at the exact values of the primary observables,

$$f_\alpha = \left.\frac{\partial f}{\partial O_\alpha}\right|_{\langle O_\alpha \rangle}, \quad f_{\alpha\beta} = \left.\frac{\partial^2 f}{\partial O_\alpha \partial O_\beta}\right|_{\langle O_\alpha \rangle, \langle O_\beta \rangle}. \tag{6.7}$$

It follows from eq. (6.6) that $\hat{F}$ is a biased estimator unless the function f is linear [171],

$$\langle \hat{F} - F \rangle \approx \frac{1}{2} \sum_{\alpha,\beta} f_{\alpha\beta} C_{\widehat{\alpha\beta}}, \tag{6.8}$$

where

$$C_{\widehat{\alpha\beta}} = \left\langle \left(\hat{O}_\alpha - \langle O_\alpha \rangle\right) \left(\hat{O}_\beta - \langle O_\beta \rangle\right) \right\rangle \tag{6.9}$$

is the covariance matrix of the estimators $\hat{O}_\alpha$ and $\hat{O}_\beta$.

To derive an expression for the covariance matrix of the estimators in the presence of autocorrelation, we define the cross-autocorrelation function of the primary observables O_α and O_β [171],

$$C_{\alpha\beta}(\Delta\hat{t}) = \left\langle [O_\alpha(\hat{t}) - \langle O_\alpha(\hat{t}) \rangle][O_\beta(\hat{t} + \Delta\hat{t}) - \langle O_\beta(\hat{t} + \Delta\hat{t}) \rangle] \right\rangle \tag{6.10}$$

and its normalized equivalent

$$\Gamma_{\alpha\beta}(\Delta\hat{t}) = \frac{C_{\alpha\beta}(\Delta\hat{t})}{C_{\alpha\beta}(0)}. \tag{6.11}$$

Here, $C_{\alpha\beta}(0) = C_{\alpha\beta}$, the covariance matrix of the (primary) observables O_α and O_β. By performing a calculation similar to eq. (6.4) one finds for the covariance matrix of the estimators [171],

$$C_{\widehat{\alpha\beta}} = \frac{C_{\alpha\beta}}{N} \sum_{t=-N}^{N} \Gamma_{\alpha\beta}(|t|) \left(1 - \frac{|t|}{N}\right) \approx \frac{C_{\alpha\beta}}{N} 2\tau_{\alpha\beta,\mathrm{int}}(N), \tag{6.12}$$

where $\tau_{\alpha\beta,\mathrm{int}}(N)$ is defined analogously to eq. (6.5).

Using this result, we can derive an expression for the variance of the estimator $\hat{F}$ of a secondary observable F [171],

$$\sigma^2_{\hat{F}} = \langle (\hat{F} - F)^2 \rangle \approx \sum_{\alpha,\beta} f_\alpha f_\beta C_{\widehat{\alpha\beta}} \approx \sum_{\alpha,\beta} f_\alpha f_\beta \frac{C_{\alpha\beta}}{N} 2\tau_{\alpha\beta,\mathrm{int}}(N), \tag{6.13}$$

where we have neglected terms of $\mathcal{O}\left((\hat{O}_\alpha - \langle O_\alpha \rangle)^3\right)$ and the factor $1 - |t|/N$ as in eq. (6.12). If we define the variance of F,

$$\sigma^2_F = \sum_{\alpha,\beta} f_\alpha f_\beta C_{\alpha\beta}, \tag{6.14}$$

and the integrated autocorrelation time for F,

$$\tau_{F,\mathrm{int}}(N) = \frac{1}{\sigma^2_F} \sum_{\alpha,\beta} f_\alpha f_\beta C_{\alpha\beta} \tau_{\alpha\beta,\mathrm{int}}(N) = \frac{1}{2\sigma^2_F} \sum_{t=-N}^{N} \sum_{\alpha,\beta} f_\alpha f_\beta C_{\alpha\beta} \Gamma_{\alpha\beta}(|t|), \tag{6.15}$$

we can bring eq. (6.13) into the familiar form [171]

$$\sigma^2_{\hat{F}} \approx \frac{\sigma^2_F}{N} 2\tau_{F,\mathrm{int}}(N). \tag{6.16}$$

In order to evaluate the integrated autocorrelation time and thus obtain a value for the variance $\sigma^2_{\hat{F}}$ in the presence of autocorrelation, one needs to find a suitable estimator for the autocorrelation function $\Gamma_{\alpha\beta}$. This is achieved by the Gamma method which we will explain in the next subsection.

A popular alternative which completely avoids the calculation of the autocorrelation function as well as of any derivatives f_α is binning combined with specific techniques for small data sets which can estimate the error of secondary observables. We will introduce these methods, which we use for our main analysis, after the discussion of the Gamma method.

6.1.1 The Gamma method

The Gamma method, which was introduced in Ref. [171], owes its name to the fact that it explicitly estimates the autocorrelation function $\Gamma_{\alpha\beta}(\Delta\hat{t})$. For this purpose, one replaces the unknown true expectation values in eq. (6.10) by the average over configurations,

$$\hat{C}_{\alpha\beta}(\Delta\hat{t}) = \frac{1}{N-\Delta\hat{t}} \sum_{\hat{t}=1}^{N-\Delta\hat{t}} [O_\alpha(\hat{t}) - \hat{O}_\alpha][O_\beta(\hat{t}+\Delta\hat{t}) - \hat{O}_\beta], \tag{6.17}$$

where $\hat{O}_{\alpha,\beta}$ are the usual estimators defined by the average over all N configurations (*cf.*, eq. (3.49)). Because the ensemble mean is used for both the "outer" and the "inner" expectation values in eq. (6.17), this estimator is not unbiased. However, the bias is of $\mathcal{O}(1/N)$ and can hence usually be neglected compared to the statistical error of $\mathcal{O}(1/\sqrt{N})$ [171]. For secondary observables, one additionally needs the Taylor coefficients f_α (*cf.*, eq. (6.7)), which are estimated by evaluating the derivative at the estimators $\hat{O}_\alpha$ instead of at the true expectation values $\langle O_\alpha \rangle$. We will call the thus defined coefficients $\hat{f}_\alpha$.

Using these replacements in eq. (6.4) and eq. (6.13), respectively, one can in principal compute the variances of both primary and secondary observables. The only remaining issue is the truncation of the sum over t in eq. (6.5) and eq. (6.15), respectively. As already mentioned, the statistical error of the integrated autocorrelation time is growing with the summation window W (more concretely, it is proportional to $\sqrt{W/N}$). However, the truncation also induces a bias, *i.e.*, a systematic error, which behaves like $\exp(-W/\tau)$, *i.e.*, it is diminishing with W. Here, τ characterizes the asymptotic exponential decay of $\Gamma_{\alpha\beta}(\Delta\hat{t})$ with $\Delta\hat{t}$. Based on these estimates of the statistical and systematic errors of $\tau_{F,\mathrm{int}}(W)$, an optimal window W has been derived in Ref. [171]. To automatically find the optimal W, one assumes that $\tau \sim S\tau_{F,\mathrm{int}}$ with some factor S and solves for τ. S should be chosen such that $\tau_{F,\mathrm{int}}(W)$ exhibits a plateau at the automatically selected value of W [171].

For the ratios of three- and two-point functions defined in eq. (6.43) below, which are directly related to the form factors we intend to extract, we find that $\tau_{\mathrm{int}} \sim 0.5 - 1$ on our ensembles. This indicates that autocorrelations are largely absent in the quantities we are mainly interested in. For the two- and three-point functions considered individually, this is no longer the case. However, we will always normalize the three-point functions as in eq. (6.43) before proceeding with the analysis, and the individual two-point functions are only used to compute the pion and nucleon masses. Although the latter will serve as input for our main analysis, their errors are completely subdominant compared to the ones of the ratios.

In fig. 6.1, the projected normalized autocorrelation $\rho(\Delta\hat{t}) = \sum_{\alpha,\beta} \hat{f}_\alpha \hat{f}_\beta \hat{\Gamma}_{\alpha\beta}(\Delta\hat{t})$ and the integrated autocorrelation time are shown for the zero-momentum-projected two-point function on D200 at $t_{\mathrm{sep}} = 16a$. Note that we have taken into account a scaling factor of 2 for the computer time here because on D200 only every second configuration is used for the two-point function from which the nucleon mass is extracted (*cf.*, table 5.1). The vertical red line indicates the automatically chosen window $W = 46$. In order to ensure that this value coincides with the onset of the plateau of $\tau_{\mathrm{int}}(W)$, we had to set $S = 3.5$. In the left panel of fig. 6.1, one can nicely see that the onset of the plateau of the integrated autocorrelation time corresponds to the point where the autocorrelation function becomes compatible with zero. As can be read off from the right panel of fig. 6.1, the Gamma method yields $\tau_{\mathrm{int}} = 4.4(1.3)$ for this example of the two-point function, *i.e.*, a clearly non-negligible autocorrelation is present.

This is to be contrasted with the ratio of three- and two-point functions shown in fig. 6.2 ($R_{V_0}^{u+d,\mathrm{conn}}$ at the first non-vanishing momentum transfer on D200, $t_{\mathrm{sep}} = 16a$ and $t = 8a$). Here, the autocorrelation function becomes almost compatible with zero already at $\Delta\hat{t} = 1$.[1] This

[1] Note that $\hat{t}$ has not been scaled here because every configuration is used (*cf.*, table 5.2).

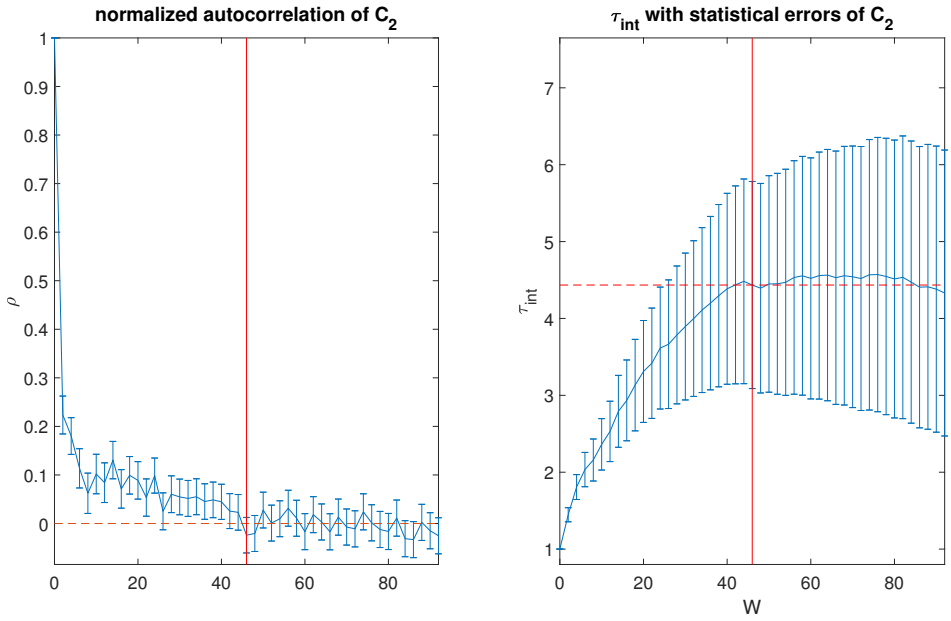


Figure 6.1: Projected normalized autocorrelation and integrated autocorrelation time of $C_2(0; t_{\text{sep}} = 16a)$ on the ensemble D200. The vertical red line indicates the automatically chosen window.

indicates that autocorrelations are barely resolvable. Using the default value of $S = 1.5$, we obtain an optimal window of $W = 4$, coinciding with the first plateau in τ_{int} at $\tau_{\text{int}} = 0.628(55)$.

6.1.2 Binning

The discussion in the previous subsection already suggests that a very sophisticated calculation of the autocorrelation time might not be required for our analysis. Furthermore, the Gamma method relies on a linear approximation of secondary observables around their true expectation value in order to determine the effects of error propagation. A method which does not require such an approximation and avoids the sometimes costly calculation of the autocorrelation function is binning. Here, one divides the raw data into sub-blocks of a specified size and averages the primary quantities first in these bins. The obtained bin averages themselves can then be considered as results of single measurements. If the bin size is large enough, these bin averages can be treated as uncorrelated [83].

To find the optimal bin size, one can study the dependence of the obtained error estimate on the bin size B. In this comparison, the error of the error needs to be taken into account. The relative statistical error of the error due to the finite number of bins can be estimated by [171]

$$\frac{\delta_{\text{stat}}(\sigma_{\hat{F},\text{bin}})}{\sigma_{\hat{F},\text{bin}}} = \sqrt{\frac{2B}{N}}, \tag{6.18}$$

i.e., it is rising with the bin size because a larger B means that fewer bins are available to estimate $\sigma_{\hat{F}}$. In practice, at a given number of available configurations, one needs to balance the improved removal of autocorrelation with the growing error of the error.

An example for such a binning analysis is shown in fig. 6.3 for the two observables corresponding to figs. 6.1 and 6.2. Here, we have estimated the errors using Jackknife resampling (*cf.*, section 6.1.4 below). For the two-point function (left panel), the error tends to increase

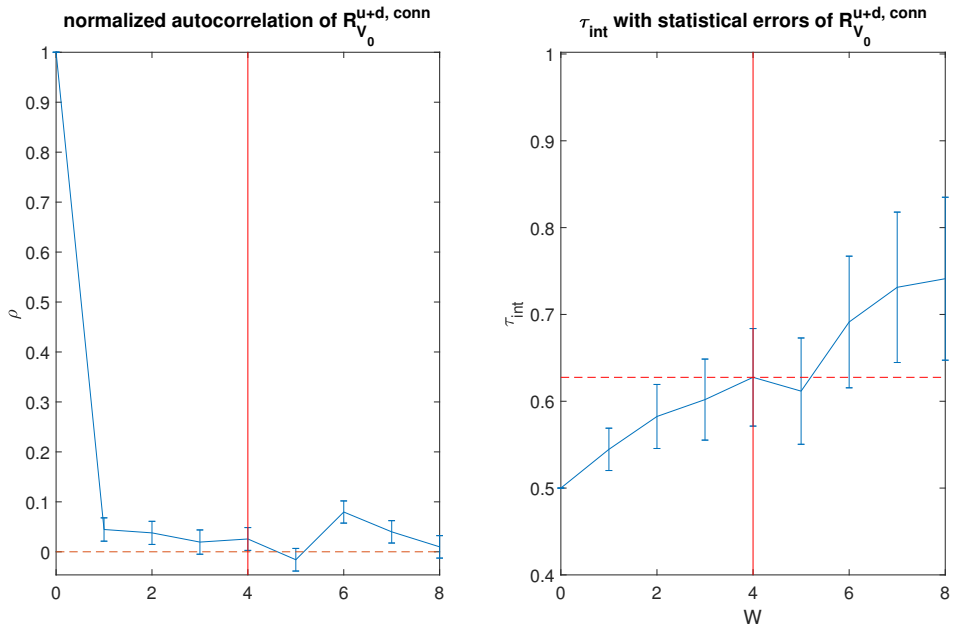


Figure 6.2: Projected normalized autocorrelation and integrated autocorrelation time of $R_{V_0}^{\mathrm{conn},u+d}(1; t_{\mathrm{sep}} = 16a, t = 8a)$ on the ensemble D200. The vertical red line indicates the automatically chosen window.

with the bin size until an approximate plateau is reached, while the error of the error keeps growing. The error increase is due to the presence of autocorrelations which are successively removed by considering larger bins. For the ratio (right panel), on the other hand, there is barely any significant error increase visible. Note that we have also scaled the horizontal axis with a factor of 2 for the two-point function here in order to enable a like-by-like comparison with the ratio. In practice, the data for the two-point function are binned with half the bin sizes shown in the plot.

Since the picture is very similar to the right panel of fig. 6.3 for the ratios on all ensembles used in this thesis, we generally employ a bin size of 2, if the spacing between two analyzed configurations in terms of molecular dynamics time does not already account for this factor. The latter is the case for the disconnected contribution on C101 and D200, for the connected contribution on E250, and for both the connected and the disconnected contributions on E300 and J303 (*cf.*, tables 5.2 and 5.3; for J303, also note the comment at the end of section 4.2.5). In order to enable a consistent Jackknife analysis, we employ the same bin sizes also for the determination of the nucleon and pion masses from the corresponding two-point functions even though more autocorrelation is present in these observables. As already mentioned, the errors of the masses are completely subdominant compared to the ones of the ratios, so that this choice does not significantly influence the error estimates of our final results.

6.1.3 Bootstrap resampling

Once the data have been binned appropriately and can be considered as uncorrelated, the central values and error estimates of the observables of interest can be determined. For secondary observables, however, this is not straightforward since they are defined as functions of the averages of primary observables. To determine the effect of error propagation and thus obtain a reasonable estimate for the error of a secondary observable without using an approximation

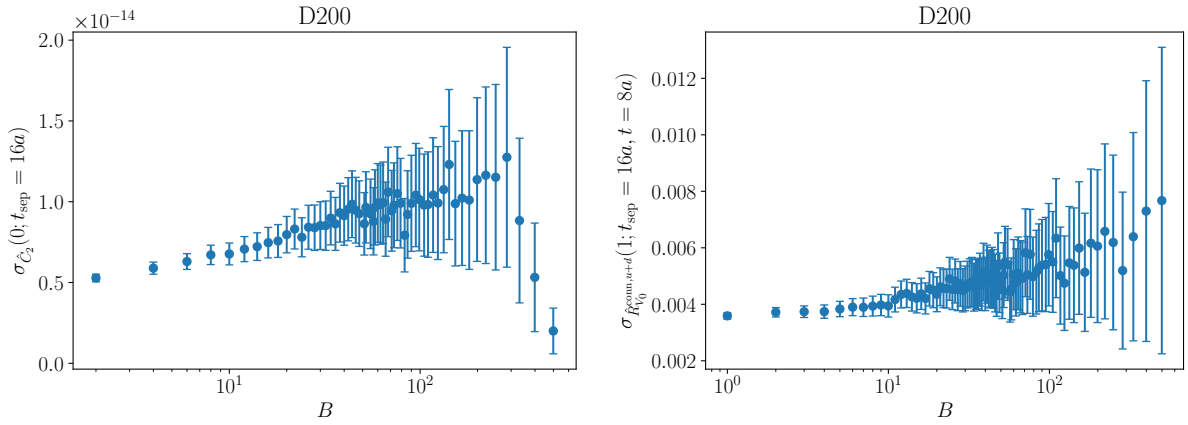


Figure 6.3: Error of $C_2(0; t_{\text{sep}} = 16a)$ (left panel) and $R_{V_0}^{\text{conn},u+d}(1; t_{\text{sep}} = 16a, t = 8a)$ (right panel) on the ensemble D200 as a function of the bin size

as in the Gamma method, one can use the technique of bootstrap resampling introduced in Ref. [172].

Consider a size-N sample of primary observables drawn from an unknown probability distribution P, $\mathbf{O}_\alpha = (O_\alpha(1), O_\alpha(2), \ldots, O_\alpha(N))$. This sample gives rise to an empirical distribution $\hat{P}$: it is the discrete distribution that puts probability $1/N$ on each value $O_\alpha(i)$, $i = 1, \ldots, N$. A bootstrap sample is now defined as a random sample of size N drawn from $\hat{P}$, and we will denote it by $\mathbf{O}^*_\alpha = (O^*_\alpha(1), O^*_\alpha(2), \ldots, O^*_\alpha(N))$. In other words, the bootstrap data points $O^*_\alpha(1), O^*_\alpha(2), \ldots, O^*_\alpha(N)$ are obtained as a random sample of size N drawn with replacement from the population $(O_\alpha(1), O_\alpha(2), \ldots, O_\alpha(N))$ [173]. Based on the bootstrap sample of the primary observables, we can also define the bootstrap replication of a secondary observable F as follows [173],

$$\hat{F}^* = f(\hat{O}^*_\alpha), \quad \hat{O}^*_\alpha = \frac{1}{N}\sum_{i=1}^{N} O^*_\alpha(i). \tag{6.19}$$

In total, there are $\binom{2N-1}{N}$ distinct bootstrap samples [173], a huge number even if N is only moderately large. Since it is practically impossible to compute all these samples, one randomly draws $\mathcal{B}$ independent bootstrap samples and evaluates the secondary observables on them. The variance of $\hat{F}$ can then be estimated by

$$\left(\sigma^{(B)}_{\hat{F}}\right)^2 = \frac{1}{\mathcal{B}}\sum_{b=1}^{\mathcal{B}}\left[\hat{F}^*(b) - \hat{F}\right]^2. \tag{6.20}$$

Similarly, for K secondary observables F_k, $k = 1, \ldots, K$, one can compute the covariance matrix of the estimators $\hat{F}_k$. The entry of this matrix with the indices k and k' is given by

$$C^{(B)}_{\widehat{kk'}} = \frac{1}{\mathcal{B}}\sum_{b=1}^{\mathcal{B}}\left[\hat{F}^*_k(b) - \hat{F}_k\right]\left[\hat{F}^*_{k'}(b) - \hat{F}_{k'}\right]. \tag{6.21}$$

A major advantage of the bootstrap is that it can be straightforwardly applied to arbitrarily complicated secondary observables without the need to keep track of derivatives as for the Gamma method. A particularly important example are fits (*cf.*, section 6.2): one simply performs the fit on each bootstrap sample $b \in \{1, \ldots, \mathcal{B}\}$ to obtain the bootstrap replica $\hat{F}^*(b)$

of the corresponding derived observable. It is not required to differentiate through the fitting procedure or perform any other theoretical calculations.

As derived in eq. (6.8), $\hat{F}$ is not unbiased as an estimator for F unless the function f is linear. The bias can be estimated from the bootstrap samples as the difference between the mean over the bootstrap replica and the mean over all data [173],

$$\langle \hat{F} - F \rangle^{(B)} = \frac{1}{\mathcal{B}} \sum_{b=1}^{\mathcal{B}} \hat{F}^*(b) - \hat{F}, \tag{6.22}$$

which clearly vanishes if f is linear. In principal, one could use eq. (6.22) to replace $\hat{F}$ by a bias-corrected estimator [173],

$$\hat{F}_{\text{unbiased}} = \hat{F} - \langle \hat{F} - F \rangle^{(B)} = 2\hat{F} - \frac{1}{\mathcal{B}} \sum_{b=1}^{\mathcal{B}} \hat{F}^*(b). \tag{6.23}$$

In practice, however, the error of the bias is usually much larger than the error of the error at the same number $\mathcal{B}$ of bootstrap samples. Thus, correcting for the bias might significantly increase the variance of our estimator [173]. For this reason, we do not apply the bias correction eq. (6.23) in our analysis. A better approach is to compute the bias estimate eq. (6.22), but not apply it to the final result. If the obtained bias is small compared to the error, as is the case for the observables considered in this thesis, it is safe to use the uncorrected $\hat{F}$. If a large bias is found, this is usually an indication that the chosen statistics is not an appropriate estimate of the parameter one is actually interested in [173]. In practice, this can also mean that (part of) the data-analysis procedure is not statistically sound and should be revised.

6.1.4 Jackknife resampling

An alternative to bootstrap is Jackknife resampling which was first introduced for bias correction in Ref. [174] and for error estimation in Ref. [175]. The general idea is similar to the bootstrap, but one focuses on the samples that leave out one observation at a time. Accordingly, the j-th Jackknife sample is constructed by removing the j-th observation from the original data set, $\mathbf{O}_\alpha^{(J)}(j) = (O_\alpha(1), \ldots, O_\alpha(j-1), O_\alpha(j+1), \ldots, O_\alpha(N))$ [173]. The j-th Jackknife replication of the primary observable O_α is hence given by [80],

$$\hat{O}_\alpha^{(J)}(j) = \frac{1}{N-1} \sum_{i \neq j} O_\alpha(i). \tag{6.24}$$

Analogously to the bootstrap, the j-th Jackknife replication of a secondary observable F is calculated as $\hat{F}^{(J)}(j) = f(\hat{O}_\alpha^{(J)}(j))$ [173].

If the sample size of the original data set is N, there are exactly N distinct Jackknife samples. Therefore, it is perfectly possible to compute all of them and estimate the variance of $\hat{F}$ as [83, 171],

$$\left(\sigma_{\hat{F}}^{(J)}\right)^2 = \frac{N-1}{N} \sum_{j=1}^{N} \left[\hat{F}^{(J)}(j) - \hat{F}\right]^2. \tag{6.25}$$

Covariance matrices are analogously estimated by

$$C_{\widehat{kk'}}^{(J)} = \frac{N-1}{N} \sum_{j=1}^{N} \left[\hat{F}_k^{(J)}(j) - \hat{F}_k\right] \left[\hat{F}_{k'}^{(J)}(j) - \hat{F}_{k'}\right]. \tag{6.26}$$

The Jackknife estimate of bias is given by [173, 174],

$$\langle \hat{F} - F \rangle^{(J)} = (N-1) \left[\frac{1}{N} \sum_{j=1}^{N} \hat{F}^{(J)}(j) - \hat{F} \right]. \tag{6.27}$$

However, we do not correct our final results for this bias for the reasons given in our discussion of the bootstrap bias estimate above.

Comparing eqs. (6.25) to (6.27) to their analogues for the bootstrap, eqs. (6.20) to (6.22), one notices a common inflation factor of $(N-1)$. Intuitively, this is required because the Jackknife samples are more similar to the original data set than typical bootstrap samples. For a stricter derivation of this factor, we refer to Ref. [173]. The closer similarity to the original data set also represents one of the major advantages of the Jackknife over the bootstrap: the occurrence of numerical issues is much less likely. For our analysis, this will become important due to the square root in the ratio eq. (6.43). The probability that the argument of the square root becomes negative due to statistical fluctuations is significantly smaller for Jackknife than for bootstrap resampling. For this reason, we generally use Jackknife instead of bootstrap resampling in our analysis. Still, we have exemplarily checked that for low momenta, where the fluctuations in our data are small enough to enable a bootstrap analysis without numerical issues, the errors obtained from bootstrap resampling agree very well with those obtained from Jackknife resampling.

Similar to the derivation of the $(N-1)$-factor, one can show that the Jackknife estimates for the error (bias) agree with the bootstrap ones for linear (quadratic) functions f[2] [173]. For nonlinear functions, the Jackknife can essentially be thought of as a linear approximation of the bootstrap [173, 175].

An important shortcoming of the Jackknife is that it is only applicable to smooth, *i.e.*, differentiable, functions f. For non-smooth functions, the Jackknife is inconsistent, *i.e.*, the Jackknife error estimate eq. (6.25) does not converge to the true error $\sigma_{\hat{F}}$ of $\hat{F}$ for $N \to \infty$ [173]. For example, it is not possible to replace the mean by the median in a Jackknife analysis, while in a bootstrap analysis, this is in principal possible.

Both bootstrap and Jackknife resampling can be combined with binning by first organizing the data for the primary observables in blocks of length B and treating the block averages as entries of the original data set $\mathbf{O}_\alpha$. For the Jackknife, this is equivalent to deleting entire blocks of B consecutive data points each in eq. (6.24) instead of only single values,

$$\frac{1}{N/B - 1} \sum_{\substack{i=1 \\ i \neq j}}^{N/B} \frac{1}{B} \sum_{l=1}^{B} O_\alpha(Bi + l) = \frac{1}{N - B} \sum_{i \notin \{Bj, \ldots, B(j+1)-1\}}^{N} O_\alpha(i). \tag{6.28}$$

As a final remark, we note that there exist different conventions in the literature for the bootstrap and Jackknife (co-)variances: one can either subtract the mean over all data [83, 171], as we have done, or the mean over the bootstrap/Jackknife replica [80, 173]. The difference is exactly of the order of the bias which we have found to be small compared to the error of our observables.

6.1.5 The parametric bootstrap

As explained above, we need to use Jackknife instead of bootstrap resampling for the analysis of the ratios in order to suppress numerical issues due to statistical fluctuations. This, however,

[2] up to a definitional factor of $\sqrt{(N-1)/N}$ which is negligible for large enough N.

comes at the price of having a different number of resamples on each ensemble because the number of configurations differs between ensembles (*cf.*, table 4.1). For this reason, it is impossible to perform a global analysis across all ensembles within the Jackknife framework.

To circumvent this issue, we generate parametric bootstrap samples on each ensemble before moving to the global step of the analysis. In the parametric bootstrap, instead of $\mathcal{B}$ times drawing N samples with replacement from the empirical distribution $\hat{P}$ and each time averaging over these N points, one draws $\mathcal{B}$ samples from a parametric estimate $\hat{P}_{\text{par}}$ of the distribution [173]. These are then treated in the same way as the $\hat{F}_k^*(b)$. In our case, we assume on each ensemble $\hat{P}_{\text{par}}$ to be a (multivariate) Gaussian with central values $\hat{F}_k$ and the covariance matrix $C^{(J)}_{\widehat{kk'}}$ estimated from the original Jackknife samples according to eq. (6.26). For data derived from a Monte-Carlo simulation, this is in general a good assumption due to the central limit theorem [173]. We have also visually checked that the histograms of the marginal Jackknife distributions indeed look Gaussian.

From the parametric bootstrap samples, one can again build a covariance matrix according to eq. (6.21) and compare it to the original covariance matrix of the Jackknife samples. To sufficiently accurately reproduce also the smaller elements of the original covariance matrix, many more bootstrap samples are required than for a single observable, where one just needs to reproduce two numbers (central value and variance). In this thesis, we employ 10 000 parametric bootstrap samples and thus arrive at a bootstrap covariance matrix which is consistent with the original (Jackknife) one.

6.2 Regression analysis

A common problem in Lattice-QCD analyses is regression, *i.e.*, finding the parameter values of a certain functional form that best fit a given set of data. In the following, we will first define the corresponding problem and introduce general methods to solve it. Next, we will explain the VarPro method which can facilitate the fit if the function is given as a linear combination of nonlinear functions. Finally, we will introduce the concept of Bayesian priors which can be used to constrain and stabilize fits.

6.2.1 Least-squares fits

Let y_i, $i = 1, \ldots, m$ be data measured as a function of an independent variable t with corresponding values t_i. The covariance matrix of the data, which we will denote by C_{ij} in the following, can for example be estimated by bootstrap or Jackknife resampling (*cf.*, eqs. (6.21) and (6.26), respectively). Consider now a function $f(t; \mathbf{a})$ with some free parameters a_k, $k = 1, \ldots, n$ as a model for the data point y corresponding to the value t of the independent variable. The aim of a correlated least-squares fit is to find the parameters a_k^*, $k = 1, \ldots, n$ which minimize the correlated sum of squared deviations between data and model,

$$\chi^2(\mathbf{a}^*) = \min_{\mathbf{a}} \sum_{i,j=1}^{m} [y_i - f(t_i; \mathbf{a})] C_{ij}^{-1} [y_j - f(t_j; \mathbf{a})] = \min_{\mathbf{a}} \sum_{i=1}^{m} d_i^2(\mathbf{a}). \tag{6.29}$$

Here, we have defined the generalized deviations

$$d_i(\mathbf{a}) = \sum_{j=1}^{m} [y_j - f(t_j; \mathbf{a})] L_{C^{-1},ji}, \tag{6.30}$$

where $L_{C^{-1}}$ is the Cholesky decomposition of C^{-1}, defined such that $L_{C^{-1}} L_{C^{-1}}^T = C^{-1}$. The Cholesky decomposition is guaranteed to exist because the covariance matrix and thus also its inverse are by definition symmetric and positive definite.

Many methods have been developed over the years and are available for the numerical solution of least-squares problems with nonlinear functions f. In this thesis, we use the `optimize` module of `SciPy` [176] to interface either the Levenberg-Marquardt algorithm [177, 178] or the Trust Region Reflective algorithm [179]. Both of these algorithms find local minima of the residual $\chi^2(\mathbf{a})$. If the function f is linear in the fit parameters $\mathbf{a}$, a solution to the problem eq. (6.29) can be derived analytically. Nevertheless, we also solve linear systems numerically since the numerical algorithms quoted above are very efficient and stable for the solution of linear systems.

The residual sum of squares at the best-fit point, $\chi^2(\mathbf{a}^*)$, can serve as an indication how well the function f is able to describe the data. If the data has indeed been generated from a process following the function f, any deviations should be purely statistical in nature. Due to the central limit theorem, the y_i would then follow a multivariate Gaussian distribution around the values $f(t_i; \mathbf{a}^*)$. In this case, the residual sum of squares $\chi^2(\mathbf{a}^*)$ would follow a χ^2-distribution with $m - n$ degrees of freedom (hence the name). To judge the quality of fit, one can thus perform a χ^2-test. Its p-value (the probability to obtain a value for $\chi^2(\mathbf{a}^*)$ at least as large as the actually observed one under the assumption that the above hypothesis is correct) is given by

$$p = 1 - \mathrm{cdf}_{\chi^2}(\chi^2(\mathbf{a}^*), m - n). \tag{6.31}$$

Here, $\mathrm{cdf}_{\chi^2}(x, m - n)$ is the cumulative distribution function of the χ^2-distribution with $m - n$ degrees of freedom. If the p-value is smaller than a predefined threshold (5 % and 1 % are common values), one has to reject the hypothesis that the data has been generated from a process described by the fit function f. On the contrary, if the p-value is close to 1, this indicates overfitting since the data fluctuate much less than expected from statistical reasons around the values predicted by the model. For the one-sided test defined in eq. (6.31), the p-value should hence ideally be around 1/2.

In practice, it happens frequently that the covariance matrix is badly estimated. In this case, performing a fully correlated fit can lead to artefacts in the results for the fit parameters. To circumvent this issue, one can either regularize the covariance matrix in some way, *e.g.*, by damping the off-diagonal elements or by imposing a cut on the singular values, or directly perform an uncorrelated fit. The latter corresponds to using $\tilde{C}_{ij} = C_{ii}\delta_{ij}$ (no sum over i) instead of C_{ij} in eq. (6.29) [83]. One has to be aware, however, that any artificial change of the covariance matrix renders the residual sum of squares $\chi^2(\mathbf{a}^*)$ and thus the p-value meaningless as a measure for the fit quality. A modified fit-quality criterion based on the expectation value of the modified $\chi^2(\mathbf{a}^*)$ has been derived in Ref. [180], but is not used in this thesis because we generally employ fully correlated fits.

Regression analysis can be combined with bootstrap or Jackknife resampling by performing the fit on each data sample to obtain sample values for the best-fit parameters, as mentioned in section 6.1.3 above. The covariance matrix, on the other hand, is kept fixed between samples.

6.2.2 The VarPro method

If the model function can be written as a linear combination of nonlinear functions, the VarPro method [181] can be used to simplify and stabilize the numerical minimization procedure. In this context, one considers model functions of the form

$$f(t_i; \mathbf{a}, \boldsymbol{\alpha}) = \sum_{k=1}^{p_\mathrm{l}} a_k \phi_k(t_i; \boldsymbol{\alpha}), \tag{6.32}$$

i.e., f depends linearly on the coefficients a_k, $k = 1, \ldots, p_\mathrm{l}$ and nonlinearly on the parameters α_k, $k = 1, \ldots, p_\mathrm{nl}$. The total number of fit parameters is thus $n = p_\mathrm{l} + p_\mathrm{nl}$. If we define the

matrix of nonlinear model functions $\Phi_{ij}(\boldsymbol{\alpha}) = \phi_j(t_i;\boldsymbol{\alpha})$, we can write the correlated sum of squared deviations as

$$\chi^2(\mathbf{a},\boldsymbol{\alpha}) = [\mathbf{y} - \Phi(\boldsymbol{\alpha})\mathbf{a}]^T C^{-1}[\mathbf{y} - \Phi(\boldsymbol{\alpha})\mathbf{a}] = \|[\mathbf{y} - \Phi(\boldsymbol{\alpha})\mathbf{a}]^T L_{C^{-1}}\|^2. \tag{6.33}$$

At a fixed value of $\boldsymbol{\alpha}$, minimizing eq. (6.33) as a function of $\mathbf{a}$ is a linear problem and can as such be solved analytically. To find the solution $\mathbf{a}^* = \arg\min_{\mathbf{a}} \chi^2(\mathbf{a},\boldsymbol{\alpha})$, we compute the gradient of $\chi^2(\mathbf{a},\boldsymbol{\alpha})$ with respect to $\mathbf{a}$ and set it to zero,

$$\begin{aligned}\nabla_{\mathbf{a}}\chi^2(\mathbf{a},\boldsymbol{\alpha})|_{\mathbf{a}=\mathbf{a}^*} &= -\left[\Phi^T(\boldsymbol{\alpha})C^{-1}[\mathbf{y} - \Phi(\boldsymbol{\alpha})\mathbf{a}^*]\right]^T - [\mathbf{y} - \Phi(\boldsymbol{\alpha})\mathbf{a}^*]^T C^{-1}\Phi(\boldsymbol{\alpha}) \\ &= -2[\mathbf{y} - \Phi(\boldsymbol{\alpha})\mathbf{a}^*]^T C^{-1}\Phi(\boldsymbol{\alpha}) = 0.\end{aligned} \tag{6.34}$$

Here, we have made use of the fact that C^{-1} is symmetric. Equation (6.34) can only be fulfilled if

$$\mathbf{y}^T C^{-1}\Phi(\boldsymbol{\alpha}) = (\mathbf{a}^*)^T \Phi^T(\boldsymbol{\alpha})C^{-1}\Phi(\boldsymbol{\alpha}), \tag{6.35}$$

or, what is equivalent because C^{-1} is symmetric,

$$\Phi^T(\boldsymbol{\alpha})C^{-1}\mathbf{y} = \Phi^T(\boldsymbol{\alpha})C^{-1}\Phi(\boldsymbol{\alpha})\mathbf{a}^*. \tag{6.36}$$

Hence, the sought solution must be given by

$$\mathbf{a}^* = [\Phi^T(\boldsymbol{\alpha})C^{-1}\Phi(\boldsymbol{\alpha})]^{-1}\Phi^T(\boldsymbol{\alpha})C^{-1}\mathbf{y}. \tag{6.37}$$

This can now be plugged into eq. (6.33) to define a modified residual which only depends on $\boldsymbol{\alpha}$.

In Ref. [181], it has been shown that, roughly speaking, $(\mathbf{a}^*, \boldsymbol{\alpha}^*)$ is a critical point of χ^2 if and only if $\boldsymbol{\alpha}^*$ is a critical point of the modified residual function and $\mathbf{a}^*$ satisfies eq. (6.37). Therefore, one only needs to solve numerically for the nonlinear parameters $\boldsymbol{\alpha}$, while the corresponding linear coefficients can be calculated exactly using eq. (6.37). This eliminates in particular the need for initial guesses for the coefficients $\mathbf{a}$. Since nonlinear functions are not protected against having multiple local minima and numerical optimizers can in general only find local solutions, whether one obtains the desired solution to a nonlinear fit can crucially depend on the initial guesses for the fit parameters. In such a situation, not depending on initial guesses at least for part of the parameters can help to stabilize the fit.

6.2.3 Bayesian priors

Another issue that can also occur for sufficiently complicated linear fits is that some of the parameters might not be well constrained by the given data. This can render the whole fit, *i.e.*, also the results for the other parameters, unstable. Under these circumstances, it is desirable to use *a priori* knowledge about the badly constrained parameters to stabilize the fit. If firm and precise knowledge about the values of these parameters exists, the known values can directly be plugged into the model and thus fixed. A much more common situation is, however, that one only has a rough and/or uncertain prior knowledge about the parameter values.

In this case, one can use Bayesian priors to constrain the fit. The method has been suggested in Ref. [182] and is motivated by Bayes' theorem,

$$P(\mathbf{a}|\mathbf{y}) = \frac{P(\mathbf{y}|\mathbf{a})P(\mathbf{a})}{P(\mathbf{y})} \propto P(\mathbf{y}|\mathbf{a})P(\mathbf{a}). \tag{6.38}$$

The denominator is the probability to obtain the data $\mathbf{y}$ from any model; it is thus independent of the parameters $\mathbf{a}$ and can be neglected. The first term of the numerator is the probability to obtain the data $\mathbf{y}$ given the parameter values $\mathbf{a}$ and can, in the context of least-squares fits,

be calculated as $P(\mathbf{y}|\mathbf{a}) = \exp[-\chi^2(\mathbf{a})/2]$ with $\chi^2(\mathbf{a})$ as in eq. (6.29). The second term of the numerator, $P(\mathbf{a})$, quantifies the prior probability of the parameters.

In practice, it is unusual to have *a priori* knowledge about the full distribution of the parameters. One rather wants to specify only the central values and corresponding error margins, *i.e.*, one assumes $a_k \approx \tilde{a}_k \pm \tilde{\sigma}_{a_k}$. It has been argued in Ref. [182] that in such a situation, the least biased prior distribution one can assume is a Gaussian,

$$P(\mathbf{a}) = e^{-\chi^2_{\text{prior}}(\mathbf{a})/2}, \quad \chi^2_{\text{prior}}(\mathbf{a}) = \sum_{k=1}^{n} \frac{(a_k - \tilde{a}_k)^2}{\tilde{\sigma}^2_{a_k}}. \tag{6.39}$$

This leads to the posterior distribution $P(\mathbf{a}|\mathbf{y}) \propto \exp[-\chi^2_{\text{aug}}(\mathbf{a})/2]$, where the augmented residual is defined as $\chi^2_{\text{aug}}(\mathbf{a}) = \chi^2(\mathbf{a}) + \chi^2_{\text{prior}}(\mathbf{a})$, *i.e.*, one simply adds the prior term to the usual residual. Of course, one does not need to set priors on *all* parameters. One can also choose to set priors only on some of them (*e.g.*, the ones which are least constrained by the data) by restricting the sum in eq. (6.39) to these parameters.

A full Bayesian analysis would require the calculation of the expectation values of the observables of interest under the posterior distribution $P(\mathbf{a}|\mathbf{y})$. These take the form of integrals over parameter space. In practice, however, these integrals are quite costly and difficult to evaluate because the posterior distribution is typically very sharply peaked around its maximum [182]. In lattice field theory analyses, it is therefore common to just quote the posterior mode $\mathbf{a}^*_{\text{PM}} = \arg\max_{\mathbf{a}} P(\mathbf{a}|\mathbf{y})$ which can be obtained by minimizing $\chi^2_{\text{aug}}(\mathbf{a})$ with the same methods that one uses to minimize $\chi^2(\mathbf{a})$ in a conventional least-squares fit. Uncertainties are then quantified in the usual way by bootstrap or Jackknife resampling. To incorporate the effects of the priors into the bootstrap or Jackknife distribution of the parameters $\mathbf{a}$, one has to use resamples also for the central values of the priors $\tilde{\mathbf{a}}$. If the priors are obtained from a previous analysis (*e.g.*, a fit to different [subsets of] data constraining the same physical parameters), the resamples outputted by it can simply be used as sample values for the priors[3]. If the prior central values are set by hand (*e.g.*, to 0), one can resample them using a procedure analogous to the parametric bootstrap explained in section 6.1.5, as suggested in Ref. [182].

6.3 Exceptional configurations

On some ensembles, we observe that individual measurements on a small number of configurations are located very far outside the distribution of the vast majority of configurations. The appearance of these outliers is likely to be connected to configurations with localized low modes of the Dirac operator which are picked up by our point sources [83, 124]. Such exceptionally small eigenvalues of the Dirac operator can occur because the Wilson Dirac operator breaks chiral symmetry explicitly. In this context, we refer to the discussion at the beginning of section 4.2.2, where a twisted-mass regulator was introduced in order to mitigate exactly this issue. Nevertheless, the probability that exceptionally small eigenvalues occur is still nonzero, in particular at low quark masses and coarse lattice spacings. Similar observations of outliers have already been reported for previous analyses on CLS ensembles [183, 184].

Normally, one would expect a Gaussian distribution for the Jackknife samples due to the central limit theorem. The outliers make the actual distribution deviate significantly from this expected shape. Keeping the exceptional configurations in the sample thus leads to a drastically increased error and, more importantly, to an unexpected scaling of the error with the source-sink separation, *i.e.*, the error is in many cases inflated strongly only for single values

[3] provided that the bootstrap resampling of the data entering the two fits is done consistently, *i.e.*, the same seed for the random-number generator is used for observables correlated between the two fits.

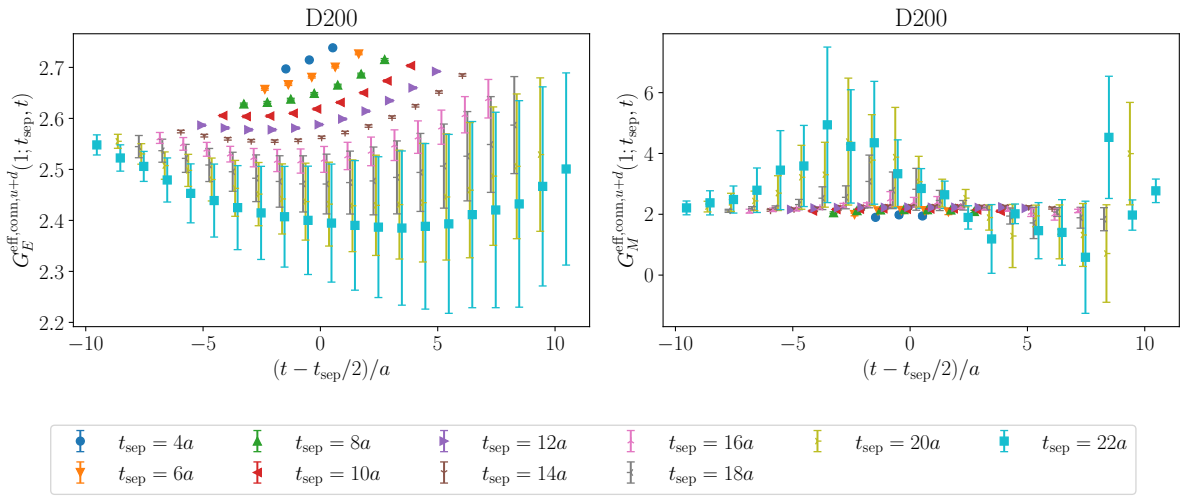


Figure 6.4: Connected contribution to the isoscalar effective form factors at the first non-vanishing momentum on the ensemble D200 ($Q^2 \approx 0.091$ GeV^2). The data have been extracted from all 2000 gauge configurations using Jackknife resampling without binning. The left plot shows the electric and the right one the magnetic form factor. For each source-sink separation t_{sep}, the effective form factors are displayed as a function of the operator insertion time t, offset to the midpoint between nucleon source and sink. The data points are horizontally displaced for better visibility.

of t_{sep}. The most prominent example is D200, where we identify one configuration (r000n757) to be the root cause of the gross overestimation of errors.

This is illustrated in figs. 6.4 and 6.5 for the connected contribution to the isoscalar effective form factors at the first non-vanishing momentum transfer. In fig. 6.4, the data extracted from all 2000 gauge configurations are shown. It is clearly visible that in particular for the source-sink separations corresponding to the last iteration ($t_{\text{sep}} = 16 - 22a$, *cf.*, table 5.2) the statistical fluctuations and errors are drastically enhanced. This suggests that mostly some of the point sources added in the last iteration are problematic.

The origin of this behavior is studied in more detail in fig. 6.5, where we show the Jackknife distributions at $t_{\text{sep}} = 16a$. No binning has been applied to the data displayed in figs. 6.4 and 6.5, so that each Jackknife sample corresponds to leaving out exactly one configuration. Obviously, the sample where the configuration r000n757 is left out deviates strongly from the remaining 1999 samples. In the interpretation of fig. 6.5, one has to keep in mind that Jackknife samples are shown, not configurations. This implies that the majority of *configurations* follow the curve highlighted in blue, not the black curves.

In order to identify the problematic configurations, we first extract the effective form factors (*cf.*, section 6.5 below) using Jackknife resampling on unbinned data. We then scan the isovector, connected isoscalar and disconnected contributions for all relevant values of Q^2, t_{sep} and t, employing the procedure introduced in the supplementary material of Ref. [183]. All configurations for which the corresponding Jackknife samples are more than $6\,\sigma$ removed from the central value are omitted during the whole main analysis. The numbers in tables 4.1 and 5.1 to 5.4 already reflect this.

It is well known that the mean and standard deviation as estimators for the central value and width of a distribution, respectively, are very sensitive to outliers. Therefore, we need

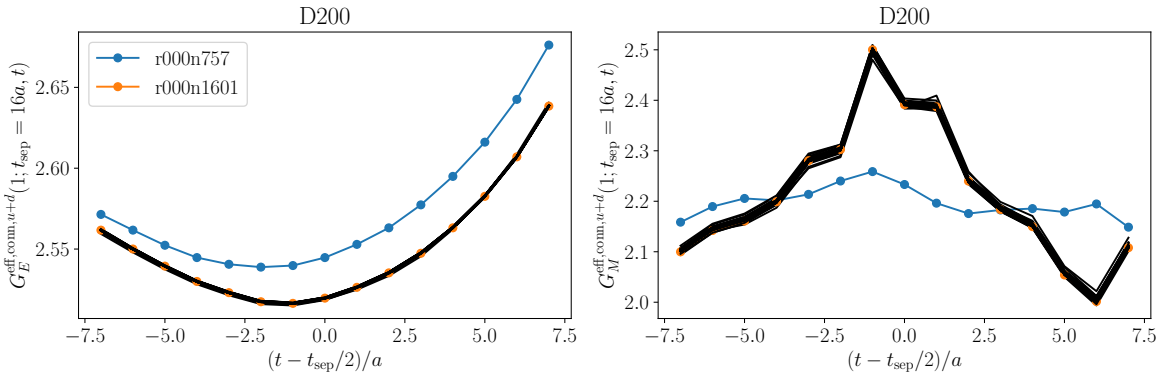


Figure 6.5: Jackknife distribution of the connected contribution to the isoscalar effective form factors at the first non-vanishing momentum and $t_{\mathrm{sep}} = 16a$ on the ensemble D200. No binning has been applied, so that each Jackknife sample corresponds to leaving out exactly one configuration. Highlighted are the samples where one of the flagged configurations is left out.

to substitute them by more robust estimators for the purpose of finding the exceptional configurations. We use the median, *i.e.*, the value separating the higher half from the lower half of the population, as a replacement for the mean, and the median absolute deviation,

$$\mathrm{MAD}(y) = \mathrm{med}(|y - \mathrm{med}(y)|), \tag{6.40}$$

as a replacement for the standard deviation [183]. In order to make contact with the standard deviation of a Gaussian distribution, we need to multiply the MAD by a constant scale factor. This has to be chosen such that ±MAD covers the central 50 % of the cumulative distribution function of a Gaussian distribution,

$$\begin{aligned}\frac{1}{2} &= P(|y-\hat{y}| \leq \mathrm{MAD}) = P\left(\left|\frac{y-\hat{y}}{\sigma}\right| \leq \frac{\mathrm{MAD}}{\sigma}\right) = \mathrm{cdf}\left(\frac{\mathrm{MAD}}{\sigma}\right) - \mathrm{cdf}\left(-\frac{\mathrm{MAD}}{\sigma}\right) \\ &= \mathrm{cdf}\left(\frac{\mathrm{MAD}}{\sigma}\right) - \left[1 - \mathrm{cdf}\left(\frac{\mathrm{MAD}}{\sigma}\right)\right] = 2\,\mathrm{cdf}\left(\frac{\mathrm{MAD}}{\sigma}\right) - 1.\end{aligned} \tag{6.41}$$

It follows that $\mathrm{MAD}/\sigma = \mathrm{cdf}^{-1}(3/4)$ or $\sigma = \mathrm{MAD}/\,\mathrm{cdf}^{-1}(3/4) \approx 1.4826\,\mathrm{MAD}$. Additionally, we multiply the thus obtained value for σ by the factor $\sqrt{N-1}$ needed to transform the width of the Jackknife distribution to that of the distribution of the original data.

Using the method described above, the following configurations get flagged as exceptional and are consequently removed from the analysis,

C101 r014n293, r014n427, r014n645, r014n793, r014n794, r014n1655, r014n1787

N101 r001n22, r001n71, r003n300, r003n390, r004n203, r006n271

H105 r001n536, r001n926

D450 r010n208, r010n250

E250 r001n49, r001n345

D200 r000n757, r000n1601

N200 r000n466, r001n327, r001n537

On the ensembles not listed here, we do not find any exceptional configurations. In particular, there are no exceptional configurations on the ensembles with our finest lattice spacing corresponding to $\beta = 3.70$.

6.4 Hadron masses

The masses of the nucleon and pion on our ensembles are required as input for the further analysis: The nucleon mass enters the kinematic prefactors in the extraction of the effective form factors (*cf.*, section 6.5) as well as the definition of Q^2 (*cf.*, eq. (5.57)). The pion mass, on the other hand, is needed for the extra- (or rather inter-)polation to the physical point because it is employed as a proxy for the light-quark mass. Moreover, it will be used to set the parameter τ_{cut} in the z-expansion fits (*cf.*, section 6.7.2).

In order to enable a consistent and (as far as possible) self-contained analysis, we need to obtain Jackknife samples for the hadron masses which can subsequently serve as input for our main analysis of the form factors. Details on how this is done are provided in the following.

6.4.1 The nucleon mass

The nucleon mass is determined from the two-point functions eq. (5.20) at zero momentum. For this purpose, we use the unpolarized nucleon, *i.e.*, $\Gamma^p = {}^1\!/_2(1 + \gamma_0)$ (the polarization term proportional to $\gamma_5\gamma_j$ does not contribute to the spectral representation of the two-point function and would thus only add unnecessary noise), and the highest available statistics in terms of sources (*cf.*, table 5.1). Details on the employed set of sources, also in the context of the averaging over the forward- and backward-propagating nucleon, can be found in section 5.3.5.

The two-point functions are fitted to a single-exponential *ansatz* $Ae^{-mt_{\text{sep}}}$ as motivated by eq. (5.29). For the fits, we use the VarPro method (*cf.*, section 6.2.2) which only needs an initial guess for the mass m. Our preferred fit range is determined by scanning over all possible fit ranges with a minimal length of 0.8 fm and choosing the one with the smallest Akaike Information Criterion (AIC) [185, 186]. This criterion is also employed for our final model averages and discussed in greater detail at the beginning of section 6.10 below. In order to use the AIC to compare different fit ranges, it needs to be modified to the AIC_P (*cf.*, eq. (6.99) below). The latter contains a term which punishes the cutting of data points fairly strongly. Hence, one needs to ensure that it does not get too aggressive in terms of excited-state contamination (at the lower end of the fit interval) and, on ensembles with open boundary conditions, also in terms of boundary effects (at the upper end of the fit interval). To that end, we further restrain the scan over fit ranges to such which fulfill $t_{\text{sep}}^{\text{min}} \geq 0.5$ fm and $t_{\text{sep}}^{\text{max}} \leq 2.0$ fm. Since the signal for the nucleon two-point function is largely lost for $t_{\text{sep}} \geq 2.0$ fm anyway, we impose the latter constraint on all ensembles, *i.e.*, also on such with periodic boundary conditions. In principle, one should average over all allowed fit ranges using weights as in eq. (6.101) and assign a corresponding systematic error. As the errors of the nucleon masses are completely subdominant in our analysis of the form factors, however, we refrain from this complication here and simply use the one fit range with the smallest AIC_P.

In fig. 6.6, the result of the single-exponential fit to the two-point function is compared with the effective mass calculated according to eq. (5.50) for the ensemble N451. The green band shows the fit result drawn in the preferred fit range $14 \leq t_{\text{sep}}/a \leq 24$. As one can see, this nicely agrees with the plateau observed in the effective mass.

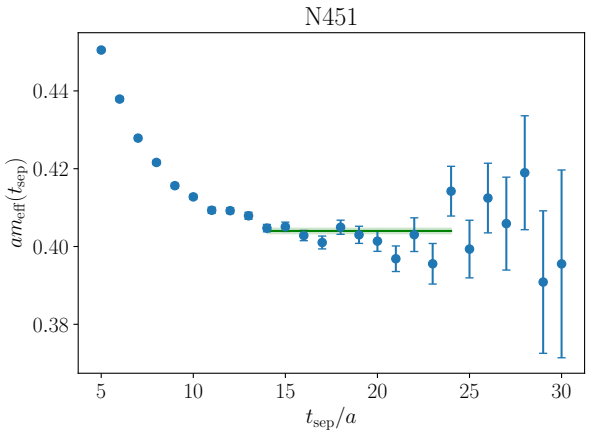


Figure 6.6: Effective nucleon mass (blue points) and result of the single-exponential fit to the nucleon two-point function (green band) on the ensemble N451

6.4.2 The pion mass

The extraction of the pion mass proceeds in a similar way to the nucleon. The pion two-point functions eq. (5.96) are computed on a subset of the sources employed for the nucleon (*cf.*, table 5.4). On ensembles with open boundary conditions, we also fit to a single-exponential *ansatz* $Ae^{-mt_{\mathrm{sep}}}$ as motivated by eq. (5.101), whereas on ensembles with periodic boundary conditions, we need to take the backward-propagating pion into account. Here, we symmetrize the pion two-point function around $T/2$ according to eq. (5.107) and then fit to a hyperbolic cosine $Ae^{-mT/2}\cosh[m(T/2 - t_{\mathrm{sep}})]$ as motivated by eq. (5.102). For the fit range scan, we demand a minimal length of $\max[0.75\ \mathrm{fm}, (T/2 - 2.25\ \mathrm{fm})/2]$ and 2.5 fm on open and periodic boxes, respectively. Moreover, we apply the restriction $t_{\mathrm{sep}}^{\mathrm{min}} \geq 0.75$ fm on all ensembles as a safeguard against excited-state effects. On ensembles with open boundary conditions, we further require that $t_{\mathrm{sep}}^{\mathrm{max}} \leq T/2 - 1.5$ fm. From the thus allowed fits, we again select the one with the smallest AIC_P.

In particular on ensembles with periodic boundary conditions, where we have relatively long plateaux and associated fit ranges, we observe that the results of correlated fits visually do not fit the data well. This is presumably due to the very strong point-to-point correlations which are present in the pion two-point function and the resulting difficulties in accurately estimating the covariance matrix. Nevertheless, for the purpose of determining the fit ranges, we need to perform correlated fits since the residuals, which enter the AIC, otherwise loose their meaning (*cf.*, the discussion close to the end of section 6.2.1). For the final fits from which our values of the pion mass are taken, however, we use uncorrelated fits.

In fig. 6.7, we compare the result of the uncorrelated cosh-fit to the pion two-point function on the ensemble N451 to the cosh-effective mass determined from the numerical solution of eq. (5.108).[4] The uncorrelated fit in the preferred range $18 \leq t_{\mathrm{sep}} \leq 51$ agrees very well with the plateau of the cosh-effective mass.[5]

The nucleon and pion masses resulting from the analysis described in this section are listed in table 6.1.

[4] We remark that the data shown in the right panel of fig. 5.8 and in fig. 6.7 are identical.

[5] We note that the corresponding correlated fit lies about $1\,\sigma$ higher; the issue is even worse on the ensembles D450 and E250 due to the smaller number of available gauge configurations.

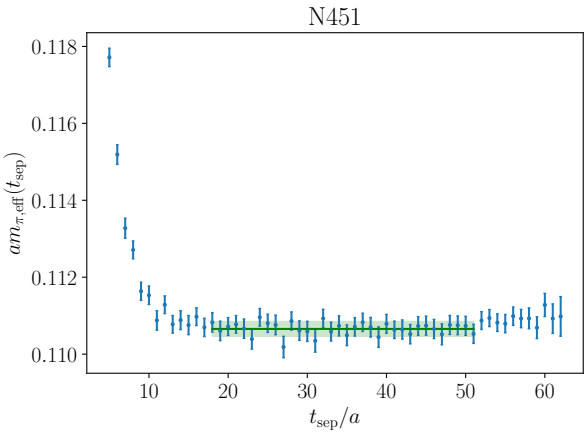


Figure 6.7: Effective pion mass (blue points) and result of the cosh-fit to the pion two-point function (green band) on the ensemble N451

Table 6.1: Pion and nucleon masses used in this thesis (in units of $\sqrt{t_0}$). The quoted errors include the error of $\sqrt{t_0^{\mathrm{sym}}}/a$ [115].

ID	$\sqrt{t_0}M_\pi$	$\sqrt{t_0}m_N$
C101	0.1662(10)	0.7160(35)
N101	0.20777(78)	0.7555(22)
H105	0.2073(19)	0.7704(35)
D450	0.16010(62)	0.7160(40)
N451	0.21167(60)	0.7728(22)
E250	0.09560(59)	0.6882(21)
D200	0.15162(74)	0.7261(23)
N200	0.20626(93)	0.7796(27)
S201	0.2163(13)	0.8323(49)
E300	0.12898(70)	0.7151(23)
J303	0.19476(64)	0.7667(22)

6.5 Ratios and effective form factors

After having discussed our determination of the relevant hadron masses, we can now turn to the extraction of the form factors. The relations connecting the nucleon three-point functions of the vector current with the electromagnetic form factors can be obtained by plugging the traces eqs. (5.58) to (5.60) into eq. (5.55). The resulting expressions still contain energy exponentials as well as the *a priori* unknown overlap factors c_N. The latter are defined in the same way as for the nucleon two-point function (*cf.*, eqs. (5.23) and (5.24)). Therefore, these factors can be cancelled by taking appropriate ratios of the three- and two-point functions, an idea initially proposed in Ref. [187]. Since eq. (5.55) is only valid in the asymptotic limit, the form factors extracted from the relations we will provide below are still contaminated by excited states heavier than the nucleon. We will hence call them effective form factors.

In our analysis, we average the nucleon two-point functions entering the ratios over equivalent momentum classes as in Ref. [188]. We call all three-momenta $\mathbf{p}$ which share the same modulus $|\mathbf{p}|$ equivalent and assign them the equivalence class $\mathfrak{p} = \{\tilde{\mathbf{p}} \in \tilde{\Lambda}^{(3)} : |\tilde{\mathbf{p}}| = |\mathbf{p}|\}$. Here, $\tilde{\Lambda}^{(3)}$ is

the set of possible lattice momenta, which is given by the spatial part of the dual lattice defined in eq. (3.26). The momentum-averaged two-point functions are then defined as

$$\langle \bar{C}_2 \rangle (\mathfrak{p}; t_{\text{sep}}) = \sum_{\tilde{\mathbf{p}} \in \mathfrak{p}} \langle C_2 \rangle (\tilde{\mathbf{p}}; t_{\text{sep}}) \Bigg/ \sum_{\tilde{\mathbf{p}} \in \mathfrak{p}} 1 \, . \tag{6.42}$$

Afterwards, we calculate the ratios [187]

$$R^{\Gamma^p}_{V_\mu}(\mathbf{0}, \mathbf{q}; t_{\text{sep}}, t) = \frac{\langle C^{\Gamma^p}_{3,V_\mu} \rangle (\mathbf{0}, \mathbf{q}; t_{\text{sep}}, t)}{\langle C_2 \rangle (\mathbf{0}; t_{\text{sep}})} \sqrt{\frac{\langle \bar{C}_2 \rangle (\mathfrak{q}; t_{\text{sep}} - t) \langle C_2 \rangle (\mathbf{0}; t) \langle C_2 \rangle (\mathbf{0}; t_{\text{sep}})}{\langle C_2 \rangle (\mathbf{0}; t_{\text{sep}} - t) \langle \bar{C}_2 \rangle (\mathfrak{q}; t) \langle \bar{C}_2 \rangle (\mathfrak{q}; t_{\text{sep}})}}. \tag{6.43}$$

The full expression relevant for $\mathbf{p}' \neq \mathbf{0}$ (not used in this thesis) can be found in Ref. [188]. From eqs. (5.29) and (5.55), it follows that the ratios are given by

$$R^{\Gamma^p}_{V_\mu}(\mathbf{0}, \mathbf{q}; t_{\text{sep}}, t) = \frac{1}{4 m_N \sqrt{2 E_{\mathfrak{q}} (m_N + E_{\mathfrak{q}})}} \frac{c_N(\mathbf{0})}{|c_N(\mathbf{0})|} \frac{c_N^*(-\mathbf{q})}{|\bar{c}_N(\mathfrak{q})|} \mathcal{T}^{\Gamma^p}_{V_\mu}(\mathbf{0}, \mathbf{q}), \tag{6.44}$$

where

$$|\bar{c}_N(\mathfrak{q})|^2 = \sum_{\tilde{\mathbf{q}} \in \mathfrak{q}} |c_N(\tilde{\mathbf{q}})|^2 \Bigg/ \sum_{\tilde{\mathbf{q}} \in \mathfrak{q}} 1 \, . \tag{6.45}$$

Moreover, the trace factor $\mathcal{T}^{\Gamma^p}_{V_\mu}(\mathbf{0}, \mathbf{q})$ (*cf.*, eq. (5.56)) is understood to be defined in terms of the effective form factors $G^{\text{eff}}_{E,M}(Q^2)$. The expression

$$e^{i\phi_{\mathbf{q}}} = \frac{c_N(\mathbf{0})}{|c_N(\mathbf{0})|} \frac{c_N^*(-\mathbf{q})}{|\bar{c}_N(\mathfrak{q})|} \tag{6.46}$$

is obviously a pure phase factor which we will always assume to be equal to unity in the following. In this case, by taking the ratio according to eq. (6.43), we have successfully eliminated all overlap factors and energy exponentials.

The ratio eq. (6.43) is constructed for the isovector, connected isoscalar and disconnected light − strange contributions to the three-point functions separately. For the connected part, we employ for each value of t_{sep} matching statistics in terms of sources for the two- and three-point functions entering eq. (6.43) (*cf.*, table 5.2). This preserves the full correlation between them, which slightly reduces the statistical fluctuations in the ratio. For the disconnected part, on the other hand, the highest statistics at our disposal is utilized for all values of t_{sep}, both for the two-point functions used to construct $\langle C_{3,V_\mu} \rangle^{\text{disc}}_F$ (*cf.*, the second term in eq. (5.70)) and for the ones entering eq. (6.43). In all cases, the same projection matrix is employed for both the two- and three-point functions entering eq. (6.43), again to ensure full correlation. For the connected part, this is only $\Gamma^p_3 = 1/2(1 + \gamma_0)(1 + i\gamma_5\gamma_3)$. For the disconnected part, we employ all three polarization directions, $\Gamma^p_j = 1/2(1 + \gamma_0)(1 + i\gamma_5\gamma_j)$, $j = 1, 2, 3$, and average the thus obtained effective form factors (see below). Moreover, we average over the forward- and backward-propagating nucleon for the disconnected part.

In our analysis, we hence extract all form factors using polarized projectors of the form $\Gamma^p_j = 1/2(1 + \gamma_0)(1 + i\gamma_5\gamma_j)$. During the writing of this thesis, however, it has been noticed that a more advantageous choice would be to extract the electric form factor from the 0-component of the vector current using the unpolarized projector $\Gamma^p_0 = 1/2(1 + \gamma_0)$ (*cf.*, eq. (5.58)) and the magnetic form factor from the orthogonal spatial components using $\tilde{\Gamma}^p_j = 1/2(1 + \gamma_0) i\gamma_5\gamma_j$. This is because in the 0-component of the vector current, the $\gamma_5\gamma_j$-term only contributes noise which has to average to zero (as mentioned in section 5.4.1). Therefore, directly omitting this part

of the projector would lead to a significantly reduced error for G_E. Similarly, the 1 in the second bracket of Γ_j^p yields the imaginary part proportional to G_E in eq. (5.60) and only noise in the real part, and could thus be omitted for the extraction of G_M. These improvements are only applicable to the disconnected part because in the connected part, only three-point functions with one projector from which all form factors can be extracted have been calculated in order to save computing time (*cf.*, section 5.4.3). Since the error of the full isoscalar form factors is already completely dominated by the connected part in our calculation, our somewhat disadvantageous choice of projectors should not strongly affect any of our final results.

By plugging eqs. (5.58) to (5.60) into eq. (6.44), we find that the ratios for the projector $\Gamma_j^p = {}^1\!/_2(1+\gamma_0)(1+i\gamma_5\gamma_j)$ can be decomposed as,

$$R_{V_0}^{\Gamma_j^p}(\mathbf{0},\mathbf{q};t_{\text{sep}},t) = \sqrt{\frac{m_N+E_\mathfrak{q}}{2E_\mathfrak{q}}}\,G_E^{\text{eff}}(Q^2;t_{\text{sep}},t), \tag{6.47}$$

$$R_{V_j}^{\Gamma_j^p}(\mathbf{0},\mathbf{q};t_{\text{sep}},t) = \frac{iq_j}{\sqrt{2E_\mathfrak{q}(m_N+E_\mathfrak{q})}}\,G_E^{\text{eff}}(Q^2;t_{\text{sep}},t), \tag{6.48}$$

$$R_{V_{k\neq j}}^{\Gamma_j^p}(\mathbf{0},\mathbf{q};t_{\text{sep}},t) = \frac{1}{\sqrt{2E_\mathfrak{q}(m_N+E_\mathfrak{q})}}\left[iq_k G_E^{\text{eff}}(Q^2;t_{\text{sep}},t) + \epsilon_{jkl}q_l G_M^{\text{eff}}(Q^2;t_{\text{sep}},t)\right]. \tag{6.49}$$

Here, G_E^{eff} and G_M^{eff} are the effective electric and magnetic form factors, respectively, which still depend on the source-sink separation t_{sep} and the operator insertion time t because they contain contributions from excited states. The spatial components of the vector current are not employed to compute the electric form factor as they are more noisy. Hence, we use the following estimators [60, 67],

$$G_E^{\text{eff}}(Q^2;t_{\text{sep}},t) = \sqrt{\frac{2E_\mathfrak{q}}{m_N+E_\mathfrak{q}}}\sum_{\tilde{\mathbf{q}}\in\mathfrak{q}} R_{V_0}^{\Gamma_j^p}(\mathbf{0},\tilde{\mathbf{q}};t_{\text{sep}},t)\Bigg/\sum_{\tilde{\mathbf{q}}\in\mathfrak{q}} 1\,, \tag{6.50}$$

$$G_M^{\text{eff}}(Q^2;t_{\text{sep}},t) = \sqrt{2E_\mathfrak{q}(m_N+E_\mathfrak{q})}\sum_{\substack{\tilde{\mathbf{q}}\in\mathfrak{q}\\ \sum_{k\neq j}\tilde{q}_k^2\neq 0}} \frac{\sum_{k,l}\epsilon_{jkl}\tilde{q}_l\,\mathrm{Re}\,R_{V_k}^{\Gamma_j^p}(\mathbf{0},\tilde{\mathbf{q}};t_{\text{sep}},t)}{\sum_{k\neq j}\tilde{q}_k^2}\Bigg/\sum_{\substack{\tilde{\mathbf{q}}\in\mathfrak{q}\\ \sum_{k\neq j}\tilde{q}_k^2\neq 0}} 1\,, \tag{6.51}$$

where we average over all three-momenta $\tilde{\mathbf{q}}$ belonging to the equivalence class $\mathfrak{q}$ and thus yielding the same Q^2 (except for those with $\sum_{k\neq j}\tilde{q}_k^2 = 0$ in case of the magnetic form factor). As mentioned above, we use only z-polarization ($j=3$) for the connected part, while for the disconnected part, we average over $j=1,2,3$.

Furthermore, we assume the relativistic dispersion relation $E_\mathfrak{q} = \sqrt{m_N^2+|\mathbf{q}|^2}$. We have checked explicitly that employing the extracted ground-state energies also for non-vanishing momenta instead of the above dispersion relation does not change our results for the ground-state form factors significantly. This is illustrated in figs. 6.8 and 6.9 for the ensemble D200: In fig. 6.8, we compare the extracted ground-state nucleon energy at different values of $|\mathbf{q}|^2$ (where we have averaged over all equivalent momenta) with the relativistic dispersion relation based on the nucleon mass. We note that we have used the same fit range for all momenta, namely the one determined at zero momentum as detailed in section 6.4.1. In this way, we achieve a consistent energy extraction across momenta, albeit at the price of potentially suboptimal fit ranges for individual momenta.

In fig. 6.9, we compare the ground-state form factors in the connected isoscalar channel corresponding to the two sets of energies, where the ground state has been identified by the methods described in sections 6.6.2 and 6.6.4 below. Even though the energies themselves can

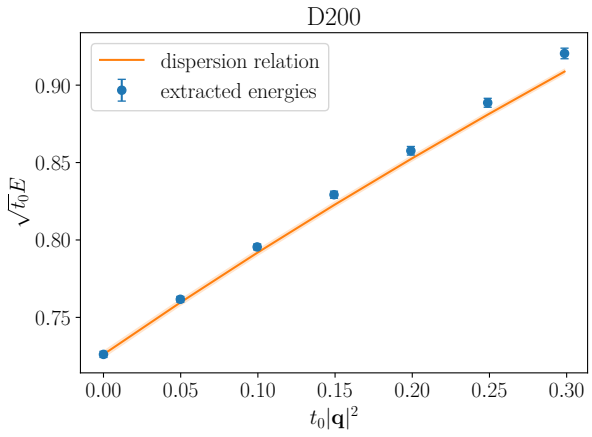


Figure 6.8: Extracted ground-state nucleon energy as a function of the three-momentum squared and relativistic dispersion relation on the ensemble D200

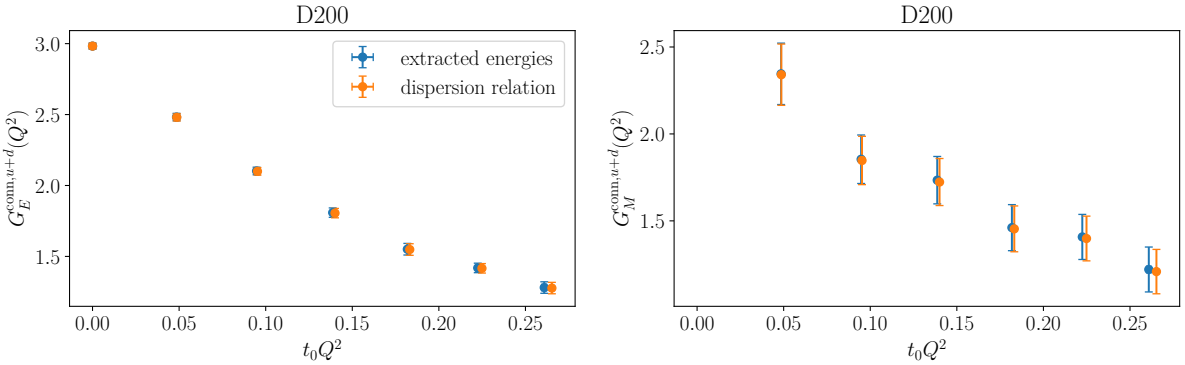


Figure 6.9: Connected contribution to the isoscalar electromagnetic form factors on the ensemble D200 as a function of Q^2. For the blue points, the extracted nucleon energies have been employed, and for the orange points the relativistic dispersion relation. The ground state has been identified using the summation method with the window average (*cf.*, sections 6.6.2 and 6.6.4 below).

differ quite significantly at higher momenta, this is of no consequence for the ground-state form factors which are very stable under such variations. Therefore, we conclude that using the dispersion relation in order to avoid potentially unstable fits to the two-point functions at high momenta does not introduce any bias in the calculation of the form factors. We also remark that if one wanted to employ the extracted nucleon energies for the final data set and not just at the level of a crosscheck, a more sophisticated way of setting fit ranges across momenta would be desirable. Moreover, to judge if the difference in the energies is actually as significant as fig. 6.8 suggests, one should average over several different fit ranges and assign a corresponding systematic error.

At the end of section 5.4.4, we have postponed a comparison of the form factors obtained with different criteria for the minimal distance to the boundaries on ensembles with open boundary conditions. Now that we have explained the extraction of the effective form factors, we want to come back to this point which only concerns the disconnected contribution in our setup. In fig. 6.10, we show in the upper panel the disconnected contribution to the effective form

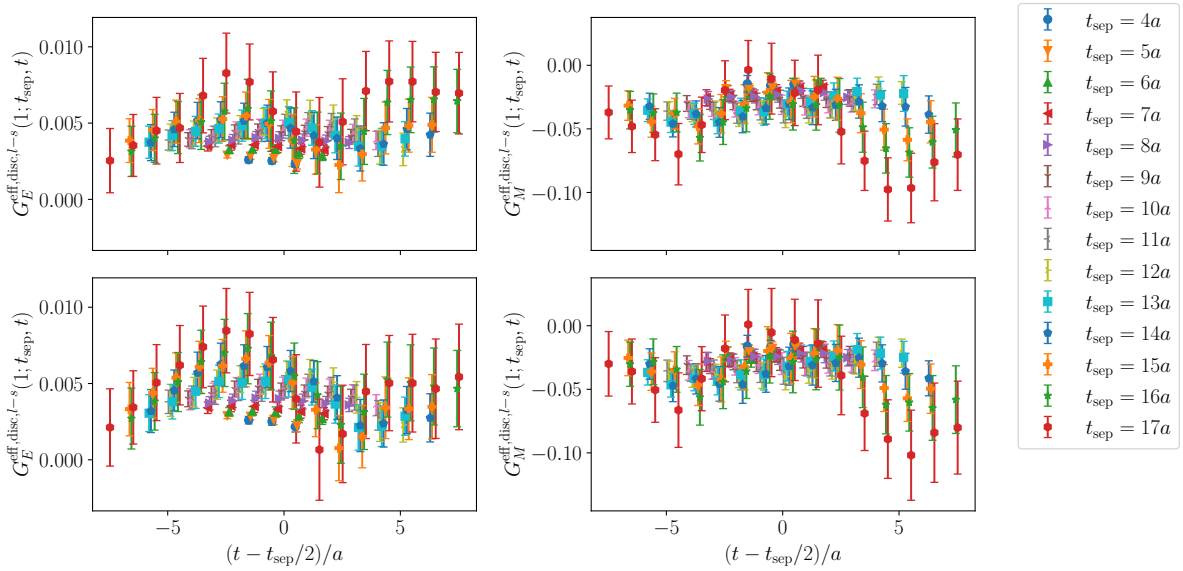


Figure 6.10: Disconnected contribution to the effective form factors at the first non-vanishing momentum on the ensemble C101 ($Q^2 \approx 0.089$ GeV^2) imposing a minimal distance of 2.5 fm (upper panel) or 3.5 fm (lower panel) to the boundary for the nucleon propagating towards it. In each row, the left plot shows the electric and the right one the magnetic form factor. For each source-sink separation t_{sep}, the effective form factors are displayed as a function of the operator insertion time t, offset to the midpoint between nucleon source and sink. The data points are horizontally displaced for better visibility.

factors at the first non-vanishing momentum on C101 obtained from our default criterion of 2.5 fm which the source of the nucleon propagating towards the boundary has to keep from the latter. In the lower panel, we show the same form factors obtained from a stricter criterion of 3.5 fm. Apart from some fluctuations at individual time slices, which are well covered by the errors, one cannot discern any systematic difference between the two data sets. It also clear to see that the errors are significantly larger for the data set imposing the stricter criterion because fewer sources from a more restricted region of the lattice are used in the measurement of the two-point functions. With the discussion at the end of section 5.4.4 in mind, and since the situation is very similar to fig. 6.10 on the other ensembles with open boundary conditions as well, we opt for the more relaxed criterion of 2.5 fm corresponding to the upper panel of fig. 6.10.

As mentioned above, the effective form factors eqs. (6.50) and (6.51) are constructed in the isospin basis, *i.e.*, for the isovector $(u-d)$ and the connected isoscalar $(u+d)$ combinations, as well as for the disconnected contributions of the light and strange quarks. The full isoscalar (octet) combination $u+d-2s$ (*cf.*, eq. (5.70)) is only built on this level, *i.e.*, we combine the effective form factors as

$$G_{E,M}^{\mathrm{eff},u+d-2s} = G_{E,M}^{\mathrm{eff,conn},u+d} + 2G_{E,M}^{\mathrm{eff,disc},l-s}. \tag{6.52}$$

Here, $l-s$ stands for the difference of the light- and strange-quark disconnected contributions. We drop the disconnected contribution $G_E^{\mathrm{eff,disc},l-s}$ at $Q^2 = 0$, as it has to be zero. Our data, on the other hand, show fluctuations around the exact zero due to the stochastic estimation of the quark loops and the application of the truncated-solver method for the calculation of the

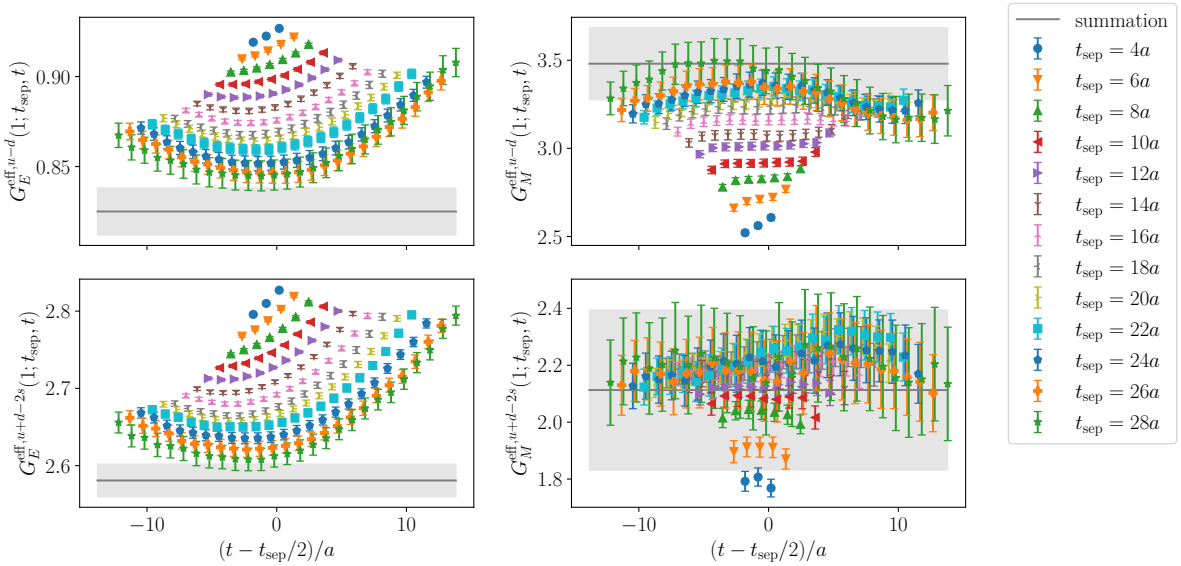


Figure 6.11: Effective form factors for the isovector (upper panel) and isoscalar (lower panel) combinations at the first non-vanishing momentum on the ensemble E300 ($Q^2 \approx 0.067$ GeV^2). In each row, the left plot shows the electric and the right one the magnetic form factor. For each source-sink separation t_{sep}, the effective form factors are displayed as a function of the operator insertion time t, offset to the midpoint between nucleon source and sink. The data points are horizontally displaced for better visibility. The gray bands and curves depict the results of the leading-order summation method using the window average, as detailed in sections 6.6.2 and 6.6.4.

two-point functions. Thus, explicitly adding this superfluous term, which is always compatible with zero, would artificially enhance noise in all data points for G_E, because they are normalized by $G_E(0)$ (*cf.*, section 6.7).

The resulting effective form factors are shown in fig. 6.11 for the first non-vanishing momentum on the ensemble E300. This demonstrates that we obtain a clear signal including the disconnected contributions: the families of points for the different source-sink separations can be clearly distinguished in all cases except for the larger values of t_{sep} in $G_M^{\mathrm{eff},u+d-2s}$. Figure 6.11 is also a good illustration of the baryonic signal-to-noise problem: even though we have significantly increased our statistics in terms of sources for the connected part (which dominates the errors) for the larger source-sink separations compared to the smaller ones (*cf.*, table 5.2), the errors are still growing quite drastically with t_{sep}. The excited-state contamination, on the other, is clearly reduced by going to larger t_{sep}, t and $t_{\mathrm{sep}} - t$, *i.e.*, towards the midpoint at large t_{sep}.

6.6 Excited-state analysis

Due to the baryonic signal-to-noise problem, it is in current lattice studies of baryon structure observables typically not possible to access source-sink separations at which it can be guaranteed that contributions from excited states are sufficiently suppressed. This necessitates an explicit treatment of the excited-state systematics in order to extract the ground-state form factors from the effective ones. Several methods for this have been developed and are in use; for a relatively recent review, see Ref. [189]. The arguably simplest approach is to fit a constant to

the effective form factor around the midpoint at the largest few source-sink separations, which is commonly referred to as plateau fits. Our data in particular for the electric form factor (which is more precise than the magnetic one) does, however, not show a clear plateau even at the largest source-sink separations (*cf.*, fig. 6.11). This indicates that some excited-state contamination is still present there and more sophisticated techniques are required.

In this thesis, we study two complementary approaches: One is based on fits to the effective form factors themselves, similar to the plateau fits, but taking an additional excited state explicitly into account. This will be discussed in section 6.6.1. The other method constructs an improved observable from the effective form factors in which the excited-state contamination is parametrically more strongly suppressed. This makes it possible to use a fit *ansatz* without explicit higher-order terms. The method in general and its application to leading order (LO) is the topic of section 6.6.2, while the prospect of including next-to-leading order (NLO) terms will be explored in section 6.6.3. In section 6.6.4, we will come back to the LO method and introduce a way to average over different fit ranges.

6.6.1 Two-state fits to the effective form factors

In order to perform two-state fits to the effective form factors, we need to derive a generalization of eq. (6.44) which takes a second state into account. By truncating the sums in eqs. (5.29) and (5.54) only after the terms corresponding to the states with the second-lowest energy, one can show that the effective form factors take the following form [60],

$$\begin{aligned} G_{E,M}^{\text{eff}}(Q^2; t_{\text{sep}}, t) = r_{00}(Q^2) \Bigg\{ 1 + \frac{\rho(Q^2)}{2} \left[e^{-\Delta(Q^2)(t_{\text{sep}}-t)} - e^{-\Delta(Q^2) t_{\text{sep}}} \right] \\ + \frac{\rho(0)}{2} \left[e^{-\Delta(0) t} - e^{-\Delta(0) t_{\text{sep}}} \right] \Bigg\} \\ + r_{01}(Q^2) e^{-\Delta(Q^2) t} + r_{10}(Q^2) e^{-\Delta(0)(t_{\text{sep}}-t)} + r_{11}(Q^2) e^{-\Delta(Q^2) t} e^{-\Delta(0)(t_{\text{sep}}-t)}. \end{aligned} \tag{6.53}$$

Here, $r_{00}(Q^2) = G_{E,M}(Q^2)$ are the ground-state form factors, $\Delta(Q^2) = E_1(Q^2) - E_0(Q^2)$ is the energy gap between the excited state and the nucleon ground state, and $\rho(Q^2) = c_1(Q^2)/c_0(Q^2)$ is the corresponding ratio of overlap factors which we assume to be independent of the momentum direction. The r_{01} and r_{10} terms stem from the summands in eq. (5.54) where one state is the ground state and the other one excited, while the r_{11} term describes an excited-to-excited transition. We fit the electric and magnetic effective form factors simultaneously to eq. (6.53), with the energy gaps $\Delta(Q^2)$, $\Delta(0)$ and the overlap factors $\rho(Q^2)$, $\rho(0)$ as common fit parameters.

To achieve stable fits, priors on the energy gaps and also on the overlap factors are required. In order to determine these, we perform two-state fits to the (momentum-averaged) two-point functions,

$$\langle \bar{C}_2 \rangle (\mathfrak{p}; t_{\text{sep}}) = c_0(\mathfrak{p}) e^{-E_0(\mathfrak{p}) t_{\text{sep}}} + c_1(\mathfrak{p}) e^{-E_1(\mathfrak{p}) t_{\text{sep}}}, \tag{6.54}$$

where we have absorbed the kinematic prefactors appearing in eq. (5.29) into the overlap factors $c_{0,1}$ and again assumed the latter to be independent of the momentum direction. From the fits to eq. (6.54), we extract the energy gaps $\Delta(Q^2) = E_1(Q^2) - E_0(Q^2)$ and the overlap factors $\rho(Q^2) = c_1(Q^2)/c_0(Q^2)$.

We fix the upper bound of the fit interval to the one used for the one-state fit to the zero-momentum two-point function (*cf.*, section 6.4.1) because it exerts only a marginal influence on the results. The systematic uncertainty originating from the choice of the lower bound of the fit interval, on the other hand, is accounted for by averaging over all reasonable options using

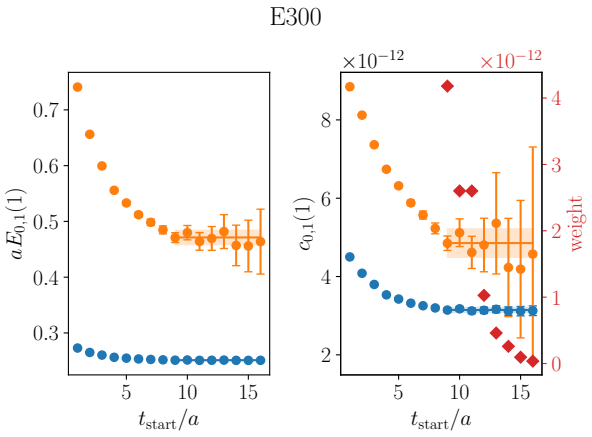


Figure 6.12: Results of the two-state fits to the nucleon two-point function at the first non-vanishing momentum on the ensemble E300 ($Q^2 \approx 0.067$ GeV^2) as a function of the lower bound of the fit interval t_{start}. The bands depict the averaged results in the interval of t_{start} values we average over. The red diamonds show the weights derived from eq. (6.101).

AIC$_P$ weights (*cf.*, eqs. (6.99) and (6.101) below) and adding Gaussian noise to the Jackknife distribution according to the systematic covariance matrix [190],

$$C_{\mathrm{syst}}(Q^2) = \sum_{\mu=1}^{N_M} w_\mu^{\mathrm{AIC}_P} \begin{pmatrix} \Delta_\mu(Q^2) \\ \rho_\mu(Q^2) \end{pmatrix} \begin{pmatrix} \Delta_\mu(Q^2) & \rho_\mu(Q^2) \end{pmatrix} - \left(\sum_{\mu=1}^{N_M} w_\mu^{\mathrm{AIC}_P} \begin{pmatrix} \Delta_\mu(Q^2) \\ \rho_\mu(Q^2) \end{pmatrix} \right) \left(\sum_{\mu=1}^{N_M} w_\mu^{\mathrm{AIC}_P} \begin{pmatrix} \Delta_\mu(Q^2) & \rho_\mu(Q^2) \end{pmatrix} \right) . \tag{6.55}$$

Here, the sums run over the N_M models (*i.e.*, values for the lower bound of the fit interval t_{start}) and $w_\mu^{\mathrm{AIC}_P}$ is as in eq. (6.101), but using the AIC$_P$ instead of the BAIC$_P$. The lowest t_{start} value entering the average is taken as the first one for which the results for $E_{0,1}$ and $c_{0,1}$ from this fit are compatible within errors with the corresponding averaged ones. The largest t_{start} value we consider is the last one for which all four fit parameters have a relative error ≤ 100 %.

The averaging over different fit intervals is illustrated in fig. 6.12 for the first non-vanishing momentum on the ensemble E300. In the left panel, we show the results for the two fitted energy levels, and in the right panel the corresponding overlap factors, both as a function of t_{start}. The bands depict the averaged results in the interval of t_{start} values we average over. One can see that our method selects the first t_{start} value entering the average such that it marks the onset of an approximate plateau, and such that the weights (shown in red) monotonically decrease in our averaging interval. This means that we only take models into account for which the weights are suppressed due to a large cut in the number of data points, and not such where the weights are suppressed due to a bad fit quality (*cf.*, the discussion around eqs. (6.99) and (6.100) below).

As the ρ-factors, which are defined by the overlaps in the two-point function, only enter the spectral representation of the ratio used to derive eq. (6.53) via the expansion of the two-point function, we directly take the values obtained from the fits to eq. (6.54) as priors for them. Here, we increase the width of the priors by multiplying the error of ρ by a conservative factor of 3.

The energy gaps, on the other hand, also enter the terms in eq. (6.53) which originate from the expansion of the three-point function. Hence, the situation is less clear than for the ρ-factors because the three-point function might have a stronger overlap with different excited states than the two-point function. This is intimately connected with the issue that our calculation, as almost certainly all other current lattice calculations of nucleon matrix elements, is not in the regime where a single excited state dominates the excited-state contamination. Since it is impossible to fit more than one excited state without putting unduly strict priors on the energy gaps (which would correspond to making an *a priori* assumption about what the states are which couple most strongly, and not really letting the data decide on this), the latter must be regarded as effective gaps summarizing the contribution of several excited states. Therefore, we employ relatively loose priors set to the range between $2M_\pi$ and the energy gap obtained from fitting the two-point function. This comes of course at the expense of less stable fits to the effective form factors. But we stress again that it is necessary in order to not introduce a systematic bias by assuming a particular value of the gap.

Both the two-state fits to the two-point functions and those to the effective form factors are performed using the VarPro method (*cf.*, section 6.2.2) to eliminate the need for initial guesses for the prefactors c_j and r_{jk}, respectively. To determine the range of points which should enter the two-state fits to the effective form factors, we compare different choices of $t_{\text{sep}}^{\text{min}}$ (the minimal source-sink separation entering the fit) and t_{skip} (the number of timeslices skipped from the borders of the insertion time) in fig. 6.13. Here, we show the three ensembles E300, D200 and C101 which we have studied in detail using the two-state fits. We find that fits including a very large amount of data are unstable and frequently converge to an obviously wrong local minimum. Consequently, both of the above parameters need to be set to relatively large values in order to obtain stable fit results and p-values which are acceptable at least in the majority of cases. Our final choices are $t_{\text{sep}}^{\text{min}} \gtrsim 6.9\sqrt{t_0}$, which corresponds to the peak of the window which we will use in the summation method to average over different $t_{\text{sep}}^{\text{min}}$ values there (*cf.*, section 6.6.4 below), and $t_{\text{skip}} \gtrsim 2.6\sqrt{t_0} \approx 0.4$ fm. The latter is realized by $t_{\text{skip}} = 8a$ on E300, $6a$ on D200 and $5a$ on C101. We remark that these values correspond to omitting about half of the data points even at our largest source-sink separation.

In figs. 6.14 to 6.16, we show our chosen two-state fits to the effective form factors on the aforementioned ensembles and $Q^2 \approx 0.2$ GeV^2. One can see that the two-state fits in many cases miss the data (even if the p-value is decent) and/or lead to an unrealistically large correction compared to the largest source-sink separation we have computed.

These observations suggest that two-state fits are likely not the optimal way to treat the excited-state systematics for our data. In particular the need to use priors in this very sensitive and crucial step of the analysis is a major concern. We have found that the choice of the location as well as of the width of the priors on the energy gaps strongly influences both the central values of the resulting ground-state form factors and their errors. Furthermore, with the relatively broad priors which we have finally adopted, we observe some instabilities in the two-state fits, mostly on ensembles with less statistics than the ones shown here and at higher momenta than those included in the analysis. For these reasons, we will focus on the alternative approach provided by the summation method in the following. We will, however, present a comprehensive comparison between the two strategies in section 6.6.5.

6.6.2 The summation method

The summation method [53, 191, 192] is based on summing the effective form factors obtained from eqs. (6.50) and (6.51) over the operator insertion time, omitting t_{skip} timeslices at both

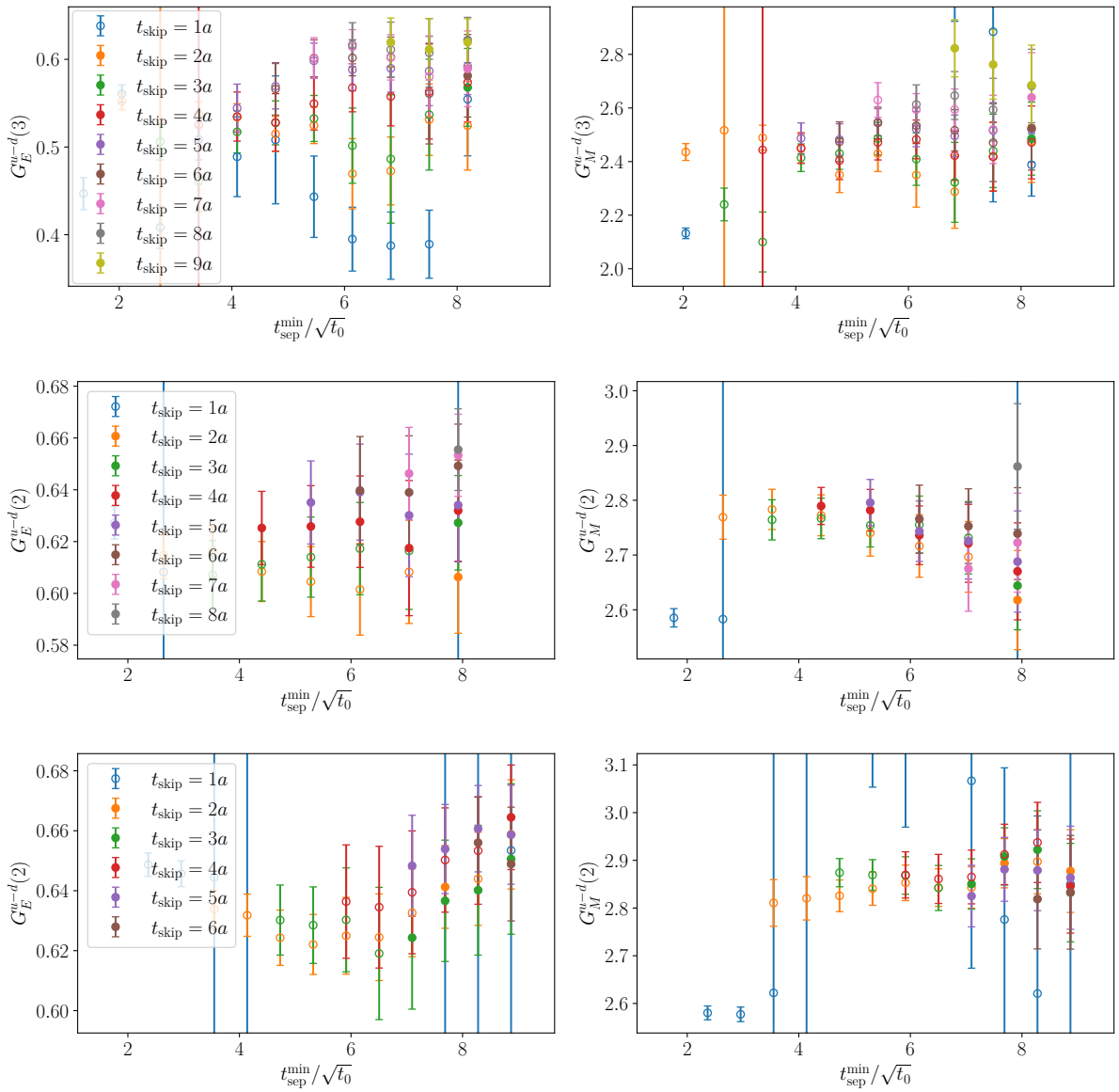


Figure 6.13: Isovector electromagnetic form factors at $Q^2 \approx 0.2$ GeV^2 on the ensembles E300 (upper panel), D200 (middle panel) and C101 (lower panel) as a function of the minimal source-sink separation entering the fits to eq. (6.53) and for different numbers of timeslices skipped from the borders. Open circles refer to fits with a p-value less than 5 %. Seemingly missing points lie, due to convergence issues, outside of the plotted range.

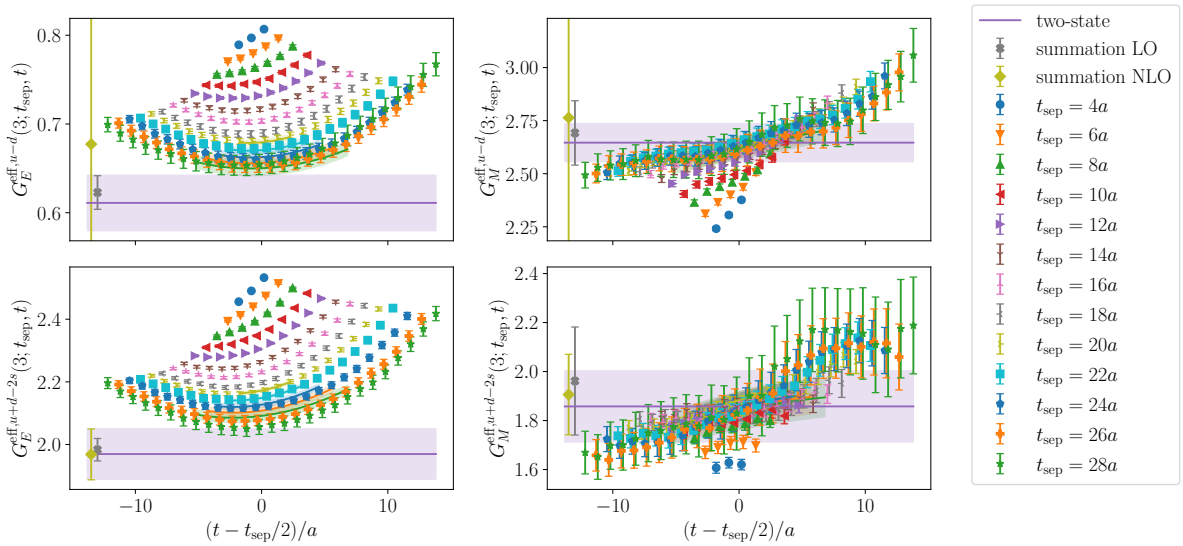


Figure 6.14: Effective form factors for the isovector (upper panel) and isoscalar (lower panel) combinations for ensemble E300 and $Q^2 \approx 0.196$ GeV^2. The data points are horizontally displaced for better visibility. The curves represent the two-state fits in their respective fit intervals, and the horizontal bands their extrapolation to $t_{\mathrm{sep}}, t \to \infty$.

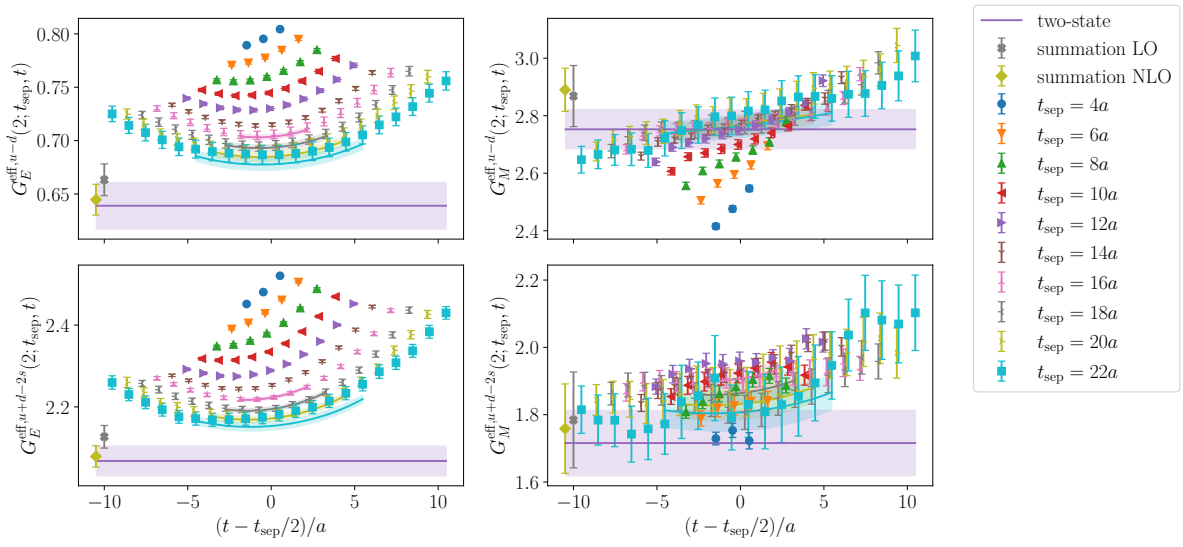


Figure 6.15: Same as fig. 6.14 for ensemble D200 and $Q^2 \approx 0.177$ GeV^2

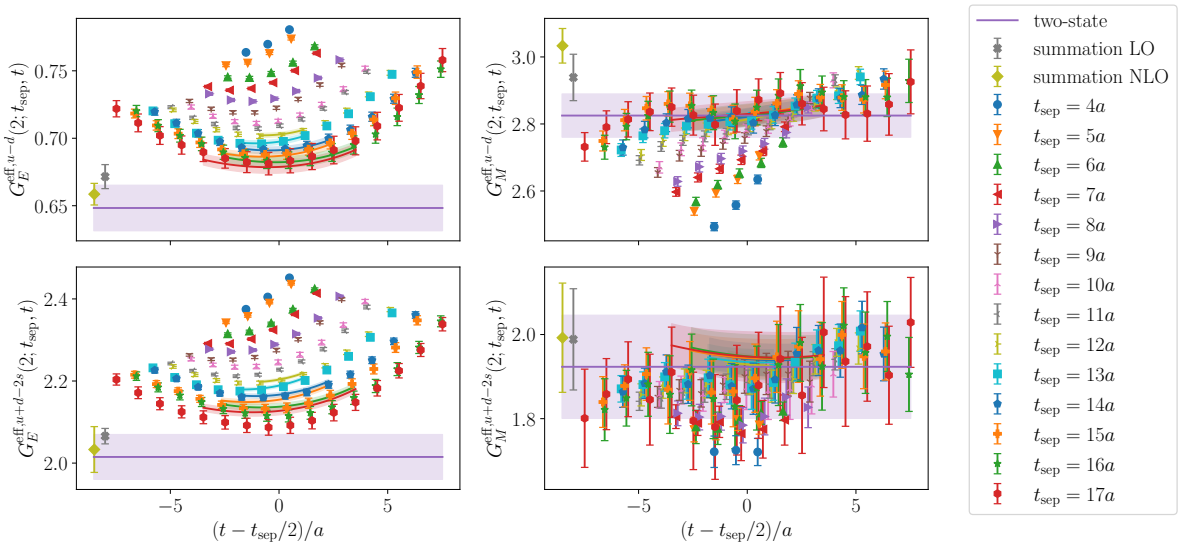


Figure 6.16: Same as fig. 6.14 for ensemble C101 and $Q^2 \approx 0.174$ GeV^2

ends,

$$S_{E,M}(Q^2;t_{\mathrm{sep}}) = \sum_{t=t_{\mathrm{skip}}}^{t_{\mathrm{sep}}-t_{\mathrm{skip}}} G^{\mathrm{eff}}_{E,M}(Q^2;t_{\mathrm{sep}},t). \tag{6.56}$$

By applying this summation to eq. (6.53), one can show that the two-state truncated spectral representation of $S_{E,M}$ is given by [60],

$$\begin{aligned} S_{E,M}(Q^2;t_{\mathrm{sep}}) = r_{00}(Q^2)\left[1 - \frac{\rho(Q^2)}{2}e^{-\Delta(Q^2)t_{\mathrm{sep}}} - \frac{\rho(0)}{2}e^{-\Delta(0)t_{\mathrm{sep}}}\right]\frac{1}{a}(t_{\mathrm{sep}} + a - 2t_{\mathrm{skip}}) \\ + \left[r_{01}(Q^2) + r_{00}(Q^2)\frac{\rho(Q^2)}{2}\right]\frac{e^{-\Delta(Q^2)(t_{\mathrm{skip}}-a)} - e^{-\Delta(Q^2)(t_{\mathrm{sep}}-t_{\mathrm{skip}})}}{e^{a\Delta(Q^2)} - 1} \\ + \left[r_{10}(Q^2) + r_{00}(Q^2)\frac{\rho(0)}{2}\right]\frac{e^{-\Delta(0)(t_{\mathrm{skip}}-a)} - e^{-\Delta(0)(t_{\mathrm{sep}}-t_{\mathrm{skip}})}}{e^{a\Delta(0)} - 1} \\ + r_{11}(Q^2)\frac{e^{-\Delta(Q^2)(t_{\mathrm{skip}}-a)-\Delta(0)(t_{\mathrm{sep}}-t_{\mathrm{skip}})} - e^{-\Delta(Q^2)(t_{\mathrm{sep}}-t_{\mathrm{skip}})-\Delta(0)(t_{\mathrm{skip}}-a)}}{e^{a\Delta(Q^2)} - e^{a\Delta(0)}}. \end{aligned} \tag{6.57}$$

Comparing this to eq. (6.53), one notices that in the summed quantity, the contributions from the excited states only enter with $e^{-\Delta t_{\mathrm{sep}}}$ rather than with $e^{-\Delta t}$, $e^{-\Delta(t_{\mathrm{sep}}-t)}$. Since both of the latter two factors must be small in order to minimize the excited-state contamination in the effective form factors themselves, one can at most achieve a suppression by $e^{-\Delta t_{\mathrm{sep}}/2}$ in that case. In the summed quantity, one gains a factor of two in the exponent relative to that.

This enhanced suppression of excited states allows one to neglect the explicit NLO terms in eq. (6.57) already at source-sink separations which are available at present. In the LO truncation with only ground-state contributions, the slope of $S_{E,M}$ as a function of t_{sep} is given by the ground-state form factor [53, 60],

$$S_{E,M}(Q^2;t_{\mathrm{sep}}) \xrightarrow{t_{\mathrm{sep}}\gg 0} C_{E,M}(Q^2) + \frac{1}{a}(t_{\mathrm{sep}} + a - 2t_{\mathrm{skip}})G_{E,M}(Q^2). \tag{6.58}$$

The constant $C_{E,M}$ can be used to capture residual excited-state effects but has no physical significance (it would vanish if there were no contributions from excited states at all). The

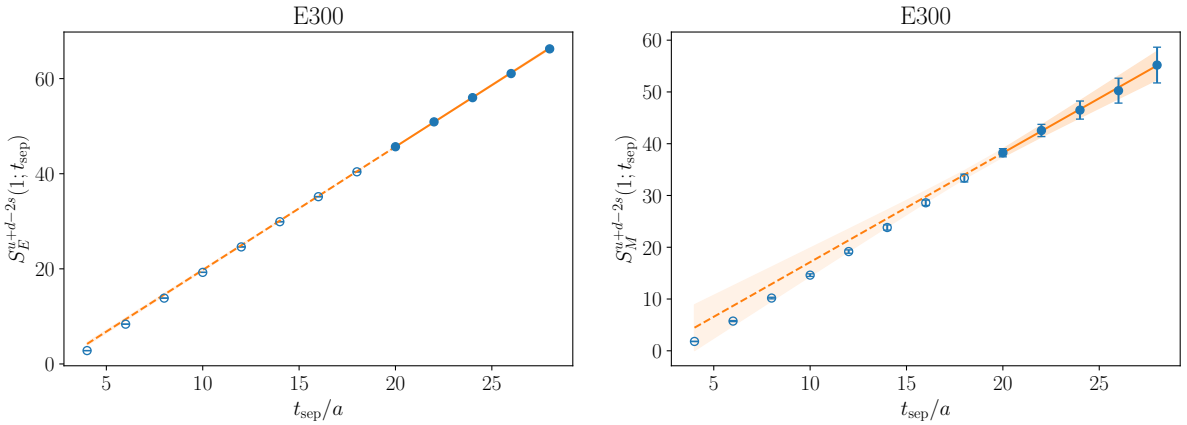


Figure 6.17: Summed isoscalar effective form factors at the first non-vanishing momentum on the ensemble E300 ($Q^2 \approx 0.067$ GeV^2) as a function of the source-sink separation. The filled blue circles show the fitted data and the solid orange line and band the corresponding fit to eq. (6.58) with its associated error. The open blue circles are not fitted, and the dashed orange line depicts an extrapolation of the fit.

linear fit to $S_{E,M}^{u+d-2s}$ according to eq. (6.58) is illustrated in fig. 6.17 for the first non-vanishing momentum on the ensemble E300 and a typical value of the minimal source-sink separation $t_{\mathrm{sep}}^{\mathrm{min}}$. While there is no significant deviation from the linear behavior in the fitted range $t_{\mathrm{sep}} \geq t_{\mathrm{sep}}^{\mathrm{min}} = 20a \approx 1$ fm, an increasing curvature becomes visible for smaller source-sink separations $t_{\mathrm{sep}} \lesssim 16a \approx 0.8$ fm.

This implies that a careful tuning of the minimal source-sink separation $t_{\mathrm{sep}}^{\mathrm{min}}$ entering the fits is required in order to reliably extract the ground-state form factors from the LO summation method. While we will postpone the greater part of this discussion to section 6.6.4, we want to briefly mention the other tunable parameter of the summation method now: the number of timeslices t_{skip} omitted at both ends of the operator insertion time (*cf.*, eq. (6.56)). Contrary to the naive expectation, a larger value of t_{skip} actually leads to a larger excited-state contamination in the summed ratio at fixed t_{sep}. This is due to the factor $\propto \exp[-\Delta(t_{\mathrm{sep}} - t_{\mathrm{skip}})]$ in the two-state truncated version of the summation method eq. (6.57). If the limit $t_{\mathrm{sep}} \to \infty$ is taken at fixed t_{skip}, however, all extractions should, independently of t_{skip}, converge to the same ground-state form factor.

Both aforementioned trends are actually observed in stability plots for the results of the summation method as a function of $t_{\mathrm{sep}}^{\mathrm{min}}$ and t_{skip}. These can be found in fig. 6.18 for the isoscalar channel at the first non-vanishing momentum on the ensembles D450 and E300. In the region of $t_{\mathrm{sep}}^{\mathrm{min}}$ values we will consider as most important for the further analysis (indicated in gray in fig. 6.18), all (not unreasonably large) values of t_{skip} already agree well. Therefore, using only one value of t_{skip} is perfectly adequate to obtain a reliable estimate of the ground-state form factors and their uncertainties. It can also be seen from fig. 6.18 that the errors tend to get smaller with rising t_{skip} because less of the somewhat more noisy data at the borders of the insertion time are employed to build the summed ratios. Consequently, $t_{\mathrm{skip}} = 2a$ appears to be a good compromise between not increasing the excited-state contamination in the summed ratio due to the effect mentioned before and excluding some of the potentially slightly less reliable data close to the source or sink. This is the value we employ for our main analysis.

Figure 6.18 also shows that the choice of $t_{\mathrm{sep}}^{\mathrm{min}}$ exerts in general a stronger influence on the extracted ground-state form factors than that of t_{skip}. The former is hence the more important

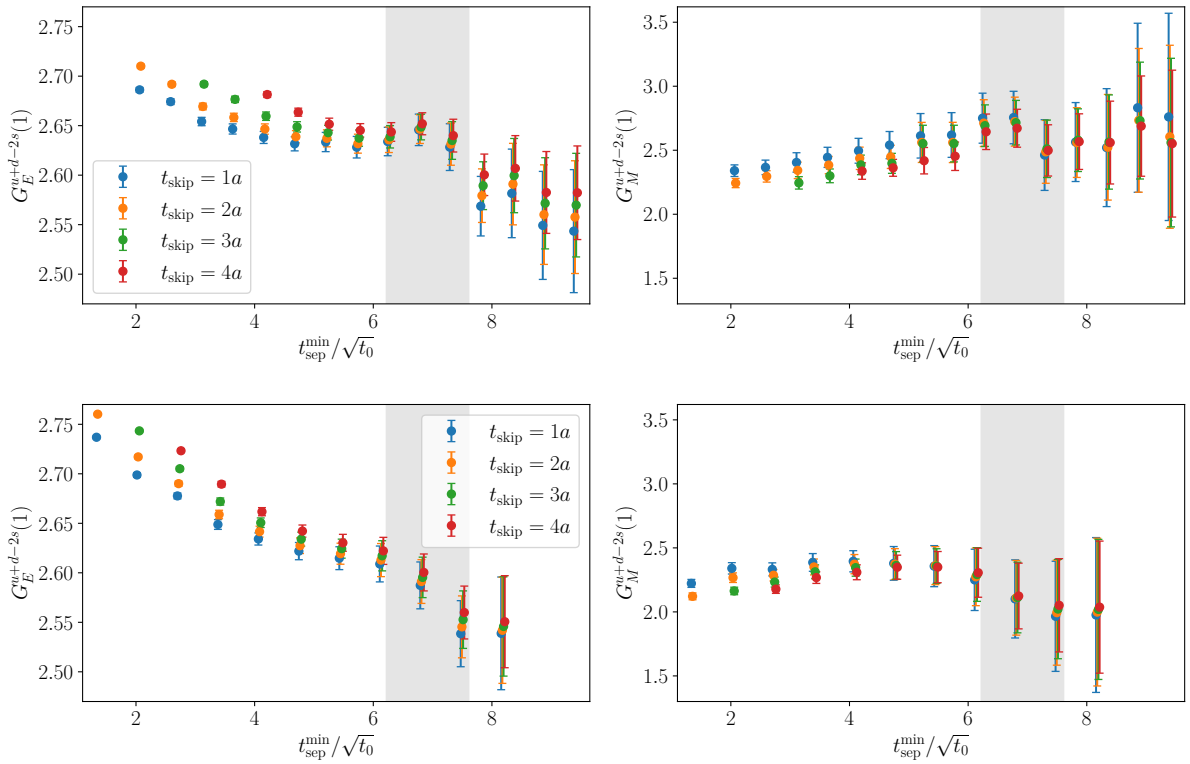


Figure 6.18: Isoscalar electromagnetic form factors at the first non-vanishing momentum on the ensembles D450 (upper panel) and E300 (lower panel) as a function of the minimal source-sink separation entering the fits to eq. (6.58) and for different values of t_{skip}. Each point represents a single fit starting at the source-sink separation given on the horizontal axis and using a value of t_{skip} as indicated by its color. The points are horizontally displaced for better visibility. The vertical gray bands mark the interval $t_{\mathrm{sep}}^{\mathrm{min}} \in [t_w^{\mathrm{low}}, t_w^{\mathrm{up}}]$ in which the window we will introduce in section 6.6.4 below has its largest support.

decision to make, and we will introduce a method to average over different values in section 6.6.4 below.

6.6.3 Two-state truncated summation method

As an alternative to the LO truncation eq. (6.58) of the summation method, one can also try to fit the full NLO expression given by eq. (6.57). This strategy has been pioneered in Refs. [193–195] for the analysis of nucleon matrix elements at zero momentum transfer. In this case, the expressions simplify and depend only on one energy gap Δ. For the momentum-dependent observables considered in the present thesis, however, no such simplification occurs; additionally, our data are less precise and more noisy.

Compared to the two-state fits to the effective form factors (*cf.*, section 6.6.1), one expects that the summed data are less sensitive to the energy gaps $\Delta(Q^2)$, $\Delta(0)$ because of the parametrically stronger suppression of the excited states. Accordingly, we observe that even with priors on the energy gaps and overlap factors similar to the ones introduced in section 6.6.1, fits to eq. (6.57) are not stable on the majority of ensembles. A promising strategy is therefore to start omitting terms from eq. (6.57) which are poorly constrained by the data. This concerns in particular the

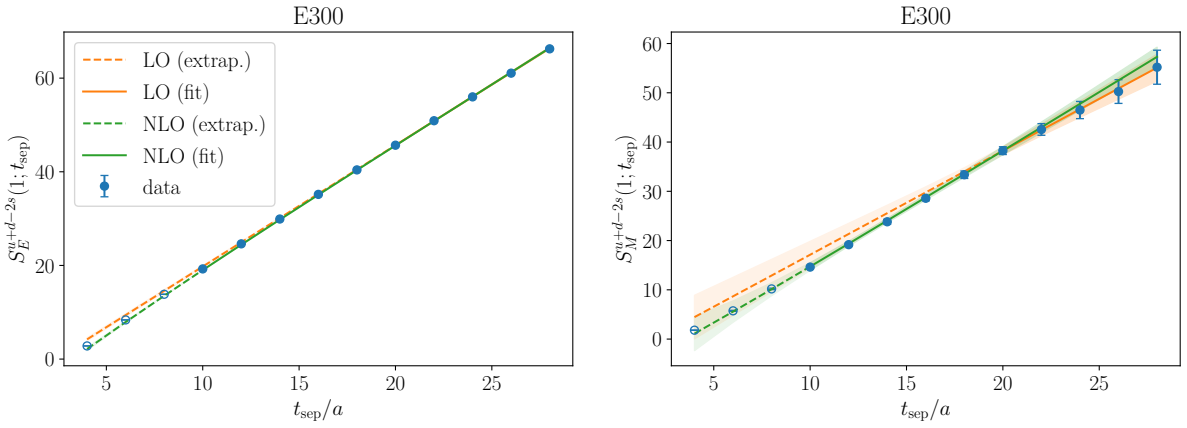


Figure 6.19: Same as fig. 6.17 extended by the fit to eq. (6.59) (solid green line) and its extrapolation (dashed green line)

term describing excited-to-excited transitions proportional to r_{11} and the terms proportional to the overlap factors $\rho(Q^2)$, $\rho(0)$. However, even with these terms omitted from the fit model and still using priors on the energy gaps, the fit can in the majority of cases not clearly distinguish contributions from two different exponentials in the data for the summed effective form factors.

Consequently, we define a very simple model with just one energy gap,

$$S_{E,M}(Q^2;t_{\mathrm{sep}}) = c_0(Q^2) + c_1(Q^2)\frac{1}{a}(t_{\mathrm{sep}} + a - 2t_{\mathrm{skip}}) + c_2(Q^2)e^{-\tilde{\Delta}(Q^2)(t_{\mathrm{sep}}-t_{\mathrm{skip}}+a)}, \qquad (6.59)$$

where we fit c_0, c_1, c_2 and $\tilde{\Delta}$. $c_1(Q^2) = G_{E,M}(Q^2)$ are the ground-state form factors, and the constant $c_0(Q^2)$ is used to capture residual effects of excited states other than the ones described by the effective gap $\tilde{\Delta}$. The fit to eq. (6.59) is illustrated in fig. 6.19 for the first non-vanishing momentum on the ensemble E300. Here, we have employed the same prior on $\tilde{\Delta}(Q^2)$ as for the two-state fit to the effective form factors (*i.e.*, the range between $2M_\pi$ and the value of $\Delta(Q^2)$ obtained from fitting the two-point function) and $t_{\mathrm{sep}}^{\mathrm{min}} = 10a \approx 0.5$ fm. We remark that the latter is already more conservative than any of the values considered in Ref. [195]. For small source-sink separations, there is a clear deviation from the linear behavior of the LO curve, which is well described by the (simplified) NLO curve in its fitted range. For large source-sink separations, on the other hand, the NLO curve converges to a straight line with a slope compatible with the one obtained from the LO fit.

We will come back to a more comprehensive comparison between the different approaches (two-state fits to the effective form factors, LO summation method and NLO summation method) in section 6.6.5. For the moment, it is worth noting that even the fit to the simplified model of eq. (6.59) is not sufficiently stable on all ensembles. A prominent example are the isovector form factors on E300 (*cf.*, fig. 6.14), where huge errors and, in some cases, also unexpected central values result. This is due to the fit converging, in spite of the prior, to extremely small or extremely large values of the energy gap $\tilde{\Delta}$ on individual Jackknife samples (or even on the central value). Because of these instabilities as well as the need for priors even in a highly simplified model such as eq. (6.59), we prefer the LO method of eq. (6.58) on which we will focus in the following.

6.6.4 The window average

The LO summation method does not make any attempt to explicitly correct for the excited-state contamination. Therefore, it is paramount to choose a minimal value of the source-sink separation at which the excited-state effects are already suppressed to a level significantly below the statistical uncertainties. To determine appropriate fit ranges, we perform several fits to eq. (6.58) for different starting values $t_{\mathrm{sep}}^{\mathrm{min}}$ of the source-sink separation.

Rather than selecting one particular value of $t_{\mathrm{sep}}^{\mathrm{min}}$ on each ensemble, we perform a weighted average over $t_{\mathrm{sep}}^{\mathrm{min}}$, where the weights are given by a smooth window function [183, 196],

$$\hat{G} = \frac{\sum_i w_i G_i}{\sum_i w_i}, \qquad w_i = \tanh \frac{t_i - t_w^{\mathrm{low}}}{\Delta t_w} - \tanh \frac{t_i - t_w^{\mathrm{up}}}{\Delta t_w}. \tag{6.60}$$

Here, t_i is the value of $t_{\mathrm{sep}}^{\mathrm{min}}$ in the i-th fit, and we choose $t_w^{\mathrm{low}} = 6.22\sqrt{t_0} \approx 0.9$ fm, $t_w^{\mathrm{up}} = 7.61\sqrt{t_0} \approx 1.1$ fm and $\Delta t_w = 0.553\sqrt{t_0} \approx 0.08$ fm. Note that the window has been shifted to the right by 0.1 fm as compared to the one originally used in Refs. [183, 196]. The reason for this is that our data for the electromagnetic form factors are statistically more precise than those for the axial form factor in Ref. [196] or the sigma term in Ref. [183]. Hence, we can resolve excited-state effects for larger values of $t_{\mathrm{sep}}^{\mathrm{min}}$, so that the plateau region is expected to start later. Accordingly, we observe that the window using larger $t_w^{\mathrm{up,low}}$ better captures the plateau on the majority of our ensembles. Thus, we opt for the choice which is more conservative regarding the suppression of excited states, and which also yields a larger error. We furthermore remark that the two options (window centered at 0.9 fm and at 1 fm, respectively) lead in most cases to well compatible results (*cf.*, fig. 9 in appendix B.1 of Ref. [5]).

We average over all available values of $t_{\mathrm{sep}}^{\mathrm{min}}$, subject to the constraint that at least three values of t_{sep} are contained in the underlying fit to eq. (6.58). It should be stressed that the only quantity that is effectively restricted by this method is the minimal source-sink separation; all fits go up to the largest t_{sep} we have computed. Essentially, the window average merely serves as a smoothing of the lower end of the fit interval.

This strategy is illustrated in fig. 6.20 for the first non-vanishing momentum on the ensembles D450 and E300. One can see that the window average agrees in all displayed cases within its error with what one might identify as a plateau in the blue points. This being valid to a similar degree on all other ensembles and for all momenta employed in the analysis, we conclude that the window method reliably identifies the asymptotic value of the effective form factors. Moreover, it reduces the human bias compared to manually picking one particular value for $t_{\mathrm{sep}}^{\mathrm{min}}$ on each ensemble, since we use the same window parameters in units of t_0 on all ensembles.

It is important to note that even if a plateau appears to be reached, this does not guarantee ground-state dominance. The situation is aggravated by the fact that in general, relatively few values of $t_{\mathrm{sep}}^{\mathrm{min}}$ are available, and correlated fluctuations in any direction can easily be mistaken for a plateau. This underlines once more the necessity of an automated strategy such as the window average which can readily be applied to all ensembles and momenta. The size of the gray error bands in fig. 6.20 furthermore shows that our window average yields, in contrast to error-weighted procedures, an error estimate which is comparable to the errors of the individual points entering the average. Thus, we are convinced that our error estimates are conservative enough to exclude any systematic bias in the identification of ground-state form factors.

The form factor values obtained from the LO summation method with the window average can be found in appendix D of Ref. [5]. These are the data we take as a basis for the further analysis described in the following sections.

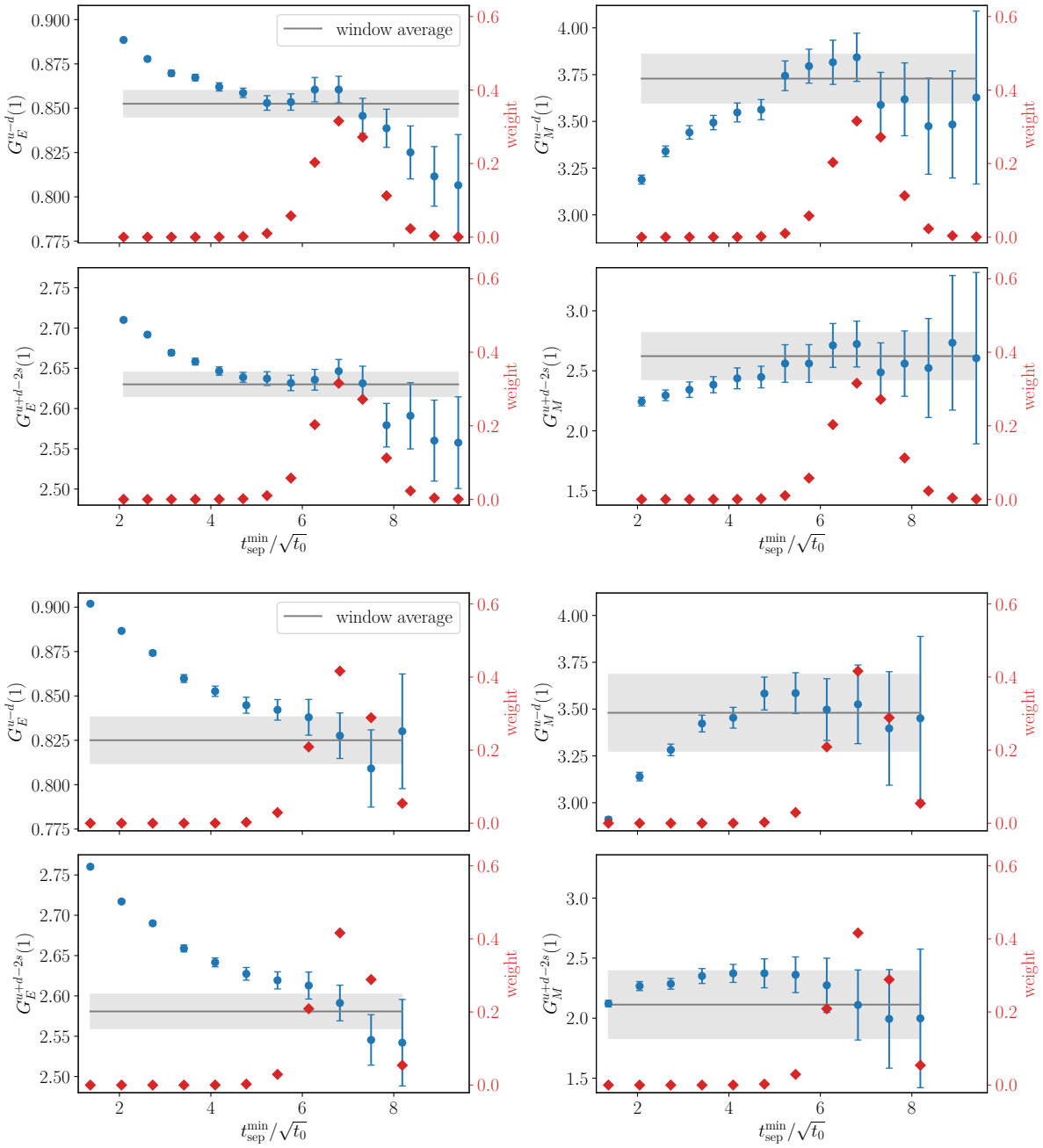


Figure 6.20: Isovector and isoscalar electromagnetic form factors at the first non-vanishing momentum on the ensembles D450 (upper panel) and E300 (lower panel) as a function of the minimal source-sink separation entering the fits to eq. (6.58). Each blue point corresponds to a single fit starting at the source-sink separation given on the horizontal axis. The associated weights derived from eq. (6.60) are represented by the red diamonds, with the gray curves and bands depicting the averaged results.

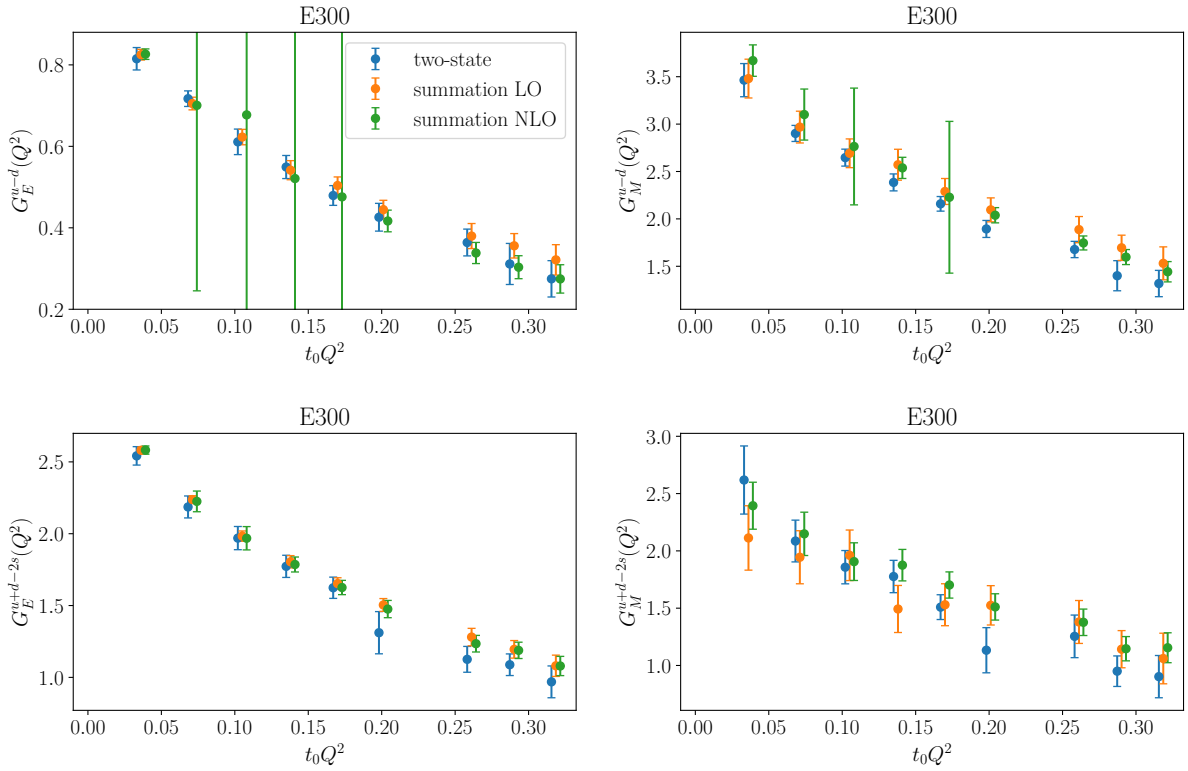


Figure 6.21: Isovector and isoscalar electromagnetic form factors on the ensemble E300 as a function of Q^2. The blue points originate from two-state fits to the effective form factors according to eq. (6.53), the orange ones have been obtained from the LO summation method using the window average and the green ones from fits to the simplified NLO summation *ansatz* eq. (6.59). The points are horizontally displaced for better visibility.

6.6.5 Comparison of different analysis techniques

We have already amply explained our reasons to favor the LO summation method with the window average over the other analyses techniques above. Nevertheless, even if the results of our preferred procedure appear to yield very conservative error estimates, a comparison between different approaches is necessary in order to exclude any systematic bias introduced by relying exclusively on the LO summation method. In figs. 6.21 to 6.23, we show plots of the Q^2-dependence of G_E and G_M on the ensembles E300, D200 and C101, comparing the three different methods discussed above.[6] These plots reveal that among the three techniques the LO summation method generally yields the largest errors, in particular for the magnetic form factors. We remark that many other lattice studies of nucleon form factors have observed a similar trend in the errors [189]. Besides, the comparison in figs. 6.21 to 6.23 does not permit the conclusion that any of the methods introduces a directed, systematic bias.

Any judgment about the reliability of a method to extract the ground-state form factors should also be based on plots of the effective form factors themselves. These can be found in figs. 6.14 to 6.16 above for the three aforementioned ensembles and $Q^2 \approx 0.2$ GeV^2. Apart from the shortcomings of the two-state fits discussed above, these plots demonstrate that the LO summation method yields entirely plausible values for the ground-state form factors.

[6] For the NLO summation method, we have chosen $t_{\text{sep}}^{\text{min}} \approx 0.5$ fm and $t_{\text{skip}} = 2a$, as in fig. 6.19.

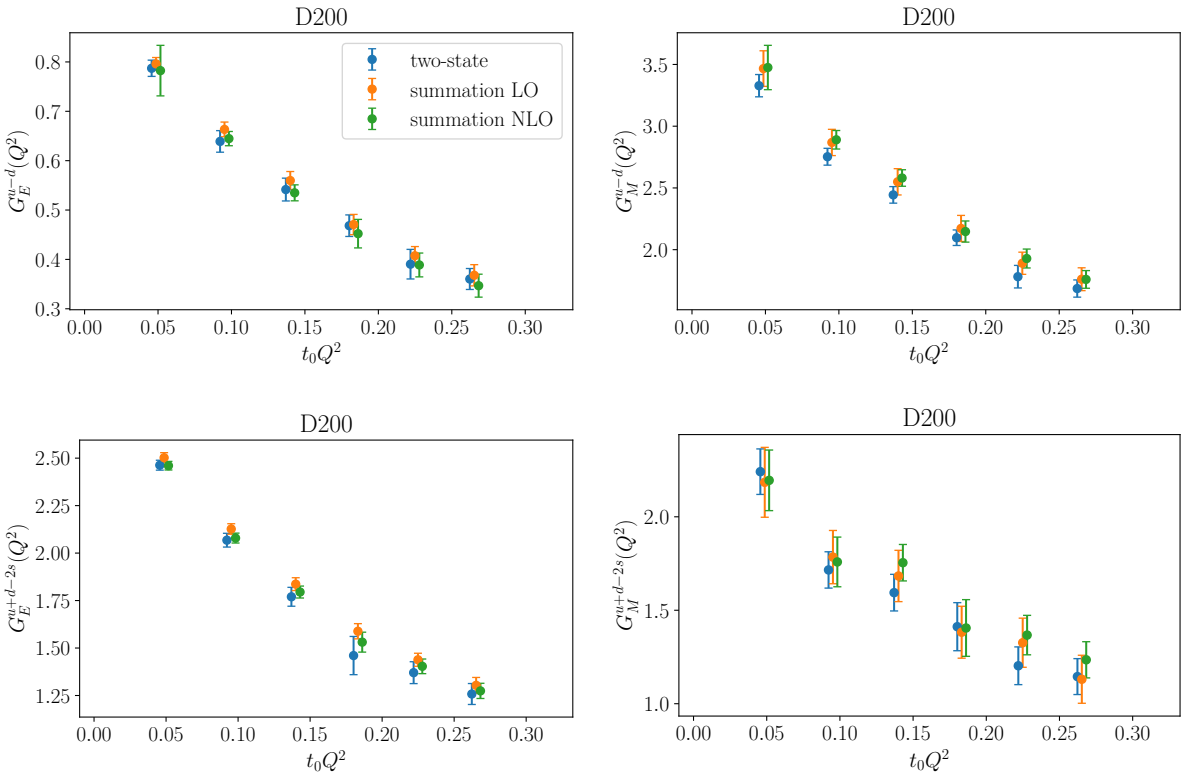


Figure 6.22: Same as fig. 6.21 for ensemble D200

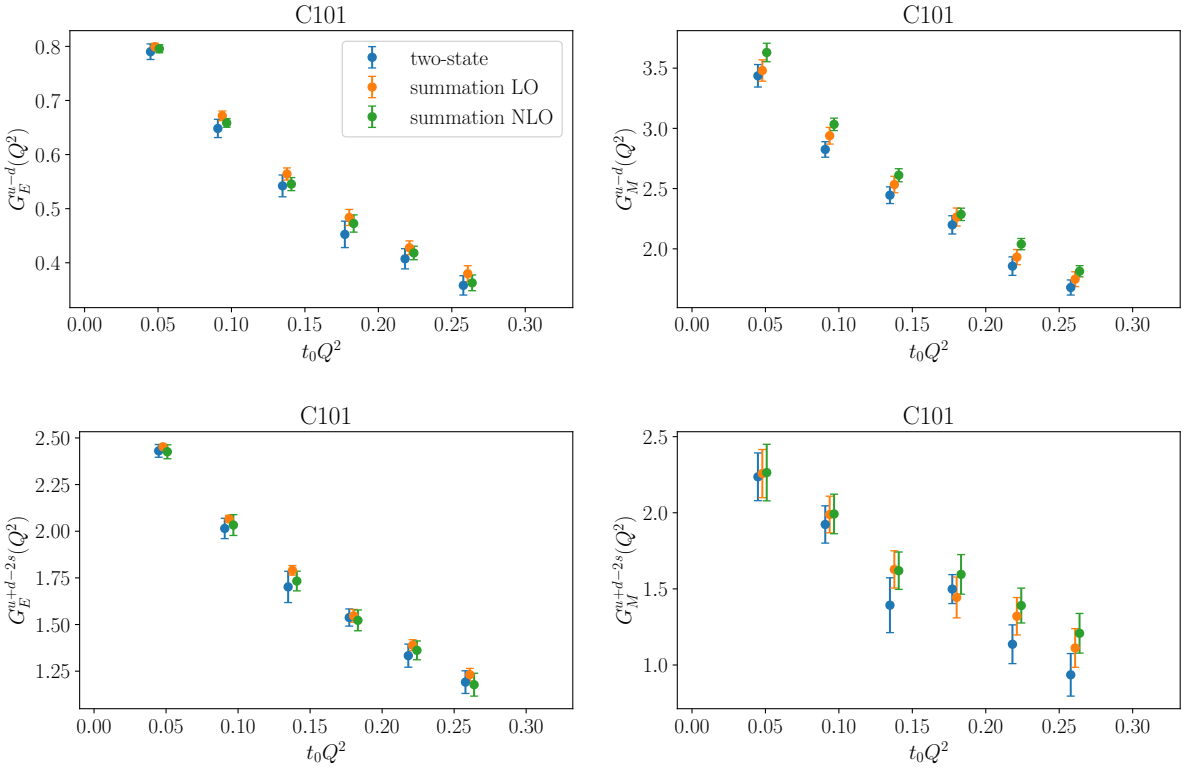


Figure 6.23: Same as fig. 6.21 for ensemble C101

To further quantify the effect on the resulting radii of choosing either the LO summation method or the two-state fits for the extraction of the ground-state form factors, we have performed BχPT fits (*cf.*, section 6.7.1 below) on individual ensembles (the ones mentioned above) for both data sets.[7] This may not be the exact same method used to obtain our final results, which we will describe in detail in section 6.7.1 below. Nevertheless, it employs the same functional forms and permits a relatively straightforward comparison of the two data sets on the level of individual ensembles. We stress that for the purpose of the subsequent comparison, we have subjected both data sets to exactly the same procedure. While we observe some variations in the results and the correlated difference of the radii extracted either from the LO summation or the two-state data can be larger than $1\,\sigma$, there is absolutely no clear pattern to see. On the contrary, the variations appear to be completely random in nature.

To summarize the discussion, we did not find any indication that the LO summation method introduces a systematic bias compared to two-state fits to the effective form factors or the NLO summation method. The opportunity to avoid the use of priors in this very sensitive and crucial step of the analysis is our main reason for preferring the LO summation method. For the full details of our arguments, which we will not repeat here, we refer to the previous subsections.

6.7 Parametrization of the Q^2-dependence and extrapolation to the physical point

One of the major goals of this thesis is to compute the electromagnetic radii of the proton and neutron. As the radii are defined in terms of the slope of the form factors at zero momentum transfer (*cf.*, eq. (5.9)), a description of their Q^2-dependence is necessary. Proceeding analogously to Refs. [53, 60], we apply two different methods: Our preferred strategy is to combine the parametrization of the Q^2-dependence with the chiral, continuum and infinite-volume extrapolation by performing a simultaneous fit to the Q^2-, pion-mass, lattice-spacing and finite-volume dependence of our form factor data directly to the NLO expressions resulting from covariant baryon chiral perturbation theory (BχPT) [197]. This is explained in detail in section 6.7.1. Alternatively, one can follow the more traditional approach of first extracting the radii on each ensemble from a generic parametrization of the Q^2-dependence and subsequently extrapolating them to the physical point. A crosscheck of our main analysis with this two-step procedure is presented in sections 6.7.2 and 6.7.3.

6.7.1 Direct BχPT fits

For our main analysis using the direct (simultaneous) fits, we fit our data for the form factors to the full expressions of Ref. [197] without explicit Δ degrees of freedom. The fits are performed for the isovector and isoscalar channels separately, but for G_E and G_M simultaneously. This allows us to take the correlation not only between different Q^2, but also between G_E and G_M into account. The ensembles, on the other hand, are treated as statistically independent. $G_E(0)$ is fixed by fitting the normalized ratio $G_E(Q^2)/G_E(0)$.

The BχPT expansion from which our fit formulae are derived is only applicable for low momentum transfers. Therefore, we incorporate the contributions from the relevant vector mesons in the expressions for the form factors. In this way, the range of validity of the resulting expressions can be extended [197–199]. In the isovector case, we take the ρ-meson contributions into account, while in the isoscalar channel, we include the leading-order terms from the ω and ϕ resonances. Because the loop diagrams involving ω or ϕ resonances only yield small numerical

[7] We have not carried out this comparison for the NLO summation method because of the observed instabilities in its results.

contributions to the form factors, our fits depend only marginally on them. This means that the corresponding low-energy constants (LECs) g_ω and g_ϕ are very poorly constrained by our data, so that we neglect these loop diagrams. The corresponding tree-level diagrams, on the other hand, only depend on the combinations $c_\omega = f_\omega g_\omega$ and $c_\phi = f_\phi g_\phi$, respectively [197]. Thus, we only use the products c_ω and c_ϕ as independent fit parameters. To summarize, the fit for the isoscalar form factors depends linearly on the LECs d_7, c_7, c_ω and c_ϕ, whereas for the isovector form factors, the relevant LECs are d_6, $\tilde{c}_6$, d_x and G_ρ [197].

The nucleon (average of the proton and neutron) and ϕ masses are fixed to their physical values [76]. Moreover, we replace the pion decay constant and the axial-vector coupling constant in the chiral limit, which appear in the BχPT formulae, by their physical values, $F_\pi = 92.2$ MeV [115] and $g_A = 1.2754$ [76], respectively. We also use the KSRF relation [200–202] $g^2 = M^2_{\rho,\mathrm{phys}}/(2F_\pi^2)$. For the numerical evaluation of one-loop integrals, we make use of `LoopTools` [203, 204], setting the renormalization scale to $\mu = 1$ GeV.

For further details on the LECs, we refer to table III in Ref. [53]. The full expressions for the form factors which we employ can be found in appendix D.2.1 of Ref. [205]. Starting from the formulae for the Dirac and Pauli form factors given there, we form the appropriate linear combinations for the electric and magnetic Sachs form factors according to eq. (5.5).

The mass of the ρ meson is set on each ensemble to the value at the corresponding pion mass and lattice spacing. This is determined from a parametrization of the pion-mass and lattice-spacing dependence of a subset of the values for M_ρ/M_π obtained in Ref. [164],

$$\frac{M_\rho}{M_\pi} = \frac{M_{\rho,\mathrm{phys}}}{M_{\pi,\mathrm{phys}}} + A\left(\frac{1}{\sqrt{t_0}M_\pi} - \frac{1}{\sqrt{t_{0,\mathrm{phys}}}M_{\pi,\mathrm{phys}}}\right) + C(\sqrt{t_0}M_\pi - \sqrt{t_{0,\mathrm{phys}}}M_{\pi,\mathrm{phys}}) + D\frac{a^2}{t_0}, \quad (6.61)$$

with the independent fit parameters A, C and D. For the fit to this formula, we disregard ensembles which are not included in our main analysis (*cf.*, table 4.1), or which are solely used to study finite-volume effects (H105 and S201), since the finite-volume dependence of the ρ masses is not sufficiently constrained by the data. The original data of Ref. [164] is shown together with the fit in fig. 6.24. This, together with the very high p-value of about 72.5 % which we obtain, demonstrates that our parametrization eq. (6.61) essentially only serves to smooth the pion-mass and lattice-spacing dependence of the ρ mass, while being compatible with the actual data within errors.

On each ensemble, we set the ω mass equal to the ρ mass obtained from eq. (6.61), because no lattice data for the ω masses on our ensembles are available, and the mass splitting between the ρ and ω mesons is small. This means that the true ω mass on our ensembles is probably much closer to the ensemble-dependent ρ mass than to the physical value of the ω mass. The ϕ resonance, on the other hand, is much heavier than the ρ and ω mesons. In the absence of lattice data for the ϕ mass on our ensembles, we thus employ the physical value $M_{\phi,\mathrm{phys}}$ [76] in our fits. The physical pion mass $M_{\pi,\mathrm{phys}}$ is fixed in units of $\sqrt{t_0}$ using its value in the isospin limit (*cf.*, eq. (4.12)), *i.e.*, we employ $\sqrt{t_{0,\mathrm{phys}}}M_{\pi,\mathrm{phys}} = 0.09881(59)$. Here, we neglect the uncertainty of $M_{\pi,\mathrm{iso}}$ in MeV since it is completely subdominant compared to that of $\sqrt{t_{0,\mathrm{phys}}}$, which enters in the unit conversion and is propagated into the fits (see below). For the pion masses we use on our ensembles, see table 6.1.

Two of the major benefits of the approach presented in this subsection as compared to the two-step procedure described in the next ones are the following: On the one hand, performing a fit across several ensembles significantly decreases the errors on the resulting radii. On the other hand, it leads to a much larger number of degrees of freedom for the fit. This increases the stability against lowering the applied momentum cut considerably. These advantages have already been noticed in Ref. [60] and apply in a similar manner to the data presented in this thesis.

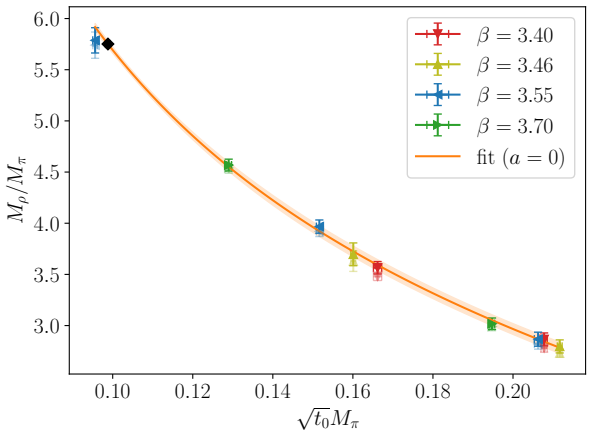


Figure 6.24: Values of M_ρ/M_π from Ref. [164] as a function of the pion mass. The faint symbols represent the original lattice data, while the opaque ones have been corrected for the continuum limit. The orange line and band depict our parametrization according to eq. (6.61) evaluated at $a = 0$. The black diamond indicates the physical point which the fit is forced to reproduce.

We perform several such fits, using different models to describe the lattice-spacing and/or finite-volume dependence and, at the same time, applying various cuts in the pion mass ($M_\pi \leq 0.23$ GeV and $M_\pi \leq 0.27$ GeV) and momentum transfer ($Q^2 \leq 0.3, \ldots, 0.6$ GeV^2), in order to estimate the corresponding systematic uncertainties. The least stringent value for the cut in Q^2 is motivated by the fact that the heaviest vector meson we consider in the isovector channel is the ρ with $M^2_{\rho,\text{phys}} \approx 0.6$ GeV^2. The variations of the results due to the cuts are in most cases much smaller than their statistical errors. In any case, these variations will be included in our systematic errors by means of a model average (*cf.*, section 6.10 below). Moreover, the p-values of all our direct fits remain on an acceptable level (*cf.*, appendix E of Ref. [5]). We conclude that we do not observe any sign of a breakdown of the BχPT expansions in the aforementioned range of pion masses and momentum transfers.

We adopt two different models for lattice artefacts, either based on an additive or a multiplicative *ansatz* [60],

$$G_E^{\text{add}}(Q^2) = G_E^\chi(Q^2) + G_E^a a^2 Q^2 + G_E^L t_0 Q^2 e^{-M_\pi L}, \tag{6.62}$$

$$G_M^{\text{add}}(Q^2) = G_M^\chi(Q^2) + G_M^a \frac{a^2}{t_0} + \kappa_L M_\pi \left(1 - \frac{2}{M_\pi L}\right) e^{-M_\pi L} + G_M^L t_0 Q^2 e^{-M_\pi L}, \tag{6.63}$$

$$G_E^{\text{mult}}(Q^2) = G_E^\chi(Q^2) + \frac{G_E^a a^2 Q^2 + G_E^L t_0 Q^2 e^{-M_\pi L}}{t_0(M_\rho^2 + Q^2)}, \tag{6.64}$$

$$G_M^{\text{mult}}(Q^2) = G_M^\chi(Q^2) + \frac{G_M^a a^2/t_0 + G_M^L t_0 Q^2 e^{-M_\pi L}}{t_0(M_\rho^2 + Q^2)} + \kappa_L M_\pi \left(1 - \frac{2}{M_\pi L}\right) e^{-M_\pi L}. \tag{6.65}$$

The precise form of the multiplicative model has been altered compared to the one used in Ref. [60], where the correction terms directly multiplied $G^\chi_{E,M}(Q^2)$. With our updated, more precise data we have found that such terms containing both $G^{a,L}_{E,M}$ and the LECs (via $G^\chi_{E,M}(Q^2)$) lead to instabilities in the determination of $G^{a,L}_{E,M}$. This is most probably due to the fit becoming nonlinear in the fit parameters. By contrast, our new model in eqs. (6.64) and (6.65) is, from a technical point of view, also purely additive and thus linear in the fit parameters, while still

capturing the essential contribution of $G^{\chi}_{E,M}(Q^2)$ to the fall-off of the form factors with rising momentum transfer.[8]

Fits leaving κ_L as a free parameter are unstable, and we therefore fix κ_L to the value from heavy-baryon chiral perturbation theory (HBχPT) [206],

$$\kappa_L = -\frac{m_{N,\mathrm{phys}} g_A^2}{4\pi F_\pi^2}\tau_3. \tag{6.66}$$

Here, $\tau_3 = +1$ for the proton and $\tau_3 = -1$ for the neutron. Following eq. (6.71) below, this implies that $\tau_3^{u-d} = +2$ and $\tau_3^{u+d-2s} = 0$. We remark that this does not mean that the isoscalar magnetic moment cannot receive any finite-volume correction at all; it merely means that it is of higher order in HBχPT. Since the value in the isovector channel is already very small, one can thus expect the one in the isoscalar channel to be even smaller, so that we indeed set it to zero. In total, we have seven different fit models: one without any parametrization of lattice artefacts, three including discretization and/or finite-volume effects with the additive model of eqs. (6.62) and (6.63), and three corresponding ones using the multiplicative prescription of eqs. (6.64) and (6.65). The results for all of them are collected in appendix E of Ref. [5].

The inclusion of a term describing lattice artefacts requires the use of Gaussian priors (*cf.*, section 6.2.3) for the relevant coefficients to stabilize the fits. For this purpose, we first perform fits to ensembles at $M_\pi \approx 0.28$ GeV only (N101, H105, N451, N200, S201 and J303; *cf.*, table 4.1). Here, we have relatively precise data in a wide range of lattice spacings and volumes. For these fits, we use a cut in Q^2 at 0.6 GeV^2 and a simultaneous description of the lattice-spacing and finite-volume dependence. The coefficients for the correction terms as determined from the fits, together with their associated errors, are then employed as priors for the final fits to the ensembles satisfying the aforementioned cuts in the pion mass. The only exception are the coefficients G_E^a parametrizing the lattice-spacing dependence of the isovector electric form factor: our data for G_E^{u-d} is sufficiently precise even at low pion masses to allow a determination of G_E^a, so that we can leave it as a free parameter.

Furthermore, we have checked that the results for the radii and magnetic moments are stable against increasing the width of the priors. This is illustrated in fig. 6.25 for the fits with $M_{\pi,\mathrm{cut}} = 0.23$ GeV and $Q^2_{\mathrm{cut}} = 0.6$ GeV^2, and corrections for both types of lattice artefacts according to the additive model of eqs. (6.62) and (6.63). The results for the electromagnetic radii and the magnetic moments using the procedure explained above are well consistent with those obtained by multiplying the prior width by a factor of 2 or 3. We find similarly good agreement between the fits with prior width factors ≥ 1 in all other cases. Moreover, we note that almost all of our priors already have a relative width of $\gtrsim 100$ % even without increasing it.

Since the number of configurations and hence the number of Jackknife samples differ between ensembles (*cf.*, table 4.1), we use the parametric bootstrap procedure described in section 6.1.5 to resample the distributions on each ensemble. With all mean values for the form factors entering a specific fit, the nucleon and the pion mass, as well as their covariance matrix, we draw 10 000 random samples from a corresponding multivariate Gaussian distribution. The covariance matrix one can build from these samples is consistent with the original covariance matrix. Moreover, the parametric bootstrap procedure enables us to account for the errors of the scale parameters t_0^{sym}/a^2 and $\sqrt{t_{0,\mathrm{phys}}}$, as well as M_ρ/M_π, which are external to this analysis. Hence, we create an independent random Gaussian distribution for $\sqrt{t_{0,\mathrm{phys}}}$, for t_0^{sym}/a^2 at each value of β, and for M_ρ/M_π on each ensemble.

[8] Note that because we employ M_ρ in place of M_ω on our ensembles, the expressions in eqs. (6.64) and (6.65) are valid for the isovector and isoscalar channels alike.

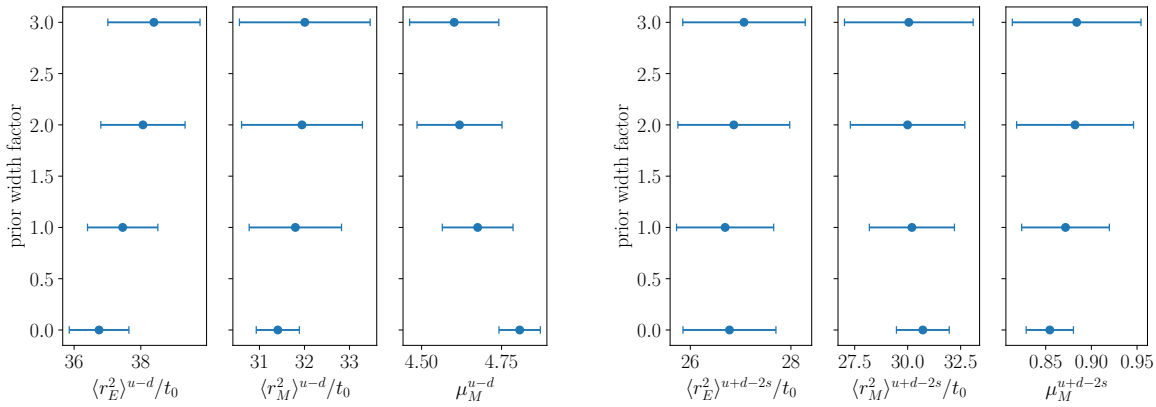


Figure 6.25: Isovector and isoscalar electromagnetic radii and magnetic moments for factors of 1×10^{-5}, 1, 2 and 3 multiplying the prior widths determined as described in the text. The numbers are obtained from direct fits with $M_{\pi,\mathrm{cut}} = 0.23$ GeV and $Q^2_{\mathrm{cut}} = 0.6$ GeV^2, employing the model of eqs. (6.62) and (6.63). We prefer the results using a factor of 1.

From the fits in the isovector and isoscalar channels, we reconstruct the form factors and all derived observables for the proton and neutron. For this purpose, we build the appropriate linear combinations of the BχPT formulae and plug in the LECs as determined from the fits to the isovector and isoscalar form factors. This is the more natural approach both from the perspective of Lattice QCD and of chiral perturbation theory: The form factors in the isospin basis are our primary[9] lattice observables, while the proton and neutron form factors can only be obtained indirectly as linear combinations of them (*cf.*, eq. (6.71) below). For the BχPT fits, the isospin basis is also advantageous because of the clear separation of the contributing resonances in the isovector and isoscalar channels, so that there are no common fit parameters between the two of them. This allows us to fit the two channels separately, thus avoiding the covariance matrices becoming impractically large.

The quality of the direct fits is illustrated in fig. 6.26 for our two most chiral ensembles E250 and E300. The fits shown here correspond to the additive model of eqs. (6.62) and (6.63) employed to parametrize discretization and finite-volume effects, with $M_{\pi,\mathrm{cut}} = 0.23$ GeV and $Q^2_{\mathrm{cut}} = 0.6$ GeV^2. In general, the fits describe the data very well. We observe that the error of the fits is significantly reduced compared to the data points on E250, but only slightly on E300. The latter is also the case on all other ensembles entering the displayed fits (D200, D450 and C101, which are not shown in fig. 6.26). We conclude that the error reduction on E250 is due to the global fit, *i.e.*, the inclusion of several ensembles in one fit, with the data at larger pion masses being more precise than near $M_{\pi,\mathrm{phys}}$.

For the electric form factors on E250, we find a slight deviation between the fit and the data, which is mostly absent in all other cases. Nevertheless, the p-values of the shown fits are acceptable, with $p > 0.15$ in both channels, because the data points are highly correlated, so that actually fewer degrees of freedom deviate than it naively appears from the plots in fig. 6.26. Also, the fits are more consistent with the data on most of the other ensembles. Consequently, the deviation on E250 could in principal either be caused by residual excited-state contamination or by more pronounced correlated statistical fluctuations which are not sufficiently suppressed

[9]The form factors in the isospin basis are also secondary quantities in the sense that they are functions of gauge averages due to the reweighting and the ratio method (*cf.*, eqs. (3.71) and (6.43), respectively).

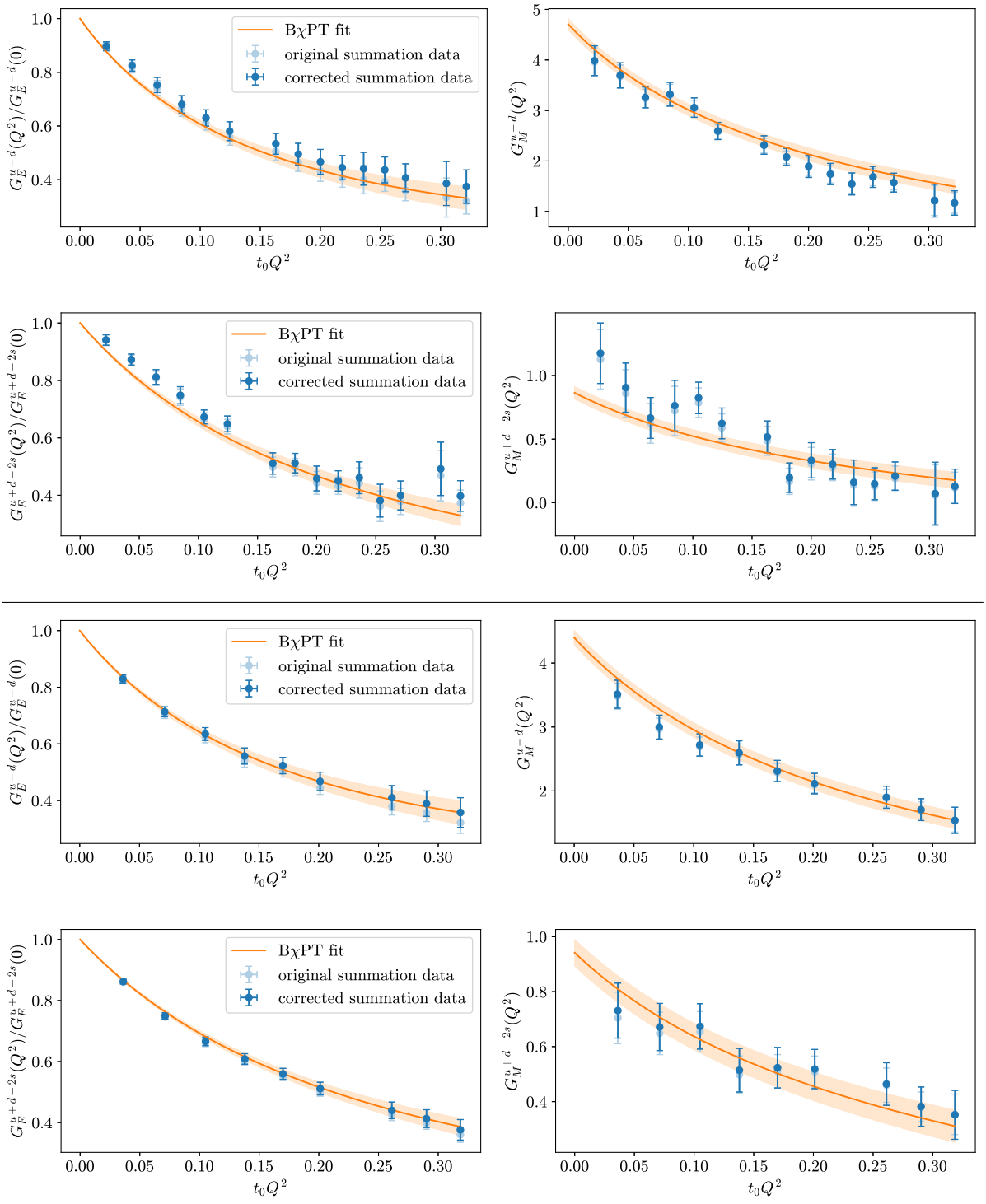


Figure 6.26: Isovector and isoscalar electromagnetic form factors on the ensembles E250 (upper panel) and E300 (lower panel) as a function of Q^2. Our original lattice data as obtained from the LO summation method using the window average are represented by the faint blue points, while the opaque ones have been corrected for the continuum and infinite-volume limit. The orange curves and bands depict direct fits with $M_{\pi,\mathrm{cut}} = 0.23$ GeV and $Q^2_{\mathrm{cut}} = 0.6$ GeV^2, evaluated at the pion mass of the respective ensemble, zero lattice spacing and infinite volume.

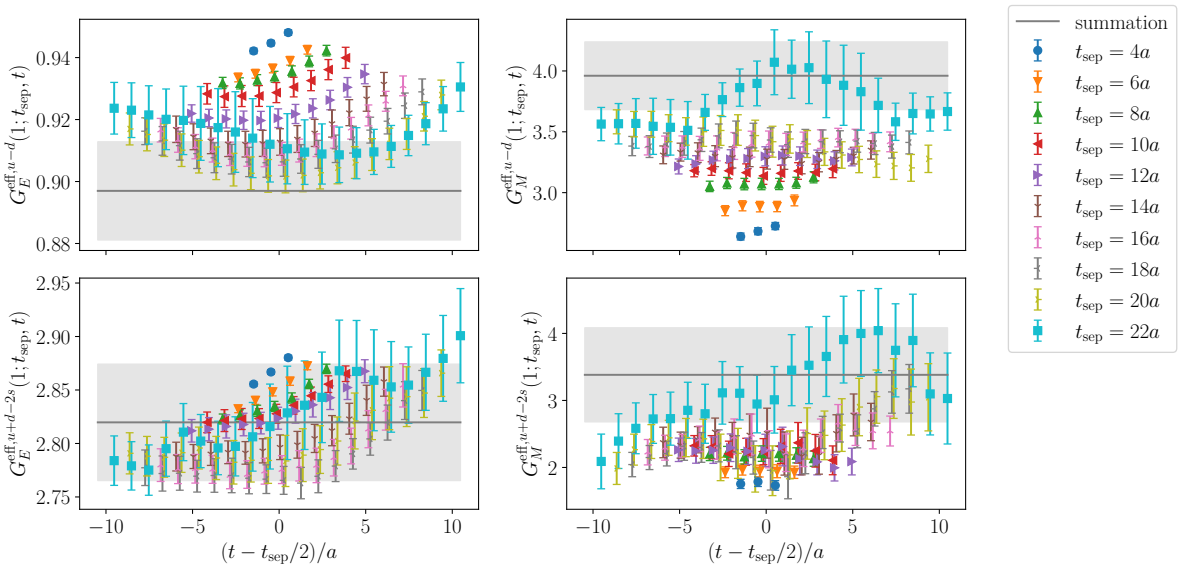


Figure 6.27: Effective form factors for the isovector (upper panel) and isoscalar (lower panel) combinations at the first non-vanishing momentum on the ensemble E250 ($Q^2 \approx 0.041$ GeV^2). The meaning of the points and bands is analogous to fig. 6.11.

by our choice of window parameters in the summation method. In the following, we will investigate this point in more detail.

In fig. 6.27, the effective form factors are shown for the first non-vanishing momentum on E250. It is obvious that the largest two source-sink separations ($t_{\mathrm{sep}} = 20a$ and $22a$) represent an upwards fluctuation. In the electric form factors, this is particularly clear because excited-state effects always have a positive sign here, so that the effective form factors are expected to monotonically decrease with rising t_{sep}. Moreover, doubling the statistics for the disconnected part from the original 398 to now 796 configurations has had (almost) no effect on the largest two source-sink separations as far as the errors are concerned, while reducing the errors of the disconnected contribution substantially for the lower values of t_{sep}. This indicates likewise that fluctuations are still dominant for the largest two source-sink separations.

The upwards fluctuation is also clearly visible in the $t_{\mathrm{sep}}^{\mathrm{min}}$-plots for the isoscalar form factors extracted from the summation method (*cf.*, fig. 6.28 (bottom)). Here, mostly the extractions with the largest two values of $t_{\mathrm{sep}}^{\mathrm{min}}$ are affected, as in these cases the influence of the effective form factors at large t_{sep} on the summation fit becomes sizable. Due to our choice for the window, these points have a significant effect on the averaged result as well. This means that on E250, the window average with our values for the parameters $t_w^{\mathrm{up,low}}$ is not able to suppress this statistical fluctuation sufficiently in the isoscalar channel, while the isovector results look somewhat more stable. On the other hand, excited-state effects are also expected to be stronger at lower pion masses. Hence, it does not appear reasonable to adjust the window to lower values of $t_{\mathrm{sep}}^{\mathrm{min}}$ on E250. Besides, the lowest point for G_E^{u+d-2s} in fig. 6.28 is still within $2\,\sigma$ of our averaged result, so that these two values are not incompatible with each other, and our error is not grossly underestimated.

As can be seen from fig. 6.26, our direct fits are stable against such fluctuations on single ensembles: the fit curves lie below the data at the first ~ 6 Q^2-points for the isoscalar form factors on E250, *i.e.*, the fit follows much more closely the trend determined by the other ensembles than this obvious fluctuation on E250. Regarding the isovector form factors, the fact

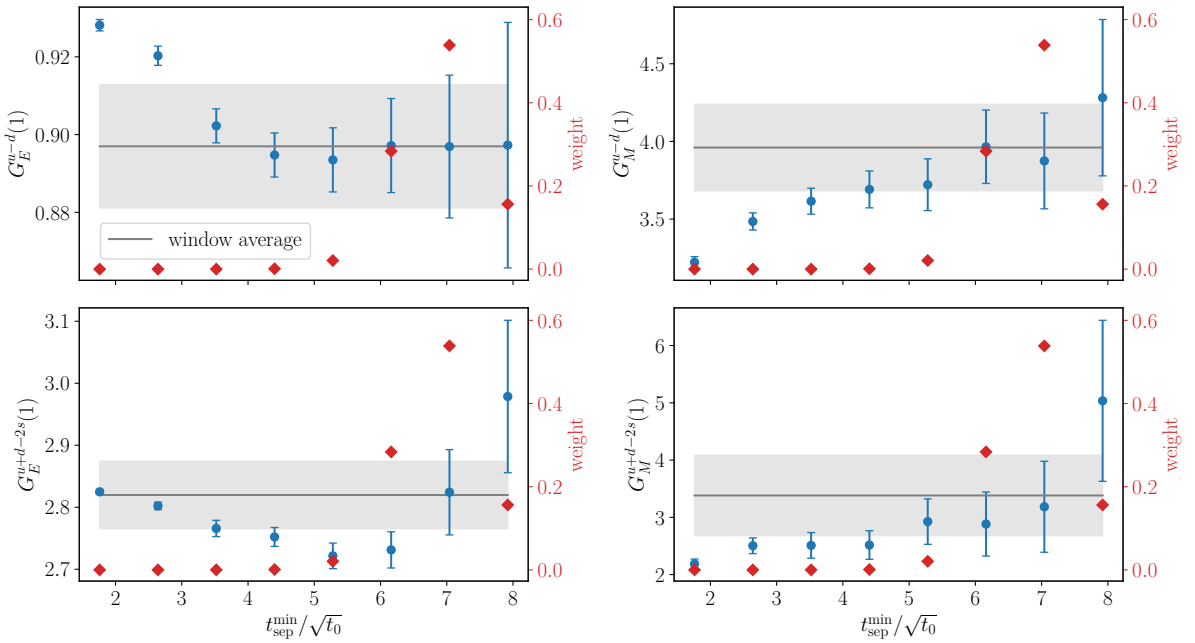


Figure 6.28: Isovector and isoscalar electromagnetic form factors at the first non-vanishing momentum on the ensemble E250 as a function of the minimal source-sink separation entering the fits to eq. (6.58). The meaning of the points and bands is analogous to fig. 6.20.

that the fit lies somewhat below the data for G_E^{u-d} and somewhat above for G_M^{u-d} on E250, but to a much lesser degree so on ensembles with heavier pion masses, has already been noticed in Ref. [60], and is qualitatively confirmed by this updated study.

Note that the curves shown in fig. 6.26 only correspond to one specific model and are thus not be interpreted as our definitive results for the form factors. These can be found in fig. 6.37 below for the proton and neutron.

6.7.2 The z-expansion

As an alternative to the direct fits, one can treat the parametrization of the Q^2-dependence and the chiral, continuum and infinite-volume extrapolation as two separate steps. For this purpose, an *ansatz* for the Q^2-dependence of the form factors on each ensemble is required.

In the early days of *ep*-scattering experiments, a popular model for the charge and magnetization distributions inside the proton was an exponential $\rho(r) \propto e^{-r/r_0}$ [13]. As noted in section 5.1, the corresponding form factors can, at low Q^2, be obtained by Fourier transformation, which yields a dipole form [207],

$$G(Q^2) = \frac{G(0)}{(1 + Q^2/M_D^2)^2}. \tag{6.67}$$

With the standard value of the dipole mass, $M_D^2 = 0.71\ \mathrm{GeV}^2$ [147], this could be shown to fit the experimental data for the electric as well as the magnetic form factor of the proton at the time [13]. This is still true within ≈ 10 % for current data [147]. From a theoretical point of view, this can be understood to be due to two nearby vector-meson poles V_1 and V_2 with opposite residua r and $-r$, respectively, in the time-like region ($Q^2 < 0$). In the space-like

region ($Q^2 > 0$), such a model leads to an approximate dipole contribution to the form factors [208],

$$G(Q^2) \sim \frac{r}{q^2 - M_{V_1}^2} + \frac{-r}{q^2 - M_{V_2}^2} = \frac{-r}{Q^2 + M_{V_1}^2} + \frac{r}{Q^2 + M_{V_2}^2} = \frac{r(M_{V_1}^2 - M_{V_2}^2)}{(Q^2 + M_{V_1}^2)(Q^2 + M_{V_2}^2)}. \quad (6.68)$$

Having a more precise look at current *ep*-scattering data, however, reveals significant deviations of the proton form factors from the simple dipole form [147]. Also from a first-principals theoretical point of view, there is no compelling argument why the form factors should *exactly* follow a dipole form. Moreover, the electric form factor of the neutron cannot be described in this way at all, since $G_E^n(0) = 0$. For these reasons, a model-independent *ansatz* is desirable.

This can be achieved by the z-expansion which only makes use of the known analytical properties of the form factors [209]: they are analytical functions of the complex variable q^2 outside of a cut at time-like values of q^2 beginning at the two-pion production threshold, $q^2 \geq 4M_\pi^2$ [210].[10] One hence maps the domain of analyticity onto the unit circle [209],

$$z(Q^2) = \frac{\sqrt{\tau_{\text{cut}} + Q^2} - \sqrt{\tau_{\text{cut}} - \tau_0}}{\sqrt{\tau_{\text{cut}} + Q^2} + \sqrt{\tau_{\text{cut}} - \tau_0}}, \quad (6.69)$$

and expands the form factors as

$$\frac{G_E(Q^2)}{G_E(0)} = \sum_{k=0}^{n} a_k z(Q^2)^k, \qquad G_M(Q^2) = \sum_{k=0}^{n} b_k z(Q^2)^k. \quad (6.70)$$

The parameter τ_0 (not to be confused with the gradient flow scale t_0, hence the slightly uncommon nomenclature) is the value of $q^2 = -Q^2$ which is mapped to $z = 0$. In principle, it can be optimized to minimize the maximum value that $|z|$ takes [209]. In our fits, however, we have not found any appreciable dependence on the value of τ_0, and thus use the simple choice $\tau_0 = 0$. On each ensemble, we set $\tau_{\text{cut}} = 9M_\pi^2$ for the isoscalar channel and $\tau_{\text{cut}} = 4M_\pi^2$ for the remaining channels, respectively, where M_π is the pion mass on the respective ensemble (*cf.*, table 6.1).

For the z-expansion analysis, we use the form factors of the proton and neutron, which can be obtained from the ones in the isospin basis according to eqs. (5.15) and (5.17),

$$G_{E,M}^{\text{eff},p} = \frac{1}{6}\left(G_{E,M}^{\text{eff},u+d-2s} + 3G_{E,M}^{\text{eff},u-d}\right), \quad G_{E,M}^{\text{eff},n} = \frac{1}{6}\left(G_{E,M}^{\text{eff},u+d-2s} - 3G_{E,M}^{\text{eff},u-d}\right). \quad (6.71)$$

In analogy to the direct fits, we fit the normalized ratio $G_E(Q^2)/G_E(0)$ and enforce the normalization by fixing $a_0 = 1$. For the exceptional case of the neutron, where $G_E^n(0) = 0$, we do not normalize $G_E^n(Q^2)$, exclude the point at $Q^2 = 0$, and set $a_0 = 0$. Evaluating eqs. (5.9) and (5.10) yields for the radii,

$$\langle r_E^2 \rangle = -\frac{3a_1}{2\tau_{\text{cut}}}, \qquad \langle r_M^2 \rangle = -\frac{3b_1}{2\tau_{\text{cut}} b_0}, \quad (6.72)$$

where the normalization of the electric radius has already been taken into account, and $b_0 = \mu_M$.

Using the same strategy to set priors on the coefficients a_k and b_k as in Ref. [60], we observe with our updated, more precise data that such priors impose too strict constraints on the Q^2-behavior of the form factors. In particular, the fits with priors tend to follow more closely the data points at large Q^2 than those at low Q^2, which is undesirable for an extraction of the radii and the magnetic moment. Hence, we resort to fits without priors, going up to order $n = 2$.

[10] In the isoscalar channel, the relevant threshold is three-pion production, $q^2 \geq 9M_\pi^2$ [210].

This represents a compromise between the fit function not being unduly stiff and avoiding overfitting especially on ensembles with a bad resolution in Q^2.

The errors on the nucleon and pion masses as well as those from the scale setting are included in the analysis using the same parametric bootstrap procedure as for the direct fits. In order to account for the correlation between G_E and G_M, the two form factors are fitted simultaneously even though they do not share any common fit parameters. However, the different channels (proton, neutron, isovector and isoscalar) are treated separately, in order to mirror the analysis with the direct fits as closely as possible.

One disadvantage of the analysis based on the z-expansion is that the lowest cut in Q^2 which is applicable for all ensembles required for a chiral and continuum extrapolation is roughly 0.6 GeV^2. We thus show in fig. 6.29 the z-expansion fits using this relatively large cut also for our near-physical pion mass ensemble E250. One can see that the fits describe the data very well. In particular for G_M, however, the form factor is (except for the isoscalar channel) very flat around $Q^2 = 0$, where there is no lattice data. This seems to be partly due to fluctuations in the data at $Q^2 > 0$, against which the z-expansion is not sufficiently stable. Consequently, it can be assumed that the magnetic moments and radii determined from these fits are considerably too small (*cf.*, fig. 6.30 below). The situation is aggravated for ensembles with a bad momentum resolution like N200 or J303. On these two ensembles, the z-expansion fits run away in opposite directions for the magnetic form factor towards $Q^2 = 0$. This demonstrates that the z-expansion is much more prone to modelling fluctuations than the BχPT fits, especially for the magnetic form factors.

Using the z-expansion, we obtain a set of results for the electromagnetic radii and magnetic moments on each ensemble. They are listed in appendix F of Ref. [5] for two different cuts in the momentum transfer ($Q^2 \leq 0.6$ GeV^2 and $Q^2 \leq 0.7$ GeV^2). We note that the ensembles dedicated to the study of finite-volume effects (H105 and S201) do not even have enough data points to permit a z-expansion with these momentum cuts.

6.7.3 Chiral, continuum and infinite-volume extrapolation

For the chiral, continuum and infinite-volume extrapolation of the z-expansion data sets, we employ fit formulae inspired by HBχPT [44]. We only take the leading-order dependence on the pion mass into account, since any higher-order coefficients are very poorly constrained by our data, and add terms $\propto a^2$ in order to account for discretization effects. As the finite-volume dependence is not sufficiently constrained by the ensembles on which z-expansion data is available (*cf.*, the discussion in section 6.7.2), we neglect it for the purpose of this crosscheck of our main analysis which is based on the direct fits (*cf.*, section 6.7.1).

This leads us to the following *ansätze* for the isovector, proton and neutron channels,

$$\frac{\langle r_E^2 \rangle}{t_0} = A + D \ln(\sqrt{t_0} M_\pi) + E \frac{a^2}{t_0}, \tag{6.73}$$

$$\frac{\langle r_M^2 \rangle}{t_0} = A + \frac{D}{\sqrt{t_0} M_\pi} + E \frac{a^2}{t_0}, \tag{6.74}$$

$$\mu_M = A + B \sqrt{t_0} M_\pi + E \frac{a^2}{t_0}. \tag{6.75}$$

The terms in the one-loop HBχPT expressions for the electromagnetic radii and the magnetic moment with the pion-mass dependence shown in eqs. (6.73) to (6.75) do not contribute in the isoscalar channel [44, 197, 211]. As argued in Ref. [44], one may nevertheless need to introduce a counterterm $\propto M_\pi^2$ parametrizing quark-mass dependent short-distance contributions to the isoscalar (anomalous) magnetic moment. This term should also enter the electric radius via the

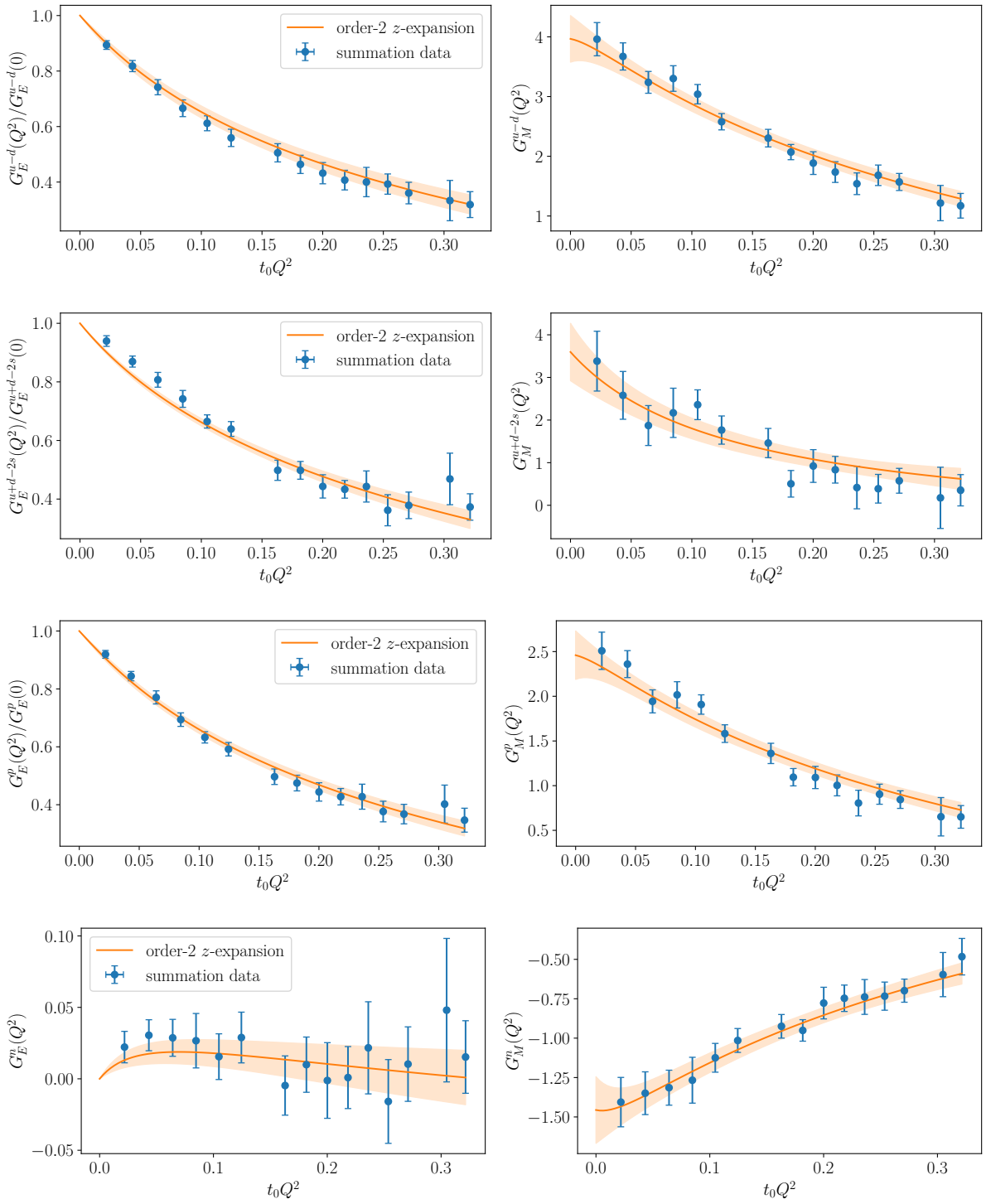


Figure 6.29: Electromagnetic form factors as obtained from the LO summation method using the window average on the ensemble E250 as a function of Q^2. From top to bottom, we show the isovector and isoscalar combinations, as well as the proton and neutron. In contrast to fig. 6.26, the orange curves and bands depict here z-expansion fits with $Q^2_{\text{cut}} = 0.6$ GeV^2.

second term in eq. (5.5) (left). In the absence of any concrete higher-order HBχPT results, we employ the following *ansatz* for all three isoscalar observables,

$$A + Ct_0 M_\pi^2 + E\frac{a^2}{t_0}. \tag{6.76}$$

The extrapolated results at the physical point are collected in table 6.2 for two different cuts each in the pion mass ($M_\pi \leq 0.27$ GeV and $M_\pi \leq 0.3$ GeV) and the momentum transfer ($Q^2 \leq 0.6$ GeV^2 and $Q^2 \leq 0.7$ GeV^2). The numbers for $\langle r_E^2 \rangle$ are, with the slight exception of the neutron, stable within their errors and compare well to the results of the direct fits (*cf.*, table 6.3 below). The magnetic radii exhibit somewhat more variation while having a considerably larger error compared with the direct fits. This is mostly due to the latter fitting G_E and G_M together with common fit parameters, thus leveraging the knowledge that both form factors are governed by the same underlying physics. For the magnetic moments, the agreement with the direct fits tends to be worse than for the radii: they are (except in the isoscalar channel) significantly smaller in magnitude than those obtained from the direct fits (*cf.*, table 6.3) which are in turn well compatible with the experimentally very precisely known values (*cf.*, fig. 6.38).

Table 6.2: Results of the chiral and continuum extrapolation for the z-expansion extractions

Channel	$M_{\pi,\mathrm{cut}}$ [GeV]	Q^2_{cut} [GeV^2]	$\langle r_E^2 \rangle$ [fm^2]	$\langle r_M^2 \rangle$ [fm^2]	μ_M	p-value
$u-d$	0.27	0.6	0.740(80)	0.38(27)	4.09(36)	0.562
$u-d$	0.27	0.7	0.727(79)	0.54(24)	4.02(34)	0.364
$u-d$	0.30	0.6	0.739(54)	0.65(16)	4.04(26)	0.0573
$u-d$	0.30	0.7	0.750(53)	0.81(14)	4.04(25)	0.0676
$u+d-2s$	0.27	0.6	0.557(31)	0.52(34)	2.43(46)	0.0557
$u+d-2s$	0.27	0.7	0.551(29)	0.52(29)	2.41(45)	0.247
$u+d-2s$	0.30	0.6	0.553(20)	0.44(23)	2.40(35)	0.167
$u+d-2s$	0.30	0.7	0.550(19)	0.47(17)	2.39(32)	0.350
p	0.27	0.6	0.643(56)	0.23(32)	2.33(24)	0.414
p	0.27	0.7	0.625(53)	0.37(29)	2.34(23)	0.453
p	0.30	0.6	0.668(37)	0.57(19)	2.37(17)	0.131
p	0.30	0.7	0.667(35)	0.70(17)	2.38(17)	0.150
n	0.27	0.6	−0.132(37)	0.54(33)	−1.65(16)	0.687
n	0.27	0.7	−0.125(36)	0.80(29)	−1.65(15)	0.272
n	0.30	0.6	−0.084(24)	0.75(20)	−1.57(12)	0.0648
n	0.30	0.7	−0.094(23)	0.93(17)	−1.59(11)	0.0634

For illustration, we display in fig. 6.30 the extrapolation for the proton using $M_{\pi,\mathrm{cut}} = 0.3$ GeV and $Q^2_{\mathrm{cut}} = 0.6$ GeV^2. The radii, especially $\langle r_M^2 \rangle$, are very small on our near-physical pion mass ensemble E250, with large errors. Moreover, individual ensembles can deviate rather strongly from the fit curve, which is most apparent for the magnetic observables on J303 and N200. This is probably due to the low momentum resolution on these two ensembles, which implies a long extrapolation to $Q^2 = 0$, where the radius and the magnetic moment are defined, but no lattice data are available. In this context, we also refer to the discussion of fig. 6.29 in section 6.7.2 above. We furthermore note that the relative weights in the extrapolation fit do not reflect the number of Q^2-points entering the z-expansion, which is different on each ensemble. In this

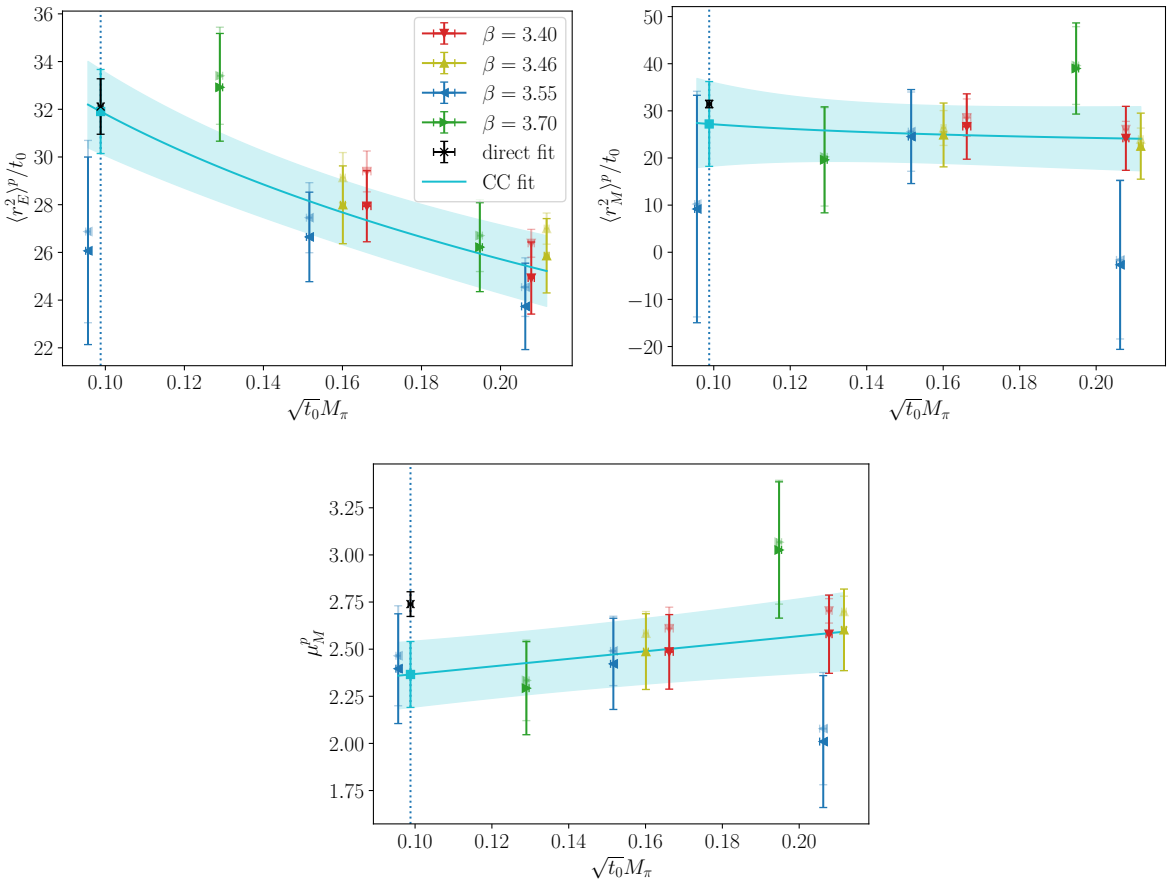


Figure 6.30: Electromagnetic radii and magnetic moment of the proton as a function of the pion mass. The faint symbols represent our original lattice data obtained from a z-expansion with $Q^2_{\text{cut}} = 0.6$ GeV^2, while the opaque ones have been corrected for the continuum limit. The cyan lines and bands depict an extrapolation fit (CC fit) according to eqs. (6.73) to (6.75). Its results at the physical point are shown as cyan squares and the model-averaged results of the direct fits as black crosses (*cf.*, table 6.3 below), with a dotted vertical line at the physical pion mass (in units of $\sqrt{t_0}$) to guide the eye.

sense the two-step process, first performing z-expansion fits and subsequently extrapolating, masks the relative paucity of data points at small momentum transfer for some ensembles.

We conclude that the direct fits are superior to the analysis using a z-expansion followed by a chiral and continuum extrapolation, in particular for the description of the magnetic form factor: they are more stable against fluctuations on individual momenta or ensembles and take more information about the physical properties of the form factors into account, which helps in reducing the errors. In cases where the two-step procedure based on the z-expansion is stable and trustworthy, both methods give consistent results. Therefore, we do not find any evidence that the functional forms employed by the direct BχPT fits introduce a systematic bias. We also remark that there is no meaningful possibility of averaging the results from the direct BχPT fits with those from the z-expansion because the latter are, in particular for the magnetic quantities, much less precise and simply not competitive.

6.8 Ratio of electric and magnetic form factor

As already mentioned at the beginning of section 6.7.2, phenomenologically motivated dipole fits to unpolarized *ep*-scattering data have traditionally found a ratio of the electric and magnetic form factor of the proton which stays rather constant close to one over a large range of Q^2. Conversely, polarization experiments, which are directly sensitive to this ratio, have seen a dramatic decrease of G_E^p/G_M^p starting at $Q^2 \approx 0.5$ GeV^2 with a continuing decline up to very large Q^2 [212]. The range of Q^2 for which a significant discrepancy can be established has recently been extended by new, precise unpolarized measurements [213]. Therefore, we want to investigate in this section whether our lattice data can provide any theory insights on this question.

For this purpose, we build the ratio G_E^p/G_M^p already on the level of the (momentum-averaged) three-point functions,

$$\frac{G_E^{\text{eff},p}(Q^2;t_{\text{sep}},t)}{G_M^{\text{eff},p}(Q^2;t_{\text{sep}},t)} = \frac{1}{m_N+E_{\mathfrak{q}}}\frac{\tilde{G}_E^{\text{eff},p}(Q^2;t_{\text{sep}},t)}{\tilde{G}_M^{\text{eff},p}(Q^2;t_{\text{sep}},t)}, \tag{6.77}$$

$$\begin{aligned}\tilde{G}_E^{\text{eff},p}(Q^2;t_{\text{sep}},t) = \sum_{\tilde{\mathbf{q}}\in\mathfrak{q}}\frac{1}{6}\Bigg(&\langle C_{3,V_0}^{\Gamma_3^p}\rangle^{\text{conn},u+d}(\mathbf{0},\tilde{\mathbf{q}};t_{\text{sep}},t) + 3\,\langle C_{3,V_0}^{\Gamma_3^p}\rangle^{u-d}(\mathbf{0},\tilde{\mathbf{q}};t_{\text{sep}},t) \\ &+ \frac{2}{3}\sum_{j=1}^{3}\langle C_{3,V_0}^{\Gamma_j^p}\rangle^{\text{disc},l-s}(\mathbf{0},\tilde{\mathbf{q}};t_{\text{sep}},t)\Bigg)\Bigg/\sum_{\tilde{\mathbf{q}}\in\mathfrak{q}}1\end{aligned} \tag{6.78}$$

$$\begin{aligned}\tilde{G}_M^{\text{eff},p}(Q^2;t_{\text{sep}},t) = \sum_{\substack{\tilde{\mathbf{q}}\in\mathfrak{q}\\ \tilde{q}_1^2+\tilde{q}_2^2\neq 0}}\frac{1}{6}\Bigg[&\sum_{k,l}\epsilon_{3kl}\tilde{q}_l\Bigg(\text{Re}\,\langle C_{3,V_k}^{\Gamma_3^p}\rangle^{\text{conn},u+d}(\mathbf{0},\tilde{\mathbf{q}};t_{\text{sep}},t) \\ &+ 3\,\text{Re}\,\langle C_{3,V_k}^{\Gamma_3^p}\rangle^{u-d}(\mathbf{0},\tilde{\mathbf{q}};t_{\text{sep}},t)\Bigg)\Bigg/(\tilde{q}_1^2+\tilde{q}_2^2) \\ &+ 2\sum_{\substack{j=1\\ \sum_{k\neq j}\tilde{q}_k^2\neq 0}}^{3}\frac{\sum_{k,l}\epsilon_{jkl}\tilde{q}_l\,\text{Re}\,\langle C_{3,V_k}^{\Gamma_j^p}\rangle^{\text{disc},l-s}(\mathbf{0},\tilde{\mathbf{q}};t_{\text{sep}},t)}{\sum_{k\neq j}\tilde{q}_k^2} \\ &\Bigg/\sum_{\substack{j=1\\ \sum_{k\neq j}\tilde{q}_k^2\neq 0}}^{3}1\Bigg]\Bigg/\sum_{\substack{\tilde{\mathbf{q}}\in\mathfrak{q}\\ \tilde{q}_1^2+\tilde{q}_2^2\neq 0}}1\end{aligned} \tag{6.79}$$

so that the two-point functions otherwise entering the ratios eq. (6.43) cancel.[11] Afterwards, we apply the LO summation method with the window average directly to the quantity defined by eq. (6.77). In this way, correlated fluctuations can cancel to a greater degree, and the uncertainty from the removal of the excited-state contamination is only incurred once. The latter effect becomes particularly important for larger Q^2-values since the growing noise makes it increasingly harder to resolve excited-state effects. We emphasize that the thus obtained data points do not enter any of our fits described in section 6.7.

[11] We remark that if one were to build the ratio G_E^p/G_M^p on the level of the effective form factors extracted from eqs. (6.50) and (6.51), the two-point functions would not cancel exactly because different statistics are used for the two-point functions entering the connected and disconnected contributions, respectively, and the combination for the proton has to be built before taking the ratio G_E^p/G_M^p.

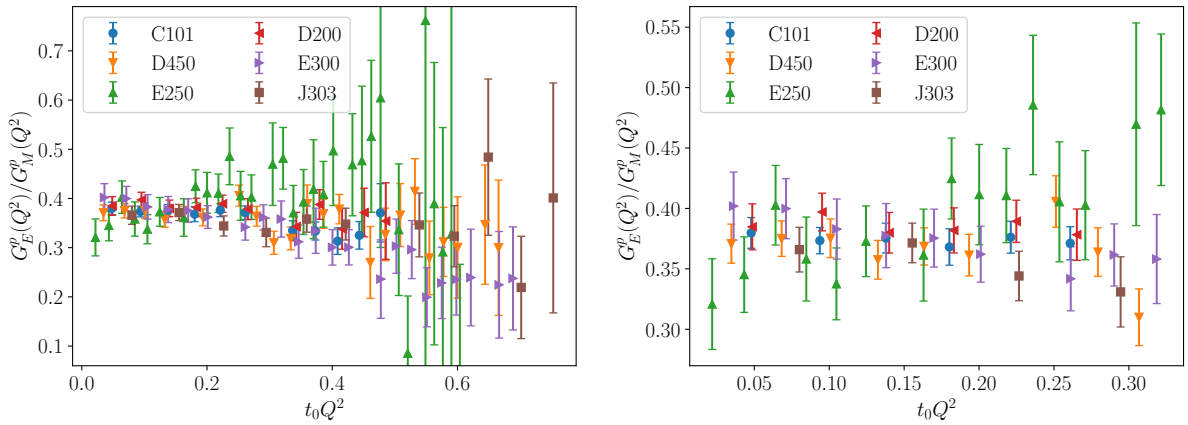


Figure 6.31: Ratio of the electric and magnetic form factor of the proton obtained as described in the text as a function of Q^2 for all ensembles entering the final BχPT fits. The right panel is zoomed in from the left one, but shows the same data.

Our lattice data for G_E^p/G_M^p obtained as explained above are plotted in fig. 6.31 as a function of Q^2 for all ensembles entering the final BχPT fits. As already noticed in Ref. [60] on the level of the isovector combination, we observe that this ratio is rather flat on most ensembles. On our finest boxes E300 and J303, the data show signs of a light downward slope, whereas on our near-physical pion mass ensemble E250, they exhibit a light upward slope. Both are, however, at the edge of significance, so that no definite conclusion on this question is possible. With respect to E250, we also remind the reader of our discussion of the relatively pronounced statistical fluctuations on this ensemble in section 6.7.1.

Moreover, the range of momenta accessible with a decent error in our lattice calculation is very limited in comparison to those considered in the context of scattering experiments. In particular, the decrease of the form factor ratio observed in polarization experiments in the range $Q^2 \leq 0.6$ GeV^2 (which is what is shown in the right panel of fig. 6.31) is only about 3 % to 5 % [212]. Such a small change could easily be masked by our statistical errors, as is obvious from fig. 6.31. We furthermore remark that a similar dip at low Q^2 is also observed in many unpolarized *ep*-scattering experiments [212, 213]. Consequently, our lattice calculation can at the current level of accuracy not contribute to the discussion of the discrepancy in G_E^p/G_M^p at larger Q^2.

6.9 The Zemach and Friar radii

In the previous section, we have seen that we cannot make a relevant contribution to the analysis of large-Q^2 observables. Therefore, we will focus on low-Q^2 observables like the radii for the remainder of this thesis, for which we can achieve a much higher precision and thus anticipate to have a significant impact.

6.9.1 Relevance of the Zemach and Friar radii

The most accurate experimental determination of the proton's electric radius is derived from the measurement of the Lamb shift in muonic hydrogen spectroscopy [16, 17]. This result exhibits a large tension with some *ep*-scattering experiments [14, 25], which is known as the proton radius puzzle we have already referred to in the introduction.

To infer the electric radius from the observed Lamb shift, higher-order nuclear-structure contributions need to be subtracted. The leading contribution is the two-photon exchange [214], the dominant, elastic part of which depends on the third Zemach moment of the proton. The general definition of the third Zemach moment of either the proton or the neutron is given by [215–217],

$$\langle r_E^3 \rangle_{(2)} = \frac{48}{\pi} \int_0^\infty \frac{dQ}{Q^4} \left[(G_E(Q^2))^2 - (G_E(0))^2 - \frac{\partial (G_E(Q^2))^2}{\partial Q^2} \bigg|_{Q^2=0} Q^2 \right]. \tag{6.80}$$

For the proton, this evaluates to

$$\langle r_E^3 \rangle_{(2)}^p = \frac{24}{\pi} \int_0^\infty \frac{dQ^2}{(Q^2)^{5/2}} \left[(G_E^p(Q^2))^2 - 1 + \frac{1}{3} \langle r_E^2 \rangle^p Q^2 \right], \tag{6.81}$$

and for the neutron to

$$\langle r_E^3 \rangle_{(2)}^n = \frac{24}{\pi} \int_0^\infty \frac{dQ^2}{(Q^2)^{5/2}} (G_E^n(Q^2))^2. \tag{6.82}$$

The radius associated with the third Zemach moment is known as the Friar radius,

$$r_F = \sqrt[3]{\langle r_E^3 \rangle_{(2)}}. \tag{6.83}$$

Historically [218], a very large Friar radius of the proton has been suggested as a possible solution to the proton radius puzzle. For this purpose, however, the Friar radius would need to be so large that the expansion in radii would break down [219, 220].

While the situation regarding the proton's electric radius is seemingly awaiting its final resolution, one also finds discrepant results for its magnetic radius, as already mentioned in the introduction. Specifically, there is a tension of $2.7\,\sigma$ between the value extracted from a z-expansion analysis of the A1 ep-scattering data alone and the estimate from the corresponding analysis applied to the remaining world data [40]. Dispersive analyses [30, 32, 34] arrive at magnetic proton radii significantly larger than the A1-data analyses [14, 40], but smaller than z-expansion on the remaining world data [40]. Therefore, it appears that the puzzle regarding the proton's electromagnetic size persists, but is shifting from the electric to the magnetic properties.

Underlining the importance of the magnetic properties of the proton, several experiments are underway to determine these from spectroscopy on (muonic) hydrogen [221–224]. This can be achieved by measuring, in addition to the Lamb shift, the hyperfine splitting (HFS) in either electronic or muonic hydrogen, which is caused by the magnetic spin-spin interaction between the nucleus and the orbiting lepton. The influence of the electromagnetic structure of the nucleus on the HFS is particularly pronounced for the S states, since the S-state wave function has a large overlap with the nucleus.

The leading-order proton-structure contribution to the S-state HFS of hydrogen depends on the Zemach radius of the proton. The general definition of the Zemach radius of either the proton or the neutron is given by [216, 225],

$$r_Z = -\frac{4}{\pi} \int_0^\infty \frac{dQ}{Q^2} \left[\frac{G_E(Q^2) G_M(Q^2)}{\mu_M} - \frac{G_E(0) G_M(0)}{\mu_M} \right]. \tag{6.84}$$

For the proton, this evaluates to

$$r_Z^p = -\frac{2}{\pi} \int_0^\infty \frac{dQ^2}{(Q^2)^{3/2}} \left[\frac{G_E^p(Q^2) G_M^p(Q^2)}{\mu_M^p} - 1 \right], \tag{6.85}$$

and for the neutron to

$$r_Z^n = -\frac{2}{\pi}\int_0^\infty \frac{dQ^2}{(Q^2)^{3/2}}\frac{G_E^n(Q^2)G_M^n(Q^2)}{\mu_M^n}. \tag{6.86}$$

A firm theoretical prediction of the Zemach radius of the proton is of vital importance for the success of the atomic spectroscopy experiments planning to measure the ground-state ($1S$) HFS in muonic hydrogen with ppm precision [221–224], from which the Zemach radius could be extracted with sub-percent uncertainty: On the one hand, it helps in narrowing down the frequency search range. On the other hand, a lattice calculation of the proton's Zemach radius would represent a first-principles prediction which will be tested by high-precision experiments and would thus allow for a crucial consistency check. It is worth noting that the interpretation of the experimental HFS results from either electronic or muonic hydrogen alone relies on theoretical input for the proton-polarizability effect, where a discrepancy has emerged between data-driven approaches and BChPT [226]. Once high-precision results for the HFS in muonic hydrogen are available, these can be combined with the existing precise results in electronic hydrogen. This will make it possible to disentangle the Zemach and polarizability contributions and compare them both to theory [41].

6.9.2 Lattice calculation: extrapolation of the form factors and integration

In the following, we will explain in detail how we use our lattice determination of the electromagnetic form factors of the proton and neutron to compute their Zemach and Friar radii. Due to their very limited range of validity in Q^2, the BχPT fits cannot be employed directly to evaluate the full integrals in eqs. (6.81), (6.82), (6.85) and (6.86). Hence, an extrapolation of the BχPT fits beyond their range of applicability is required if they are to be used to parametrize the form factors. For each model variation, we evaluate the BχPT formula for $G_E^{p,n}$ and $G_M^{p,n}$, using the low-energy constants as determined from the corresponding fit, at the physical point and at twenty evenly spaced points in $Q^2 \in [0, Q^2_{\text{cut}}]$. Here, Q^2_{cut} is the cut in the momentum transfer corresponding to the respective variation of the BχPT fit.

In the next step, we fit a model which obeys the large-Q^2 constraints on the form factors from perturbation theory [227] to these data points and their error estimates. We note that the data points exhibit an extremely high artificial correlation due to the way we generate them. Taking these correlations into account when adjusting the extrapolation model would thus not be meaningful, and also technically challenging because the resulting covariance matrices are extremely badly conditioned.

To describe the Q^2-dependence, we use the model-independent z-expansion [209] explained in section 6.7.2 (*cf.*, eqs. (6.69) and (6.70)). Here, we employ $\tau_{\text{cut}} = 4M^2_{\pi,\text{phys}}$ and again $\tau_0 = 0$. We truncate the z-expansion beyond $n = 9$ and incorporate the four sum rules from Ref. [40] for each form factor, which ensure the correct asymptotic behavior of the latter for large Q^2 (namely that G_E and G_M fall off $\propto Q^{-4}$). The normalization of the electric form factor is enforced by fixing $a_0^p = 1$ and $a_0^n = 0$, respectively. For the determination of the Zemach radius, we fit G_E and G_M simultaneously, similar to the crosscheck of our analysis in section 6.7.2, so that we have eleven independent fit parameters altogether. For the third Zemach moment, on the other hand, only the electric form factor is required, so that we fit only G_E and have five independent fit parameters here. The extrapolation fits are performed for the proton and neutron independently. Using more than twenty data points for each form factor or a higher degree of the z-expansion does not increase the overlap between the original BχPT fit and the extrapolation any further.

For the numerical integration of eqs. (6.81), (6.82), (6.85) and (6.86), we smoothly replace the BχPT parametrization of the form factors by the z-expansion-based extrapolation in a narrow

window around Q^2_{cut}. Concretely, we use the following estimate for the form factor term,

$$F(Q^2) = \frac{1}{2}\left[1 - \tanh\left(\frac{Q^2 - Q^2_{\text{cut}}}{\Delta Q^2_w}\right)\right] F^\chi(Q^2) + \frac{1}{2}\left[1 + \tanh\left(\frac{Q^2 - Q^2_{\text{cut}}}{\Delta Q^2_w}\right)\right] F^z(Q^2), \quad (6.87)$$

where $F(Q^2) \equiv G_E(Q^2)G_M(Q^2)/\mu_M$ for the Zemach radius and $F(Q^2) \equiv G^2_E(Q^2)$ for the third Zemach moment, respectively. In eq. (6.87), $F^\chi(Q^2)$ represents our fit to BχPT, while $F^z(Q^2)$ denotes the z-expansion parametrization of the form factors. For the width of the window in which we switch between the two parametrizations, we choose $\Delta Q^2_w = 0.0537 t_0^{-1} \approx 0.1$ GeV^2.

For a consistent calculation of the third Zemach moment of the proton, the replacement according to eq. (6.87) has to be applied to all terms in eq. (6.81). Thus, instead of $\partial (G^p_E(Q^2))^2/\partial Q^2|_{Q^2=0}$, we have to use

$$\begin{aligned}\left.\frac{\partial F(Q^2)}{\partial Q^2}\right|_{Q^2=0} = &\left[1 + \tanh\left(\frac{Q^2_{\text{cut}}}{\Delta Q^2_w}\right)\right] \left.\frac{\partial G^{p,\chi}_E(Q^2)}{\partial Q^2}\right|_{Q^2=0} \\ &+ \left[1 - \tanh\left(\frac{Q^2_{\text{cut}}}{\Delta Q^2_w}\right)\right] \left.\frac{\partial G^{p,z}_E(Q^2)}{\partial Q^2}\right|_{Q^2=0}. \end{aligned} \quad (6.88)$$

Here, we have made use of the fact that $G^{p,\chi}_E(0) = G^{p,z}_E(0) = 1$. Effectively, eq. (6.88) amounts to applying the replacement according to eq. (6.87) also to the value of $\langle r^2_E \rangle^p$ in eq. (6.81).

The cancellation between the different terms of eq. (6.81) at small Q^2 does not occur at the required numerical accuracy on all our bootstrap samples. To facilitate the numerical integration, we therefore regulate the small-Q^2 contribution to the integral for the proton by replacing $t_0 Q^2 \to t_0 Q^2 + \epsilon$ in the denominator. This regularization is only applied to the integral from $Q^2 = 0$ to $Q^2 = 0.00537 t_0^{-1} \approx 0.01$ GeV^2.

The value of ϵ can be tuned to balance the systematic error, which we define as the difference to the integration with $\epsilon = 0$, with the numerical error of the integration. The relative errors as a function of ϵ are displayed in fig. 6.32 for the variation of the BχPT fits with $M_{\pi,\text{cut}} = 0.23$ GeV, $Q^2_{\text{cut}} = 0.6$ GeV^2 and corrections for both types of lattice artefacts according to the additive model of eqs. (6.62) and (6.63). The calculation of the systematic error is possible because the integration works also without a regulator for the central value (in contrast to some of the bootstrap samples). For the numerical error, on the other hand, we have taken the maximum over all bootstrap samples. One can see from fig. 6.32 that $\epsilon = 1 \times 10^{-7}$ offers a good compromise between the systematic and numerical errors, thus minimizing the quadratic sum of the two. We remark that for this value, which is the one we will employ in the following, the systematic error amounts to less than 6 % of the statistical error of $\langle r^3_E \rangle^p_{(2)}$ and can hence safely be neglected.

The two parametrizations and their weighted average according to eq. (6.87) are illustrated in fig. 6.33 for the case of the Zemach radius. While the BχPT formula is clearly not reliable for $Q^2 \gtrsim 1.7$ GeV$^2 \approx 0.9 t_0^{-1}$, the z-expansion behaves well for arbitrarily large momenta, which is due to the sum rules [40] we have included. In the region where we adjust the z-expansion to the BχPT parametrization ($0 < Q^2 \leq 0.6$ GeV^2 for the case shown in fig. 6.33), however, the two curves overlap so closely that they are indistinguishable by eye. The blue curve, which is the one we use for the integration, smoothly switches from the orange (BχPT) curve to the green (z-expansion) one in a tight window around $Q^2_{\text{cut}} = 0.6$ GeV$^2 = 0.322 t_0^{-1}$.

The results for the Zemach radii and third Zemach moments of the proton and neutron obtained from all variations of the BχPT fits are collected in the appendix of Ref. [7].

Finally, we note that the major advantage of our approach based on the BχPT fits over an integration of the form factors on each ensemble is that the Zemach and Friar radii can be computed directly at the physical point, so that an extrapolation of results for the radii to the physical point, which would entail further significant systematic uncertainties, is not required.

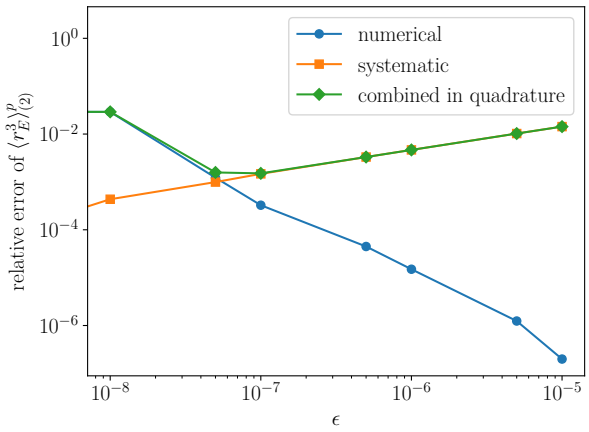


Figure 6.32: Relative numerical (integration) and systematic errors of the third Zemach moment of the proton as a function of the regulator ϵ introduced in the text. The plot has been obtained using the variation of the BχPT fits with $M_{\pi,\text{cut}} = 0.23$ GeV, $Q^2_{\text{cut}} = 0.6$ GeV^2 and corrections for both types of lattice artefacts according to the additive model of eqs. (6.62) and (6.63). For the further analysis, we use $\epsilon = 1 \times 10^{-7}$.

6.9.3 Contribution of the extrapolated form factors

Replacing the BχPT parametrization smoothly with a constant zero instead of the z-expansion-based extrapolation [*i.e.*, setting $F^z(Q^2) \equiv 0$ in eq. (6.87)] allows one to estimate the contribution of the form factors at $Q^2 > Q^2_{\text{cut}}$ to the resulting Zemach radius and third Zemach moment, respectively. In this case, $F^z(0)$ is not equal to $F^\chi(0)$ for the proton any more, so that the formulae for the Zemach radius and third Zemach moment need to be modified accordingly. For the Zemach radius of the proton, the expression in square brackets in eq. (6.85) has to be replaced by

$$F(Q^2) - F(0) = \frac{1}{2}\left[1 - \tanh\left(\frac{Q^2 - Q^2_{\text{cut}}}{\Delta Q^2_w}\right)\right] F^\chi(Q^2) - \frac{1}{2}\left[1 + \tanh\left(\frac{Q^2_{\text{cut}}}{\Delta Q^2_w}\right)\right]. \tag{6.89}$$

For the third Zemach moment of the proton, the appropriate replacement for the term in square brackets in eq. (6.81) is

$$\begin{aligned} F(Q^2) - F(0) - \left.\frac{\partial F(Q^2)}{\partial Q^2}\right|_{Q^2=0} Q^2 = {} & \frac{1}{2}\left[1 - \tanh\left(\frac{Q^2 - Q^2_{\text{cut}}}{\Delta Q^2_w}\right)\right] F^\chi(Q^2) \\ & - \frac{1}{2}\left[1 + \tanh\left(\frac{Q^2_{\text{cut}}}{\Delta Q^2_w}\right)\right] \\ & + \frac{1}{6}\left[1 + \tanh\left(\frac{Q^2_{\text{cut}}}{\Delta Q^2_w}\right)\right] \langle r_E^2 \rangle^{p,\chi} Q^2 \\ & + \frac{1}{2\Delta Q^2_w \cosh^2\left(Q^2_{\text{cut}}/\Delta Q^2_w\right)} Q^2. \end{aligned} \tag{6.90}$$

Again, $F(Q^2) \equiv G_E^p(Q^2) G_M^p(Q^2)/\mu_M$ in eq. (6.89) and $F(Q^2) \equiv (G_E^p(Q^2))^2$ in eq. (6.90), respectively. We stress that the replacements according to eqs. (6.89) and (6.90) are only correct and relevant if one sets the extrapolated form factors to zero, which we only do for the purpose of determining their contribution to the actual results.

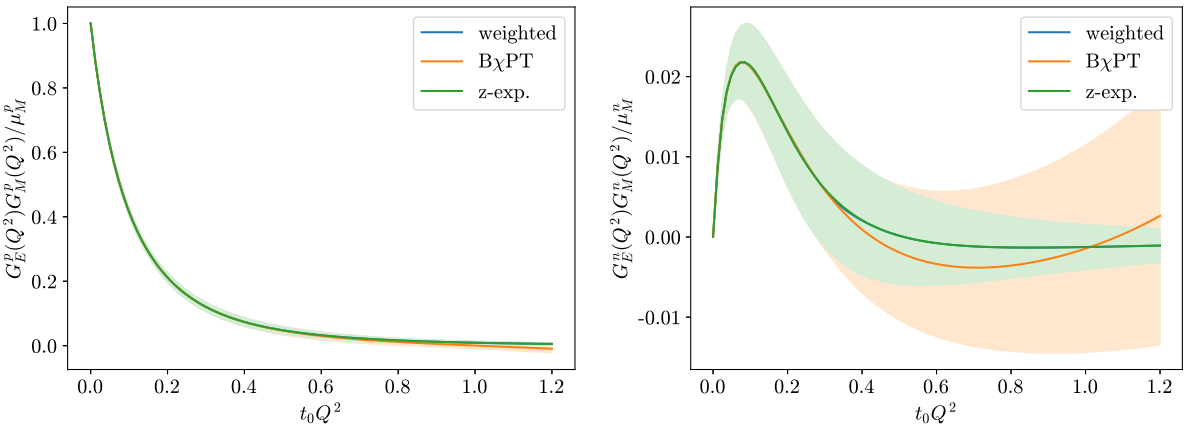


Figure 6.33: Product of the electric and normalized magnetic form factors of the proton (left panel) and neutron (right panel) at the physical point evaluated with different parametrizations. The orange curve shows the variation of the BχPT fits to our lattice data with $M_{\pi,\mathrm{cut}} = 0.23$ GeV, $Q^2_{\mathrm{cut}} = 0.6$ GeV$^2 \approx 0.322 t_0^{-1}$ and corrections for both types of lattice artefacts according to the additive model of eqs. (6.62) and (6.63). The green curve depicts the z-expansion-based extrapolation and the blue curve the weighted average of the two according to eq. (6.87).

For $Q^2_{\mathrm{cut}} = 0.6$ GeV^2 (our largest, *i.e.*, least stringent, value for the cut), we find that the relative difference of the thus obtained value for r_Z^p to the actual result using the corresponding variation of the BχPT fits is less than 0.9 %. In other words, the form factor term at $Q^2 > 0.6$ GeV^2 contributes less than 0.9 % to the Zemach radius of the proton. For the third Zemach moment, the denominator in the integrand suppresses the large-Q^2 contribution to the integral even more strongly than for the Zemach radius. Accordingly, we find a corresponding relative contribution of less than 0.3 % to the third Zemach moment of the proton.

Due to this smallness of the contribution of the extrapolated form factors, the precise form of the chosen model for the extrapolation only has a marginal influence on the resulting values for the Zemach radius and third Zemach moment. For example, if we replace the z-expansion by a dipole *ansatz* according to eq. (6.67) (which also fulfills the constraints from Ref. [227]), we find that the Zemach radius of the proton derived from any of our fit variations changes by at most ~ 20 % of the entire systematic error quoted in table 6.3 below. Thus, adding the variation in r_Z^p due to the extrapolation model quadratically to the systematic uncertainty in table 6.3 would not change the latter significantly.

6.10 Model average and final results

In sections 6.6 and 6.7 above, we have provided a comprehensive explanation of our reasons for favoring the direct fits over the two-step procedure based on the z-expansion, and for using the LO summation method with the window average for ground-state identification. We will hence restrict our presentation of the final results to our preferred methods. Within the direct-fit approach, we have no strong *a priori* preference for one specific setup, and thus determine our final results and total errors from averages over different fit models and kinematic cuts.

For this purpose, we employ weights derived from the Akaike Information Criterion (AIC) [185, 186]. In the following, we will sketch the construction of the AIC and its extensions which are relevant to our analysis. Consider a size-N sample of data drawn from an unknown true

probability distribution P_T, $\mathbf{y} = (y_1, \ldots, y_N)$. Note that in general, each y_i is a dimension-d vector. Our aim is to study parametric models M_μ depending on parameters a_j, $j = 1, \ldots, k$ with corresponding probability densities $P(\mathbf{y}|M_\mu, \mathbf{a})$, and to find associated weights w_μ to use for model averaging.

The adequacy of a model M_μ to approximate the true distribution can be measured by their Kullback-Leibler (KL) divergence [228],

$$\mathrm{KL}(M_\mu, \mathbf{a}) = \int dz\, P_T(z) \ln\left(\frac{P_T(z)}{P(z|M_\mu, \mathbf{a})}\right) = \int dz\, P_T(z)[\ln P_T(z) - \ln P(z|M_\mu, \mathbf{a})]. \quad (6.91)$$

Here, z denotes future observations drawn from $P_T(z)$, so that eq. (6.91) is an expectation value with respect to the true distribution. In words, the KL divergence measures the information loss in the estimation of $P_T(z)$ by the model distribution $P(z|M_\mu, \mathbf{a})$. Since we are dealing with parametric models, we need to fix the parameter values before we can evaluate and minimize eq. (6.91). In a frequentist context, one commonly uses the maximum-likelihood estimator $\mathbf{a}^*_{\mathrm{MLE}}$, and in a Bayesian context the posterior mode $\mathbf{a}^*_{\mathrm{PM}}$ [229]. We note that both of these estimators are based on maximizing a likelihood function calculated from the observed data $\mathbf{y}$ rather than from the true distribution P_T (which is inaccessible in practice). Simply plugging such an estimator into eq. (6.91) yields the plug-in KL divergence,

$$\mathrm{KL}_{\text{plug-in}}(M_\mu) = \int dz\, P_T(z)[\ln P_T(z) - \ln P(z|M_\mu, \mathbf{a}^*)], \quad (6.92)$$

which no longer depends on $\mathbf{a}$. The first term in eq. (6.92) does not depend on the model but only on the true distribution. If the data is fixed for all candidate models, this is a constant, and minimizing $\mathrm{KL}_{\text{plug-in}}(M_\mu)$ is equivalent to minimizing $-2\int dz\, P_T(z) \ln P(z|M_\mu, \mathbf{a}^*)$, where we have included a conventional factor of 2.

Since the true distribution is unknown, we need to estimate the expectation value with respect to it from the observed data $\mathbf{y}$,

$$-2\int dz\, P_T(z) \ln P(z|M_\mu, \mathbf{a}^*) \approx -\frac{2}{N}\sum_{i=1}^{N} \ln P(y_i|M_\mu, \mathbf{a}^*) = -\frac{2}{N} \ln P(\mathbf{y}|M_\mu, \mathbf{a}^*). \quad (6.93)$$

Because the data $\mathbf{y}$ is used to estimate both the optimal parameters $\mathbf{a}^*$ and the "outer" expectation value, eq. (6.93) is not unbiased. If one assumes that the true distribution P_T belongs to the family of candidate distributions under consideration, one can derive that the bias of eq. (6.93) can be estimated by $-2k/N$, where k is the number of parameters (*i.e.*, the dimension of the parameter vector $\mathbf{a}$) [230, 231]. Multiplying by a conventional factor of N, one obtains (in the frequentist case) the AIC [185, 186],

$$\mathrm{AIC}_\mu = -2 \ln P(\mathbf{y}|M_\mu, \mathbf{a}^*_{\mathrm{MLE}}) + 2k. \quad (6.94)$$

Even if the requirement that the true distribution is part of the collection of candidate models is not always fulfilled in practice, it has been shown in Ref. [232] that the AIC is asymptotically optimal in the sense that it asymptotically selects the candidate model which minimizes the mean-squared error of prediction.

In the context of least-squares fits (*cf.*, section 6.2.1), one assumes that the likelihood function is given by a Gaussian,

$$P(\mathbf{y}|M_\mu, \mathbf{a}) \propto \prod_{i=1}^{N} \exp\left[-\frac{1}{2}\tilde{\chi}^2_{i,\mu}(\mathbf{a})\right], \quad (6.95)$$

with the residuals $\tilde{\chi}^2_{i,\mu}(\mathbf{a})$ for the model M_μ on the data sample i,

$$\tilde{\chi}^2_{i,\mu}(\mathbf{a}) = \frac{1}{N}[y_i - f_\mu(\mathbf{a})]^T C^{-1}[y_i - f_\mu(\mathbf{a})]. \tag{6.96}$$

Here, $f_\mu(\mathbf{a})$ is the model function corresponding to the model M_μ and C denotes the data covariance matrix as usual (*i.e.*, the one of the estimators, *cf.*, eq. (6.9)). The mean-based residuals χ^2_μ defined as in eq. (6.29) can be written as the sum over the sample-based residuals, $\sum_{i=1}^N \tilde{\chi}^2_{i,\mu}$, plus a model-independent constant which can be ignored if the data is kept fixed between models [229]. Thus, the AIC takes the form [229],

$$\mathrm{AIC}_\mu = \chi^2_\mu(\mathbf{a}^*_{\mathrm{MLE}}) + 2k, \tag{6.97}$$

where the maximum-likelihood estimator for the parameters, $\mathbf{a}^*_{\mathrm{MLE}}$, is obtained by minimizing $\chi^2_\mu(\mathbf{a})$ (*cf.*, eq. (6.29)).

In the presence of Bayesian priors, the maximum-likelihood estimator is replaced by the posterior mode $\mathbf{a}^*_{\mathrm{PM}}$ which minimizes the augmented residuals $\chi^2_{\mathrm{aug}}(\mathbf{a})$ (*cf.*, section 6.2.3). It has been shown in Ref. [229] that minimizing the KL divergence in the Bayesian case yields the Bayesian Akaike Information Criterion (BAIC),

$$\mathrm{BAIC}_\mu = -2\ln P(\mathbf{y}|M_\mu, \mathbf{a}^*_{\mathrm{PM}}) + 2k = \chi^2_\mu(\mathbf{a}^*_{\mathrm{PM}}) + 2k, \tag{6.98}$$

where the latter equality holds (up to a model-independent constant) in the case of least-squares fits. This means that what enters the BAIC are the unaugmented residuals (*i.e.*, without the contribution of the priors), but evaluated at the posterior mode (which is obtained by minimizing the augmented residuals).

The applicability of the AIC and BAIC in their original forms is called into question when only subsets of data are fitted for some models as a result of applying data cuts, *i.e.*, ignoring d_c of the d dimensions of the data vectors. This is because the derivation of these criteria assumes that all models describe the same fixed set of data. To circumvent this restriction, Refs. [229, 233] have introduced a perfect-model construction. The key idea is that cutting d_c data points from a fit is equivalent to fitting these points to a perfect model with zero degrees of freedom. Such a model which passes through the cut data points exactly obviously requires d_c additional fit parameters. An explicit construction of the perfect model as well as a comprehensive discussion of the influence of correlations between the cut and the kept parts of the data can be found in Ref. [229]. In general, however, an explicit construction is not required in practice; it suffices to replace $k \to k + d_c$ in the AIC and BAIC,

$$\mathrm{AIC}_{P,\mu} = \chi^2_\mu(\mathbf{a}^*_{\mathrm{MLE}}) + 2k + 2d_c, \tag{6.99}$$

$$\mathrm{BAIC}_{P,\mu} = \chi^2_\mu(\mathbf{a}^*_{\mathrm{PM}}) + 2k + 2d_c. \tag{6.100}$$

From a practical point of view, the (B)AIC takes the goodness of fit into account, while at the same time penalizing a reduction of the degrees of freedom that may result either from the introduction of further fit parameters or from cutting away data points.

In our direct fits, we both use Bayesian priors for some fit parameters and apply kinematic cuts. Consequently, we employ the BAIC_P defined by eq. (6.100). For the proton and neutron observables, which are derived from two separate fits to the isovector and isoscalar form factors, the BAIC_P is obtained as the sum of the BAIC_P values of both fits (which would be the BAIC_P resulting from a combined fit with the cross-correlations between the two channels set to zero[12]).

[12]Estimating these correlations is not feasible because the size of the resulting covariance matrices would be much larger than the number of available configurations on some ensembles.

For the weighting of the different models, one can use [229, 234, 235],

$$w_\mu^{\mathrm{BAIC}_P} = \frac{e^{-\mathrm{BAIC}_{P,\mu}/2}}{\sum_\nu e^{-\mathrm{BAIC}_{P,\nu}/2}}. \tag{6.101}$$

When computing these weights for our set of models, it turns out that the BAIC_P strongly prefers the fits with the least stringent cut in Q^2. This is due to the relatively large number of data points which is cut away by lowering Q^2_{cut} in a fit across several ensembles. The effect is enhanced by our two most chiral ensembles E250 and E300 which feature a comparatively large density of Q^2-points. At a pion-mass cut of 0.23 GeV, for example, a reduction of Q^2_{cut} from 0.6 GeV^2 to 0.5 GeV^2 leads to an increase in d_c by 14. If the fit quality stayed the same, $\chi^2(\mathbf{a}^*_{\mathrm{PM}})$ would shrink by 14, so that the BAIC_P would in total rise by 14. According to eq. (6.101), this would imply a relative weight factor of $e^{-14/2} = e^{-7} \approx 9.1 \times 10^{-4}$, which is close to what we actually observe.

Since the radii and the magnetic moment are defined in terms of the low-Q^2 behavior of the form factors, a stricter cut in Q^2 is theoretically better motivated for an extraction of these quantities. Hence, we employ eq. (6.101) for each value of Q^2_{cut} at a time to weight the remaining variations, *i.e.*, the pion-mass cut and the modelling of lattice artefacts. Using these separately normalized BAIC_P weights, we finally apply a flat weight function to the estimates originating from the different Q^2_{cut}. This prescription, which we dub $\overline{\mathrm{BAIC}}$, ensures that the stricter cuts in Q^2, and thus our low-momentum data, have a strong influence on our final results.

For the estimation of the statistical and systematic uncertainties of our model averages, we adopt a bootstrapped variant the procedure from Ref. [10], which we briefly sketch in the following. To that end, one treats the model-averaged quantity as a random variable with a cumulative distribution function (CDF) adding up from the weighted CDFs of the individual models,

$$\mathrm{cdf}^x(y) = \sum_{\mu=1}^{N_M} \frac{w_\mu^{\overline{\mathrm{BAIC}}}}{N_B} \sum_{n=1}^{N_B} \Theta(y - x_{\mu,n}). \tag{6.102}$$

Here, the outer sum runs over our $N_M = 4 \times 2 \times 7 = 56$ models with the associated weights $w_\mu^{\overline{\mathrm{BAIC}}}$ computed as explained above. The inner sum runs over the $N_B = 10\,000$ bootstrap samples obtained from our resampled analysis (*cf.*, sections 6.1.5 and 6.7.1). Θ denotes the Heaviside step function and $x_{\mu,n}$ the estimate for the observable x on the n-th sample and using the μ-th model. Due to the large number of bootstrap samples N_B, the distribution in eq. (6.102) is effectively smoothed in spite of being a sum of step functions. The final value and the total error are easily read off from this distribution as the median and the quantiles which would correspond to the central $1\,\sigma$ of an effective Gaussian distribution, respectively.

In order to isolate the statistical and systematic errors, one can scale the (squared) width of the bootstrap distributions entering eq. (6.102) by a factor of λ. Under the assumption that such a rescaling of the errors of the individual model results only affects the statistical, but not the systematic error, one can separate these two contributions as demonstrated in Ref. [10],

$$\sigma^2_{\mathrm{stat}} = \frac{\sigma^2_{\mathrm{scaled}} - \sigma^2_{\mathrm{orig}}}{\lambda - 1}, \qquad \sigma^2_{\mathrm{syst}} = \frac{\lambda\sigma^2_{\mathrm{orig}} - \sigma^2_{\mathrm{scaled}}}{\lambda - 1}. \tag{6.103}$$

Here, σ_{scaled} denotes the total error obtained from the rescaled analysis and σ_{orig} the one from the original analysis without rescaling. We use $\lambda = 2$ as in Ref. [10], but we remark that the results of this method are essentially independent of the choice of λ for our data as long as λ does not come too close to 1. As an example of this, we display the statistical and systematic

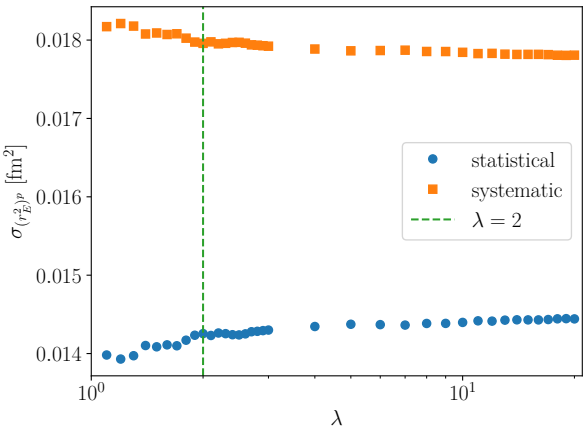


Figure 6.34: Statistical and systematic errors of the model-averaged result for $\langle r_E^2\rangle^p$ as a function of the rescaling factor λ. The total error, which combines the two in quadrature, is independent of λ by construction. We use $\lambda = 2$ as indicated by the dashed green line.

errors of $\langle r_E^2\rangle^p$ in fig. 6.34 as a function of λ. They indeed show only a negligibly small trend as λ is varied.

The collection of results for the electromagnetic radii and the magnetic moments of the proton and neutron together with the CDF obtained as explained above is displayed in fig. 6.35. One can see that approximately the expected fraction of results lie within the central 68 % quantiles of the averaged distribution. Moreover, the symmetrized errors as shown by the gray bands agree well with the (generally non-symmetric) quantiles of the distributions, which are indicated by the dashed lines. Since this statement holds for all observables under study, we quote the symmetrized errors together with our final results which are collected in table 6.3. We find that we can obtain the magnetic radii of the proton and neutron to a precision very similar to their respective electric radii.

Table 6.3: Final results for the radii and magnetic moments. In each case, the first error is statistical and the second one systematic, respectively.

Observable	Isovector	Isoscalar	Proton	Neutron
$\langle r_E^2\rangle$ [fm^2]	0.785(22)(26)	0.554(18)(13)	0.672(14)(18)	−0.115(13)(7)
$\langle r_M^2\rangle$ [fm^2]	0.663(11)(8)	0.657(30)(31)	0.658(12)(8)	0.667(11)(16)
μ_M	4.62(10)(7)	2.47(11)(10)	2.739(63)(18)	−1.893(39)(58)
r_Z [fm]	–	–	1.013(10)(12)	−0.0411(56)(40)
$\langle r_E^3\rangle_{(2)}$ [fm^3]	–	–	2.200(60)(71)	0.0078(20)(12)

The individual contributions of every variation to the total systematic error can be estimated following the procedure outlined in Refs. [10, 60]. To that end, one builds the CDF according to eq. (6.102) not over all variations but rather first iterates over a particular feature, *e.g.*, the momentum cut, and builds the CDF for every realization of that feature separately. One then constructs a secondary CDF as the weighted sum of these individual CDFs, where the weight of each realization is now given by the sum of the weights of all such fits. Next, the rescaling analysis as explained above is applied to this secondary CDF, which yields the corresponding variation-specific systematic error. Repeating this scheme for all variations (momentum cut,

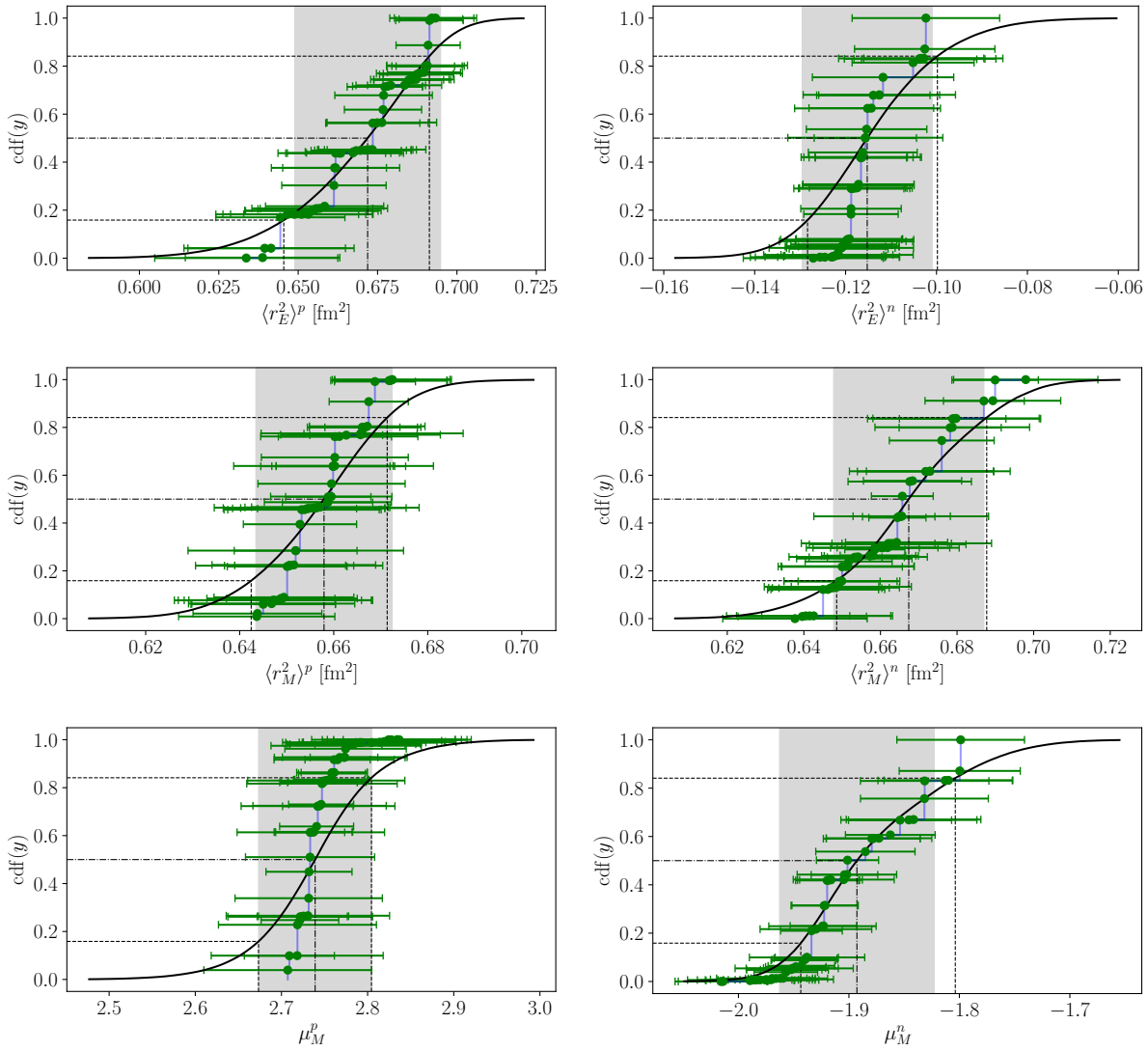


Figure 6.35: Cumulative distribution function of the electromagnetic radii and the magnetic moments of the proton (left) and neutron (right) for all fitted models. The green points depict the central values and errors of the individual fit results. The thick black line shows the weighted CDF according to eq. (6.102). For comparison, we also include a CDF based on the central values only, *i.e.*, $\mathrm{cdf}^x(y) = \sum_{\mu=1}^{N_M} w_\mu^{\overline{\mathrm{BAIC}}} \Theta(y - x_\mu)$, which is displayed by the light blue line. The dashed-dotted and dashed lines indicate the median and the central 68 % quantiles, respectively. The gray bands, on the other hand, depict the symmetrized errors quoted in table 6.3.

pion-mass cut and modelling of lattice artefacts), we obtain a complete systematic-error budget (*cf.*, table 6.4).

Table 6.4: Systematic-error budget for the radii and magnetic moments

Observable	Error	Isovector	Isoscalar	Proton	Neutron
$\langle r_E^2 \rangle$ [fm^2]	Q^2_{cut}	0.0137	0.0054	0.0102	0.0042
$\langle r_E^2 \rangle$ [fm^2]	$M_{\pi,\text{cut}}$	0.0025	0.0000	0.0020	0.0031
$\langle r_E^2 \rangle$ [fm^2]	artefacts	0.0227	0.0112	0.0157	0.0051
$\langle r_M^2 \rangle$ [fm^2]	Q^2_{cut}	0.0075	0.0084	0.0068	0.0095
$\langle r_M^2 \rangle$ [fm^2]	$M_{\pi,\text{cut}}$	0.0045	0.0116	0.0021	0.0104
$\langle r_M^2 \rangle$ [fm^2]	artefacts	0.0020	0.0294	0.0042	0.0123
μ_M	Q^2_{cut}	0.025	0.026	0.0108	0.028
μ_M	$M_{\pi,\text{cut}}$	0.028	0.040	0.0045	0.032
μ_M	artefacts	0.057	0.093	0.0081	0.054
r_Z [fm]	Q^2_{cut}	–	–	0.0075	0.0024
r_Z [fm]	$M_{\pi,\text{cut}}$	–	–	0.0013	0.0023
r_Z [fm]	artefacts	–	–	0.0101	0.0034
$\langle r_E^3 \rangle_{(2)}$ [fm^3]	Q^2_{cut}	–	–	0.0431	0.00073
$\langle r_E^3 \rangle_{(2)}$ [fm^3]	$M_{\pi,\text{cut}}$	–	–	0.0032	0.00064
$\langle r_E^3 \rangle_{(2)}$ [fm^3]	artefacts	–	–	0.0557	0.00099

We note that, due to correlations, the individual errors added in quadrature need not exactly reproduce the total systematic error given in table 6.3. In this sense, the more finely one decomposes the total error (which is the primary output of the CDF method), the more caution should be exercised in interpreting the resulting error budgets. Nevertheless, one can conclude from table 6.4 that the modelling of the lattice artefacts is the dominant source of systematic uncertainties for most observables in our analysis, while the kinematic cuts generally also play a non-negligible role. One should keep in mind, though, that excited-state contamination represents an additional, important source of systematic error in our calculation. It is, however, incorporated into the statistical error of the ground-state form factors through the window average and subsequently propagated into our final results in this way. We also remark that our statistical and total systematic errors quoted in table 6.3 are on a par, with the caveat that the statistical errors contain, in a not straightforwardly quantifiable way, the uncertainty associated with ground-state identification, which is usually regarded as a systematic error.

In fig. 6.36, we compare the numbers for the electromagnetic radii and the magnetic moments from table 6.3 to the results of two alternative averaging strategies: the BAIC$_P$ weights of eq. (6.101) applied to all variations, *i.e.*, also the cut in Q^2, or a naive (flat) average imposing a p-value cut at 1 %. For the latter, we use the average (squared) statistical uncertainty, and the variance determined from the spread of the fit results as a systematic error estimate [236],

$$\hat{x} = \frac{1}{N_M} \sum_{\mu=1}^{N_M} x_\mu, \qquad \sigma^2_{\text{stat}} = \frac{1}{N_M} \sum_{\mu=1}^{N_M} \sigma^2_\mu, \qquad \sigma^2_{\text{syst}} = \frac{1}{N_M} \sum_{\mu=1}^{N_M} (x_\mu - \hat{x})^2. \tag{6.104}$$

While this method is robust, it is also very conservative and susceptible to overestimating the true errors. The "plain" BAIC$_P$, on the other hand, drastically underestimates the systematic error for observables which display a non-negligible dependence on Q^2_{cut}, like the isovector

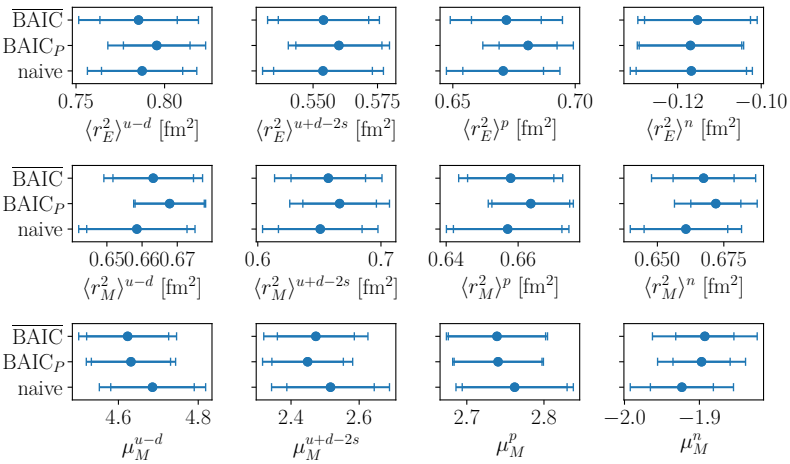


Figure 6.36: Comparison of the model-averaged results for the electromagnetic radii and the magnetic moments using different averaging methods: our preferred procedure (labelled $\overline{\text{BAIC}}$), the "plain" BAIC_P weights of eq. (6.101) applied to all variations (including Q^2_{cut}), and a naive (flat) average according to eq. (6.104).

and the proton's magnetic radius (*cf.*, also table 6.4). For these reasons, we adopt the model-averaging procedure $\overline{\text{BAIC}}$ explained above. Regarding the central values as well as the errors, its results usually lie in between those from the other two methods, all of which are compatible within errors.

6.11 Comparison to other studies

In the following, we will compare our final results to other lattice calculations, data-driven evaluations, direct experimental results and other theory approaches, where available. We will start in section 6.11.1 with the electromagnetic form factors themselves which form the basis for the extraction of all other observables we consider. Next, we will focus in section 6.11.2 on the electromagnetic radii and magnetic moments, before coming to the Zemach and Friar radii in section 6.11.3.

6.11.1 Electromagnetic form factors

In analogy to the radii and the magnetic moments, one can also average the form factors evaluated at the physical point and at particular values of Q^2 over the model variations. The results are plotted in fig. 6.37 for the proton and neutron and are directly compared to experimental data. For the proton, one can observe a moderate deviation in the slope of the electric form factor between our result and that of the A1 Collaboration [14] over the whole range of Q^2. As shown in the inset, the slope of our electric form factor at low Q^2 is much closer to that of the PRad experiment [26] than to that of Ref. [14]. The magnetic form factor, on the other hand, agrees well with that of Ref. [14]. For the neutron, we compare with the collected experimental world data [237] which are largely compatible with our curves within our quoted errors. Only the slope of our magnetic form factor differs somewhat from experiment. Furthermore, one can see that our results reproduce within their errors the experimental values of the magnetic moments both of the proton and of the neutron [76].

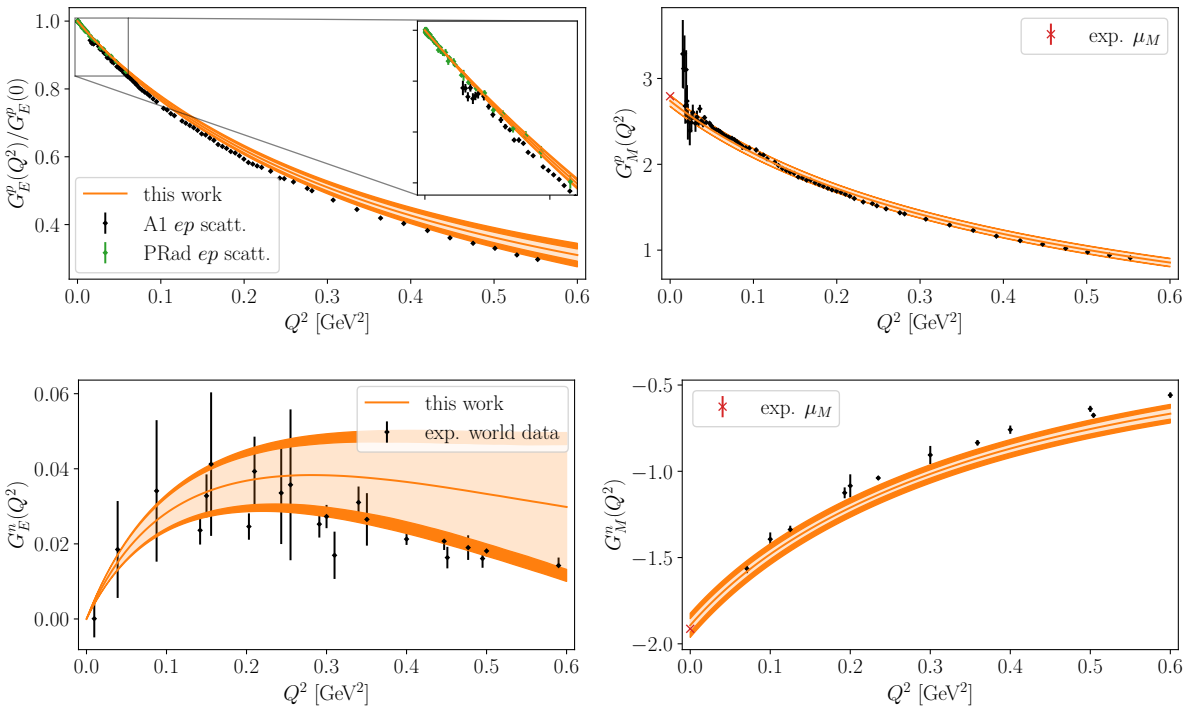


Figure 6.37: Electromagnetic form factors of the proton and neutron at the physical point as a function of Q^2. The orange curves and bands correspond to our final results with their full uncertainties obtained as model averages over the different direct fits. The light orange bands indicate the statistical uncertainties only. For the proton, the black diamonds represent the experimental *ep*-scattering data from A1 [14] obtained using Rosenbluth separation, and the green diamonds the corresponding data from PRad [26]. For the neutron, the black diamonds show the experimental world data collected in Ref. [237]. The experimental values of the magnetic moments [76] are depicted by red crosses.

6.11.2 Electromagnetic radii and magnetic moments

Regarding the electromagnetic radii and magnetic moments, we start by remarking that our values in the isovector channel, which are based on a similar, but updated and significantly extended data set, agree well with the results of Ref. [60]. We find similar errors on the electric radius and the magnetic moment and an improved error on the magnetic radius.

In fig. 6.38, we compare our results for the proton and neutron (*cf.*, table 6.3) to recent lattice determinations and to the experimental values. Before turning to a more detailed discussion of the results of the various lattice calculations, we give a brief overview of the ensembles and analysis techniques on which each of them is based:

- CSSM and QCDSF/UKQCD [50, 51] have computed the connected part of the electromagnetic form factors of the octet baryons on $N_f = 2 + 1$ ensembles with a lattice spacing of $a = 0.074$ fm and various pion masses going down to 310 MeV. From the Q^2-dependence of the form factors, they have extracted the electromagnetic radii and magnetic moments. Based on the aforementioned ensembles, they have performed an extrapolation to the physical pion mass, and they have corrected for finite-volume effects using a χPT calculation. Because they have only one lattice spacing, however, no continuum extrapolation has been performed. Moreover, they have employed only one value of the source-sink

separation and could thus not study excited-state effects.

- PACS [57] have extracted the electromagnetic radii and magnetic moments of the proton and neutron from a calculation of the connected part of the electromagnetic form factors on one $N_f = 2 + 1$ $(10.8\text{ fm})^4$ ensemble at the physical pion mass and with a lattice spacing of $a = 0.085$ fm. Consequently, no continuum extrapolation has been performed. Because of the large physical volume, their observables exhibit a much less pronounced excited-state contamination than ours, so that they did not need to perform an analysis as sophisticated as ours in this regard. Moreover, the large spatial volume gives them access to a large density of Q^2-points, and in particular also to a very small minimal value of Q^2, which is advantageous for an extraction of the radii. Their final results for the radii are based on dipole fits (except for $\langle r_E^2 \rangle^n$), and they quote the difference to a third-order z-expansion as a systematic error.

- PACS [238] have performed another calculation on a similar ensemble with a smaller lattice spacing of $a = 0.063$ fm. Here, their final results for the radii are based on a fourth-order z-expansion. Combining the two calculations, they can start to investigate the finite-lattice-spacing effects on the radii, which they find to be significant. However, they cannot yet perform a continuum extrapolation, for which at least a third lattice spacing will be required.

- ETMC [58] have extracted the electromagnetic radii and magnetic moments of the proton and neutron from a calculation of the electromagnetic form factors on one $N_f = 2 + 1 + 1$ twisted-mass ensemble at the physical pion mass and with a lattice spacing of $a = 0.080$ fm. They have included the light-quark disconnected contributions in their calculation. Because only one $N_f = 2 + 1 + 1$ ensemble enters the final results of this study, no continuum and infinite-volume extrapolation has been performed. However, they have compared two $N_f = 2$ ensembles with different volumes and did not find any significant finite-volume effects. They extract the ground-state form factors using plateau fits, but compare to the LO summation method and to two-state fits and quote a corresponding systematic error. Their final results for the radii are based on dipole fits (except for $\langle r_E^2 \rangle^n$), and they quote the difference to a fourth-order z-expansion as a systematic error.

- ETMC [59] have explored a method based on Fourier transformation which allows them to directly extract the radius without any parametrization of the Q^2-dependence, thus eliminating an important source of systematic uncertainty. They have demonstrated the method for the electric radii of the proton and neutron, using the same $N_f = 2 + 1 + 1$ ensemble as in Ref. [58], but not including the disconnected contributions.

- Ref. [60] is, as already mentioned, a previous study of the isovector electromagnetic form factors and radii on $N_f = 2+1$ CLS ensembles. They have, however, in comparison to this work fewer ensembles (in particular not E300, D450, N451 and N101), fewer source-sink separations and less statistics, especially on the near-physical pion mass ensemble E250. Otherwise, they employ an analysis relatively similar to ours. Because they do not have as many source-sink separations, they cannot investigate the dependence of the ground-state identification (*e.g.*, via the LO summation method) on $t_{\text{sep}}^{\text{min}}$, which is a major improvement of this work. Furthermore, many technical details in particular regarding the direct BχPT fits have been refined in this work.

We stress again that the only complete lattice study including disconnected contributions other than ours is Ref. [58][13], which, however, does not perform a continuum and infinite-volume

[13] and its predecessor on an $N_f = 2$ ensemble [54]

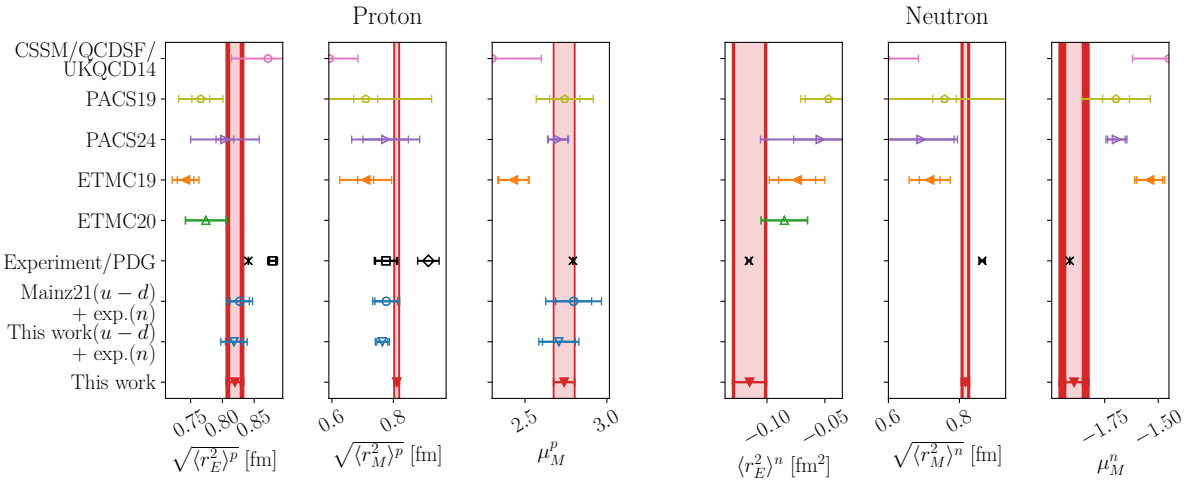


Figure 6.38: Comparison of our best estimates (red downward-pointing triangle) for the electromagnetic radii and the magnetic moments of the proton and neutron with other lattice calculations, *i.e.*, Mainz21 [60] (blue circle), ETMC20 [59] (green upward-pointing triangle), ETMC19 [58] (orange leftward-pointing triangle), PACS24 [238] (purple rightward-pointing triangle), PACS19 [57] (yellow pentagon) and CSSM/QCDSF/UKQCD14 [50, 51] (pink hexagon). Only studies with filled markers, *i.e.*, ETMC19 and this work, include disconnected contributions and hence represent a full lattice calculation. The Mainz21 values for the proton have been computed by combining their isovector results with the PDG values for the neutron [76]. We also show this estimate using our updated isovector results from this work (blue downward-pointing triangle). The experimental values for the neutron and for μ_M^p are taken from PDG [76] (black cross). The two data points for $\sqrt{\langle r_E^2\rangle^p}$ represent the values from PDG [76] (cross) and Mainz/A1 [14] (square), respectively. The two data points for $\sqrt{\langle r_M^2\rangle^p}$, on the other hand, depict the reanalysis of Ref. [40] either using the world data excluding that of Ref. [14] (diamond) or using only that of Ref. [14] (square). For ease of comparison, the red bands show our final results with the full uncertainty, with the light bands indicating the statistical errors.

extrapolation.

As can be seen from fig. 6.38, our estimates for the electric radii of the proton and neutron are larger in magnitude than the results of Refs. [57–59], while Ref. [51] quotes an even larger central value for $\sqrt{\langle r_E^2\rangle^p}$. The very recent result for $\sqrt{\langle r_E^2\rangle^p}$ from Ref. [238] is very close to ours, but has a significantly larger error, does not include disconnected contributions and is not extrapolated to the continuum limit (see above). We stress that any difference between our estimate and previous lattice calculations is not related to our preference for direct fits to the form factors, as opposed to the more traditional analysis via the z-expansion. In fact, the z-expansion approach yields similar values for our data (*cf.*, sections 6.7.2 and 6.7.3). Regarding $\langle r_E^2\rangle^n$, we remark that our result is not yet precise enough to discriminate between the value from *en* scattering [76] and the one combining a chiral effective field theory calculation of the deuteron structure radius [239] with the deuteron-proton charge-radius difference obtained from the hydrogen-deuterium isotope shift [240], between which a slight discrepancy has emerged.

Furthermore, we obtain results for the magnetic moments of the proton and neutron, as well as for $\sqrt{\langle r_M^2 \rangle^n}$, which are considerably larger in magnitude than that of Refs. [50, 58]. This improves the agreement with the experimental values [76]. In the case of the magnetic moments, the latter are very precisely known and are reproduced by our estimates within our quoted uncertainties. The values for μ_M^p from Refs. [57, 238] are well compatible with ours, while those for μ_M^n and $\sqrt{\langle r_M^2 \rangle^n}$ are smaller in magnitude than ours, in line with Ref. [58].

For $\sqrt{\langle r_M^2 \rangle^n}$, we observe, in spite of the good agreement which we achieve for μ_M^n, a $3.1\,\sigma$ tension between our result and the PDG value (after combining all errors in quadrature). On the level of the form factor G_M^n evaluated at any particular value of Q^2, however, the discrepancy is much smaller, as can be seen from fig. 6.37 (bottom right). For $\sqrt{\langle r_M^2 \rangle^p}$, our result agrees with those of Refs. [57, 58, 238] within 1.2 combined standard deviations, while that of Ref. [50] is much smaller. We note that our results for the isoscalar radii are somewhat larger than those of Ref. [58], while μ_M^{u+d-2s} compares well between our study and Ref. [58].

For the electric and magnetic radii of the proton, the experimental situation is much less clear than for the magnetic moment. As is the case for most of the other recent lattice calculations [57–59, 238], our result for $\sqrt{\langle r_E^2 \rangle^p}$ is much closer to the PDG value [76], which is completely dominated by muonic hydrogen spectroscopy, than to the A1 *ep*-scattering result [14]: while we only observe a very mild $1.5\,\sigma$ tension with the former, we disagree at the $3.7\,\sigma$ level with the latter (after combining all errors in quadrature). We note that we achieve an even better $0.6\,\sigma$ agreement with the recent *ep*-scattering experiment by PRad [26] which has also yielded a small electric radius of the proton.

For $\sqrt{\langle r_M^2 \rangle^p}$, on the other hand, our estimate is well compatible with the value inferred from the A1 experiment by the analyses [14, 40] and exhibits a sizable $2.8\,\sigma$ tension with the other collected world data [40]. As can be seen from fig. 6.37 (top right), the good agreement with A1 is not only observed in the magnetic radius, but also for the Q^2-dependence of the magnetic form factor over the whole range of Q^2 under study. We note that both the dispersive analysis of the combined Mainz/A1 + PRad data in Ref. [30] and the dispersively improved analysis of the Mainz/A1 data alone in Ref. [34] have yielded significantly larger magnetic radii [$\sqrt{\langle r_M^2 \rangle^p} = (0.847 \pm 0.004\ (\text{stat}) \pm 0.004\ (\text{syst}))$ fm and $(0.850 \pm 0.001\ (\text{stat})^{+0.009}_{-0.004}\ (\text{syst}))$ fm, respectively] than the z-expansion-based analysis of the Mainz/A1 data in Ref. [40]. The values of Refs. [30, 34] also exhibit a $3.4\,\sigma$ and $4.0\,\sigma$ tension with our result, respectively, which is partly due to their substantially smaller errors compared to Ref. [40]. Possible reasons for this discrepancy include unaccounted-for isospin-breaking effects.

Our statistical and systematic error estimates for the electric radii and magnetic moments are similar or smaller compared to the other lattice studies, while being substantially smaller for the magnetic radii. We remark that the missing data point at $Q^2 = 0$ complicates the extraction of the magnetic low-Q^2 observables in most recent lattice determinations, especially for z-expansion fits on individual ensembles. The direct approach has, additionally to combining information from several ensembles and from G_E and G_M, less freedom and by itself allows for considerably less variation in the form factors at low Q^2. We believe this to be responsible, in large part, for the small errors we achieve in the magnetic radii.

6.11.3 Zemach and Friar radii

Since there are, to our knowledge, at the time of writing no other lattice calculations of the Zemach and Friar radii, we can compare these quantities only to determinations based (at least in part) on experimental data, or on chiral perturbation theory. In the existing literature, three

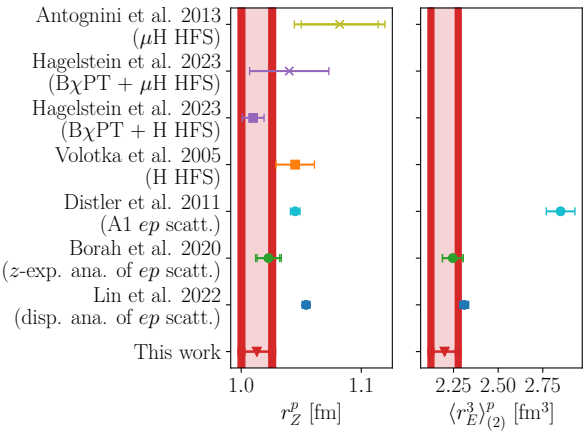


Figure 6.39: Comparison of our best estimates for the Zemach radius and third Zemach moment of the proton (red downward-pointing triangles) with determinations based on experimental data, *i.e.*, muonic hydrogen HFS [17, 226] (crosses), electronic hydrogen HFS [226, 245] (squares) and *ep* scattering [32, 219, 246] (circles).

main types of experiments have been employed to compute the Zemach radius of the proton: muonic hydrogen HFS [17], electronic hydrogen HFS [241] and *ep* scattering. In order to extract the proton Zemach radius from an HFS measurement, input on the proton-polarizability effect is required. This can be either taken from BχPT [226] or evaluated in a data-driven fashion, *i.e.*, using information on the spin structure functions [242–244] (as was done in Refs. [17, 245]). Between these two approaches, a tension has emerged, as already discussed in section 6.9.1. The form factors measured in *ep*-scattering experiments, on the other hand, can be analyzed with many different fit models, *e.g.*, by employing a (modified) power series [219], a z-expansion [246] or dispersion theory [32].

In fig. 6.39, we compare our results for the Zemach radius and third Zemach moment of the proton to such studies. While our result for r_Z^p agrees within one combined standard deviation with the extractions based on BχPT [226] and the z-expansion-based analysis of world *ep*-scattering data [246], and still within two combined standard deviations with the data-driven HFS extractions [17, 245] and the analysis of the A1 *ep*-scattering experiment [219], we observe a $2.6\,\sigma$ tension with the dispersive analysis of world *ep*-scattering data [32]. We also note that our estimate is smaller than all of the above experimental determinations except the one combining BχPT and electronic hydrogen HFS, which is slightly smaller than ours.

The proton's third Zemach moment can be extracted from *ep*-scattering experiments in the same way as its Zemach radius, and we also compare to these results in fig. 6.39. Again, our value is comparatively small, but this time in good agreement with both the z-expansion-based [246] and the dispersive analysis [32]. Against the analysis of the A1 *ep*-scattering experiment [219], on the other hand, we observe a clear tension of $5.3\,\sigma$ in $\langle r_E^3 \rangle_{(2)}^p$. We remark that the result of Ref. [246] for $\langle r_E^3 \rangle_{(2)}^p$, which is also based on the A1 data, is only as small as ours if they constrain the electric radius to the muonic hydrogen value (*cf.*, fig. 6.40 below). Without this constraint, they arrive at values for $\langle r_E^3 \rangle_{(2)}^p$ and r_Z^p compatible with those of Ref. [219].

In interpreting the aforementioned discrepancies, one must take into account that our results for the Zemach radii and third Zemach moments are not independent from those for the electromagnetic radii because they are based on the same lattice data for the form factors and the same BχPT fits. To quantify this correlation, we estimate the covariance matrix of our

model-averaged results for the different proton radii r_j,

$$C_{\widehat{jk}} = \frac{1}{(\mathrm{cdf}_{\mathcal{N}}^{-1}(3/4))^2}\,\mathrm{med}([r_j - \mathrm{med}(r_j)][r_k - \mathrm{med}(r_k)]), \tag{6.105}$$

where we have included the scale factor derived in eq. (6.41) involving the inverse CDF of a Gaussian distribution, which we denote here by $\mathrm{cdf}_{\mathcal{N}}^{-1}$. The median is calculated from the model-averaged (empirical) CDF built as in eq. (6.102). For the covariance matrix of $[\sqrt{\langle r_E^2\rangle^p}, \sqrt{\langle r_M^2\rangle^p}, r_Z^p, r_F^p]$, we thus obtain[14],

$$C = \begin{pmatrix} 1.92424878 \times 10^{-4} & 5.09426560 \times 10^{-5} & 1.86037985 \times 10^{-4} & 2.47985325 \times 10^{-4} \\ 5.09426560 \times 10^{-5} & 7.90873678 \times 10^{-5} & 1.00214335 \times 10^{-4} & 7.00503838 \times 10^{-5} \\ 1.86037985 \times 10^{-4} & 1.00214335 \times 10^{-4} & 2.44881039 \times 10^{-4} & 2.30139512 \times 10^{-4} \\ 2.47985325 \times 10^{-4} & 7.00503838 \times 10^{-5} & 2.30139512 \times 10^{-4} & 3.38165365 \times 10^{-4} \end{pmatrix}. \tag{6.106}$$

Using $\mathrm{corr}_{\widehat{jk}} = C_{\widehat{jk}}/\sqrt{C_{\widehat{jj}}C_{\widehat{kk}}}$ (no sum over j, k), eq. (6.106) corresponds to a correlation matrix of

$$\mathrm{corr} = \begin{pmatrix} 1 & 0.41294995 & 0.85702489 & 0.97214447 \\ 0.41294995 & 1 & 0.72010978 & 0.42834371 \\ 0.85702489 & 0.72010978 & 1 & 0.79974042 \\ 0.97214447 & 0.42834371 & 0.79974042 & 1 \end{pmatrix}. \tag{6.107}$$

Hence, we indeed observe a strong correlation both between $\sqrt{\langle r_E^2\rangle^p}$ and r_Z^p and between $\sqrt{\langle r_M^2\rangle^p}$ and r_Z^p, while our correlation between $\sqrt{\langle r_E^2\rangle^p}$ and r_F^p is even higher. We note that a large positive correlation between the proton's electric and Zemach radii has also been reported in the experimental literature [41, 247].

A selection of the results for r_Z^p and r_F^p from fig. 6.39 are plotted against the corresponding values for $\sqrt{\langle r_E^2\rangle^p}$ and $\sqrt{\langle r_M^2\rangle^p}$ in fig. 6.40. This demonstrates that if a particular analysis yields a larger electric radius, it generally also produces larger Zemach and Friar radii. The plots showing the magnetic radius on the horizontal axis, on the other hand, exhibit a significantly larger amount of scatter. Approximating the model-averaged distribution in each two-dimensional projection as a multivariate Gaussian with a covariance matrix given by the corresponding (2×2)-submatrix of eq. (6.106) in the vicinity of our central values, we can draw confidence ellipses as displayed in fig. 6.40. They illustrate again the strong correlation among our lattice results for the different proton radii and show that these confirm the trends observed in the data-driven evaluations as far as the correlation of the Zemach and Friar radii with the electric radius is concerned.

We conclude that our small results for $\sqrt{\langle r_E^2\rangle^p}$ and $\sqrt{\langle r_M^2\rangle^p}$ in table 6.3 naturally imply small values for r_Z^p and r_F^p. By contrast, dispersive analyses like Ref. [32] arrive at a significantly larger magnetic radius than the A1-data analyses [14, 40] and our Lattice-QCD-based extraction, as already mentioned in section 6.11.2. This may explain why we observe a larger tension in the Zemach radius (which equally probes electric and magnetic properties) with Ref. [32] than with Ref. [219], even though the situation is exactly the opposite for the third Zemach moment / Friar radius (which only probes the electric properties). For a deeper understanding of the

[14] We remark that the errors derived from the square root of the diagonal elements of eq. (6.106) differ slightly from the definition using the quantiles of the distribution which we have employed in table 6.3 above. This is because using only the covariance matrix amounts to approximating the distribution as Gaussian in the vicinity of the central values. Although this is a good approximation in our case, it does not need to be fulfilled exactly in general.

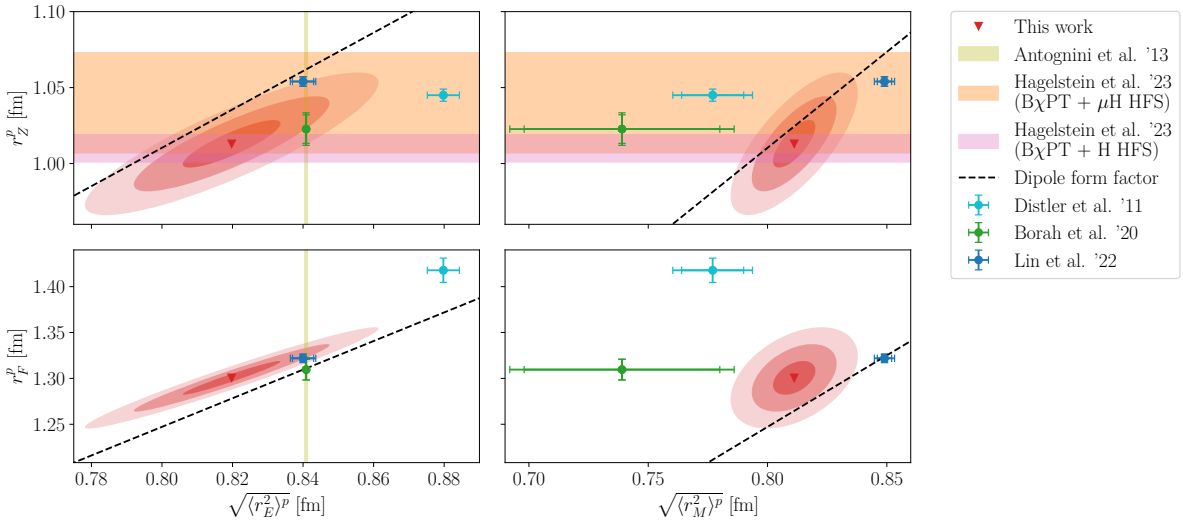


Figure 6.40: Correlation between the different proton radii. The results of this work (red downward-pointing triangles; shaded ellipses: $1\,\sigma$, $2\,\sigma$ and $3\,\sigma$ confidence regions) are compared with determinations based on experimental data (cyan, green and blue circles; *cf.*, fig. 6.39) [32, 219, 246]. The horizontal bands show the results for the Zemach radius from Ref. [226] combining BχPT with either muonic hydrogen HFS data (orange) or electronic hydrogen HFS data (pink). The vertical yellow band depicts the value for the electric radius derived from muonic-hydrogen Lamb shift [17]. The dashed black line is calculated with a dipole form (*cf.*, eq. (6.67)) for the electric and magnetic form factors assuming the same dipole mass for both form factors, which is varied along the line.

underlying differences, a comparison of the full Q^2-dependence of the form factors would be required, rather than merely of the radii. Furthermore, the role of higher-order electromagnetic corrections should be clarified.

The Zemach radius and third Zemach moment of the proton can also be computed in the framework of HBχPT [248], which yields a much larger value of $r_Z^p = 1.35$ fm and a much smaller value of $\langle r_E^3 \rangle_{(2)}^p = 0.8142$ fm^3. For the Zemach radius, the authors of Ref. [248] do not quote an error estimate and claim it to be in good agreement with the experimental results, so that the uncertainty is presumably rather large. For the third Zemach moment, they estimate the size of the uncomputed corrections which can, however, by far not explain their extremely low number.

Our results for the Zemach radius and third Zemach moment of the neutron are very well compatible with the z-expansion-based analysis of world *en*-scattering data [246], albeit with a more than two times larger error.

To summarize, the high correlation with the electromagnetic radii implies that our relatively low values for the Zemach and Friar radii of the proton do not give rise to an independent puzzle from the lattice perspective. Nevertheless, they emphasize the good agreement of our results with data-driven dispersive approaches regarding the electric properties of the proton and the corresponding tension regarding its magnetic properties (*cf.*, fig. 6.40 bottom left *vs.* top right). Thus, they make an important contribution towards reaching the goal of a complete and consistent picture of all the fundamental electromagnetic properties of the nucleon both from theory and from experiment. This is particularly relevant in light of the upcoming high-precision HFS measurements in muonic hydrogen spectroscopy [221–224].

7 Conclusions and outlook

In this thesis, we have investigated the electromagnetic form factors of the proton and neutron in Lattice QCD with $2+1$ flavors of dynamical quarks. From the Q^2-dependence of the form factors, we have extracted the electromagnetic radii which are of a special interest in light of the still unresolved puzzles surrounding their accurate experimental determination. In this context, a firm theoretical prediction from the Standard Model based on first principles is of hightened relevance in order to understand the origins of the observed discrepancies.

Our calculation includes both quark-connected and -disconnected contributions, thus representing a full lattice study. We have made use of the precise and effective computation of the relevant difference of the light- and strange-quark loops which has been achieved by the one-end trick [162, 164–166]. Combined with high-precision measurements of the nucleon two-point function also on ensembles with open boundary conditions, which have partly been performed in the context of this thesis, this enables an accurate determination of the quark-disconnected contributions.

Moreover, we have taken all relevant systematic effects into account in our analysis. In particular, we have explored three complementary approaches for the removal of excited-state contributions: two-state fits to the effective form factors, the LO summation method with the window average and the NLO summation method. We have found no evidence that any of the three methods introduces a systematic bias. However, for our data any fit *ansatz* including an explicit gap requires priors at least on these parameters. We have observed that the choice of the location as well as of the width of the priors can significantly influence the results for the ground-state form factors. Therefore, we prefer to avoid the use of priors in this step of the analysis by opting for the LO summation method. Here, we average over different choices for the minimal source-sink separation entering the fits as this parameter exerts a crucial influence on the results and we aim to minimize the human bias in this regard. We have adopted a conservative choice for the location of the averaging window with regard to the suppression of excited states.

By fitting our data for the form factors to the NLO expressions resulting from covariant baryon chiral perturbation theory (BχPT), we have combined the parametrization of their Q^2-dependence with the extrapolation to the physical point. Performing in this way a fit across several ensembles significantly increases the stability against applying strict momentum cuts which are desirable for the extraction of low-Q^2 observables such as the radii. Besides, this method yields results for the Q^2-dependence of the form factors at the physical point, which enable a more meaningful comparison to experiments. As a crosscheck, we have compared our results for the electromagnetic radii and magnetic moments from the BχPT fits to a more traditional procedure which parametrizes the form factors first on each ensemble using a model-independent z-expansion and subsequently extrapolates the resulting radii and magnetic moments to the physical point. In the cases where the z-expansion-based analysis is stable, we have found a good agreement between both approaches. Nevertheless, the BχPT fits are, in our opinion, superior in particular for the description of the magnetic form factor (the normalization of which is *a priori* unknown) because they enable a simultaneous fit of the electric and magnetic form factors with common fit parameters. Not having a strong preference for one specific setup of the BχPT fits, we use a model average to determine our final results, which allows us to quantify the systematic uncertainties associated with these variations.

We have found that little contribution is possible to the experimental debates regarding the large-Q^2 behavior of the form factors since our data quality rapidly decreases with rising momentum transfer. Consequently, we focus on low-Q^2 observables like the radii and magnetic moments, for which we can achieve a much higher precision. As an important benchmark, we reproduce the experimentally very precisely known magnetic moments within our quoted uncertainties. The precision of our results for the various proton radii (electric, magnetic, Zemach and Friar), which include a full error budget, is sufficient to make a meaningful contribution to the debate surrounding their determination. Our result for the electric (charge) radius of the proton, and correspondingly also that for the Friar radius, favors a low value. Thus, our electric radius is much closer to the one extracted from muonic hydrogen spectroscopy [17] and from the PRad *ep*-scattering experiment [26] than to the one obtained by A1 [14]. For the magnetic radius, and correspondingly also for the Zemach radius, our estimates favor low values as well. This is, in contrast to the electric radius, in better agreement with the analyses [14, 40] of the A1 data than with the corresponding analysis applied to the other collected world data [40] or with dispersive analyses of *ep*-scattering data [30, 32, 34]. Also the Q^2-dependence of our magnetic form factor of the proton over the whole range of Q^2 under study is well compatible with A1.

Consequently, we contribute additional evidence to suggest that lattice calculations agree with the emerging consensus about the experimental value of the electric proton radius [41, 249, 250]. Meanwhile, the results for the proton's magnetic radius (and its Zemach radius) require further investigation. In particular, a comparison of the full Q^2-dependence of the form factors themselves with dispersive approaches would be essential for understanding the origins of the observed discrepancies in the magnetic and Zemach radii. The ultimate goal of the ongoing and future research in this direction must be to attain a complete and consistent picture of all the fundamental electromagnetic properties of the nucleon both from theory and from experiment.

For lattice studies of the electromagnetic form factors of the nucleon, the excited-state contamination remains an important source of systematic uncertainty and a major challenge. The currently available data at sufficiently large source-sink separations is, without resorting to priors, not precise enough to clearly disentangle the contributions of multiple states, as would be required for a convincing application of multi-state fits or of the NLO summation method. Using the LO summation method, the signal gets lost in the exponentially growing noise very quickly after the plateau region is reached. This renders firm statements about the exact location of the plateau impossible and forces us to quote larger errors in case of ambiguity, which is achieved by the window average. A promising strategy to tackle this issue, besides drastically increasing statistics at large source-sink separations, is to perform a dedicated study of the excitation spectrum as initiated in Refs. [251, 252] for the axial and pseudoscalar form factors.

Of the remaining sources of systematic uncertainty, the modelling of the lattice artefacts, *i.e.*, the extrapolation to the continuum and infinite-volume limit, is generally the most important one in our analysis. This can be improved upon by using a larger range of lattice spacings and volumes at or near the physical pion mass. Small quark masses and coarse lattice spacings significantly enhance the probability of finding small eigenvalues of the Dirac operator. This raises concerns regarding the algorithmic stability of the current CLS setup in this regime. However, CLS is currently producing a fine ($\beta = 3.70$) ensemble at the physical pion mass called F300 which will be very helpful in further constraining the chiral interpolation and the continuum extrapolation. Moreover, its large volume implies a high density of Q^2-points and a small minimal value of Q^2, which is crucial for an accurate extraction of low-Q^2 observables like the radii.

A systematic error which has been left unquantified in this thesis is the one due to neglecting

higher orders in the BχPT expansion. Including them would require a derivation of the NNLO BχPT expressions for the electromagnetic form factors of the nucleon, which, to our knowledge, do not yet exist in the literature. Since the NLO expressions employed in this thesis are already very complicated, including the full NNLO terms would probably only be reasonable if one defines a more efficient scheme for their numerical evaluation. Besides, the fit quality of our NLO BχPT fits is already good (*cf.*, appendix E of Ref. [5]), so that the coefficients of NNLO terms are not likely to be well constrained by our data. Another type of corrections which has not been studied in this thesis is due to Quantum Electrodynamics (QED) and strong isospin breaking ($m_u \neq m_d$). Even though such corrections are generally believed to be negligible at our current level of precision, a precise estimate for their size does, to our knowledge, not exist at present. Calculating the relevant diagrams would, however, be numerically extremely costly. The current state of the art in this regard is computing isospin-breaking corrections to baryon masses [190] (*i.e.*, two-point functions); for three-point functions, this will be significantly more challenging.

Apart from computing the isoscalar electromagnetic form factors, the precise calculation of quark-disconnected diagrams used in this thesis can be employed for an updated study of the strangeness electromagnetic form factors of the nucleon [188]. Moreover, it can be used to investigate the isoscalar, strangeness and charm axial form factors. Such observables are relevant for elastic neutrino-nucleon scattering both via W- or Z-boson exchange. Here, the single-nucleon form factors serve as an important input for the calculation of neutrino-nucleus cross sections which are needed to interpret the results of upcoming long-baseline neutrino experiments aiming to measure the neutrino oscillation parameters to a high precision [253–255]. Furthermore, both the difference of the strangeness and charm axial form factors and the individual strangeness and charm vector form factors are required for hadronic corrections in experimental measurements of the proton weak charge at hadronic energy scales via parity-violating ep scattering [256, 257]. From the proton weak charge, the Weinberg mixing angle can be determined, the running of which sets constraints on extensions of the Standard Model [257].

Acknowledgments

Finally, I want to thank Prof. Dr. Hartmut Wittig for giving me the chance to join his working group and for his always friendly and supportive supervision, as well as for numerous pieces of advice on this project and beyond. Moreover, I have profited greatly from the PhD fellowship of the Mainz Physics Academy (MPA). This has given me the freedom to pursue my research project, while providing invaluable support through various soft-skill workshops, retreats and social activities, thus tremendously enriching my experience as a PhD student in Mainz. Here, particular thanks are due to Dr. Victoria Durant and Freya Luberg who are always giving their best to organize events according to the students' needs and without whom this rich program would not be conceivable. I also want to thank my fellow students in the MPA and in the Mainz lattice group, in particular Alexander Segner and Arnau Beltran Martínez, for accompanying me these past years. You have made this journey fun in spite of various hurdles and the not always easy times. Furthermore, I am grateful to my senior colleagues in the Mainz lattice group, Prof. Dr. Georg von Hippel, Prof. Dr. Harvey Meyer, Dr. Dalibor Djukanovic, Dr. Konstantin Ottnad and Dr. Jonna Koponen, for many inspirational discussions and valuable tips. Thanks equally go to them and to all the other group members for being more than just colleagues, for being friends. I also want to thank Dr. Franziska Hagelstein and her working group for helpful discussions on the topic of the Zemach and Friar radii. Lastly, special thanks are due to my parents and my sister who have accompanied me on my way through my studies and beyond like no one else.

Bibliography

[1] Dalibor Djukanovic *et al.* "Isoscalar electromagnetic form factors of the nucleon in $N_\mathrm{f} = 2 + 1$ lattice QCD." In: *Proceedings of The 38th International Symposium on Lattice Field Theory — PoS(LATTICE2021)*. Vol. 396. July 2022, p. 422. DOI: 10.22323/1.396.0422. arXiv: 2110.10626 [hep-lat].

[2] Miguel Salg *et al.* "Electromagnetic form factors of the proton and neutron from $N_f = 2{+}1$ lattice QCD." In: *Proceedings of The 39th International Symposium on Lattice Field Theory — PoS(LATTICE2022)*. Vol. 430. Apr. 2023, p. 121. DOI: 10.22323/1.430.0121. arXiv: 2211.17049 [hep-lat].

[3] Miguel Salg *et al.* "Proton and neutron electromagnetic radii and magnetic moments from $N_f = 2 + 1$ lattice QCD." In: *Proceedings of The 40th International Symposium on Lattice Field Theory — PoS(LATTICE2023)*. Vol. 453. Jan. 2024, p. 317. DOI: 10.22323/1.453.0317. arXiv: 2401.05404 [hep-lat].

[4] Miguel Salg *et al.* "Proton and neutron electromagnetic radii and magnetic moments from lattice QCD." In: *EPJ Web of Conferences* 303 (Sept. 2024), p. 02002. DOI: 10.1051/epjconf/202430302002. arXiv: 2401.09440 [hep-lat].

[5] Dalibor Djukanovic *et al.* "Electromagnetic form factors of the nucleon from $N_f = 2 + 1$ lattice QCD." In: *Physical Review D* 109.9 (May 2024), p. 094510. DOI: 10.1103/PhysRevD.109.094510. arXiv: 2309.06590 [hep-lat].

[6] Dalibor Djukanovic *et al.* "Precision calculation of the electromagnetic radii of the proton and neutron from lattice QCD." In: *Physical Review Letters* 132.21 (May 2024), p. 211901. DOI: 10.1103/PhysRevLett.132.211901. arXiv: 2309.07491 [hep-lat].

[7] Dalibor Djukanovic *et al.* "Zemach and Friar radii of the proton and neutron from lattice QCD." In: *Physical Review D* 110.1 (July 2024), p. L011503. DOI: 10.1103/PhysRevD.110.L011503. arXiv: 2309.17232 [hep-lat].

[8] David P. Aguillard *et al.* (The Muon $g-2$ Collaboration). "Measurement of the Positive Muon Anomalous Magnetic Moment to 0.20 ppm." In: *Physical Review Letters* 131.16 (Oct. 2023), p. 161802. DOI: 10.1103/PhysRevLett.131.161802. arXiv: 2308.06230 [hep-ex].

[9] Tatsumi Aoyama *et al.* "The anomalous magnetic moment of the muon in the Standard Model." In: *Physics Reports* 887 (Dec. 2020), pp. 1–166. DOI: 10.1016/j.physrep.2020.07.006. arXiv: 2006.04822 [hep-ph].

[10] Szabolcs Borsányi *et al.* "Leading hadronic contribution to the muon magnetic moment from lattice QCD." In: *Nature* 593.7857 (Apr. 2021), pp. 51–55. DOI: 10.1038/s41586-021-03418-1. arXiv: 2002.12347 [hep-lat].

[11] Simon Kuberski. "Muon $g-2$: Lattice calculations of the hadronic vacuum polarization." In: *Proceedings of The 40th International Symposium on Lattice Field Theory — PoS(LATTICE2023)*. Vol. 453. 2024, p. 125. DOI: 10.22323/1.453.0125. arXiv: 2312.13753 [hep-lat].

[12] Jean-Philippe Karr, Dominique Marchand, and Eric Voutier. "The proton size." In: *Nature Reviews Physics* 2.11 (Sept. 2020), pp. 601–614. DOI: 10.1038/s42254-020-0229-x.

[13] E. E. Chambers and R. Hofstadter. "Structure of the Proton." In: *Physical Review* 103.5 (Sept. 1956), pp. 1454–1463. DOI: 10.1103/PhysRev.103.1454.

[14] Jan C. Bernauer *et al.* (A1 Collaboration). "Electric and magnetic form factors of the proton." In: *Physical Review C* 90.1 (July 2014), p. 015206. DOI: 10.1103/PhysRevC.90.015206. arXiv: 1307.6227 [nucl-ex].

[15] Peter J. Mohr, Barry N. Taylor, and David B. Newell. "CODATA recommended values of the fundamental physical constants: 2010." In: *Reviews of Modern Physics* 84.4 (Nov. 2012), pp. 1527–1605. DOI: 10.1103/RevModPhys.84.1527. arXiv: 1203.5425 [physics.atom-ph].

[16] Randolf Pohl *et al.* "The size of the proton." In: *Nature* 466.7303 (July 2010), pp. 213–216. DOI: 10.1038/nature09250.

[17] Aldo Antognini *et al.* "Proton Structure from the Measurement of 2S-2P Transition Frequencies of Muonic Hydrogen." In: *Science* 339.6118 (Jan. 2013), pp. 417–420. DOI: 10.1126/science.1230016.

[18] Carl E. Carlson. "The proton radius puzzle." In: *Progress in Particle and Nuclear Physics* 82 (May 2015), pp. 59–77. DOI: 10.1016/j.ppnp.2015.01.002. arXiv: 1502.05314 [hep-ph].

[19] Axel Beyer *et al.* "The Rydberg constant and proton size from atomic hydrogen." In: *Science* 358.6359 (Oct. 2017), pp. 79–85. DOI: 10.1126/science.aah6677.

[20] Nikita Bezginov *et al.* "A measurement of the atomic hydrogen Lamb shift and the proton charge radius." In: *Science* 365.6457 (Sept. 2019), pp. 1007–1012. DOI: 10.1126/science.aau7807.

[21] Alexey Grinin *et al.* "Two-photon frequency comb spectroscopy of atomic hydrogen." In: *Science* 370.6520 (Nov. 2020), pp. 1061–1066. DOI: 10.1126/science.abc7776.

[22] Simon Scheidegger and Frédéric Merkt. "Precision-Spectroscopic Determination of the Binding Energy of a Two-Body Quantum System: The Hydrogen Atom and the Proton-Size Puzzle." In: *Physical Review Letters* 132.11 (Mar. 2024), p. 113001. DOI: 10.1103/PhysRevLett.132.113001.

[23] Hélène Fleurbaey *et al.* "New Measurement of the $1S - 3S$ Transition Frequency of Hydrogen: Contribution to the Proton Charge Radius Puzzle." In: *Physical Review Letters* 120.18 (May 2018), p. 183001. DOI: 10.1103/PhysRevLett.120.183001. arXiv: 1801.08816 [physics.atom-ph].

[24] A. D. Brandt *et al.* "Measurement of the $2S_{1/2} - 8D_{5/2}$ Transition in Hydrogen." In: *Physical Review Letters* 128.2 (Jan. 2022), p. 023001. DOI: 10.1103/PhysRevLett.128.023001. arXiv: 2111.08554 [physics.atom-ph].

[25] Miha Mihovilovič *et al.* "The proton charge radius extracted from the initial-state radiation experiment at MAMI." In: *The European Physical Journal A* 57.3 (Mar. 2021), p. 107. DOI: 10.1140/epja/s10050-021-00414-x. arXiv: 1905.11182 [nucl-ex].

[26] Weizhi Xiong *et al.* "A small proton charge radius from an electron-proton scattering experiment." In: *Nature* 575.7781 (Nov. 2019), pp. 147–150. DOI: 10.1038/s41586-019-1721-2.

[27] Patrick Mergell, Ulf-G. Meißner, and Dieter Drechsel. "Dispersion-theoretical analysis of the nucleon electromagnetic form factors." In: *Nuclear Physics A* 596.3 (Jan. 1996), pp. 367–396. DOI: 10.1016/0375-9474(95)00339-8. arXiv: hep-ph/9506375.

[28] M. A. Belushkin, Hans-Werner Hammer, and Ulf-G. Meißner. "Dispersion analysis of the nucleon form factors including meson continua." In: *Physical Review C* 75.3 (Mar. 2007), p. 035202. DOI: 10.1103/PhysRevC.75.035202. arXiv: hep-ph/0608337.

[29] Martin Hoferichter *et al.* "On the $\pi\pi$ continuum in the nucleon form factors and the proton radius puzzle." In: *The European Physical Journal A* 52.11 (Nov. 2016), p. 331. DOI: 10.1140/epja/i2016-16331-7. arXiv: 1609.06722 [hep-ph].

[30] Yong-Hui Lin, Hans-Werner Hammer, and Ulf-G. Meißner. "High-precision determination of the electric and magnetic radius of the proton." In: *Physics Letters B* 816 (May 2021), p. 136254. DOI: 10.1016/j.physletb.2021.136254. arXiv: 2102.11642 [hep-ph].

[31] Yong-Hui Lin, Hans-Werner Hammer, and Ulf-G. Meißner. "Dispersion-theoretical analysis of the electromagnetic form factors of the nucleon: Past, present and future." In: *The European Physical Journal A* 57.8 (Aug. 2021), p. 255. DOI: 10.1140/epja/s10050-021-00562-0. arXiv: 2106.06357 [hep-ph].

[32] Yong-Hui Lin, Hans-Werner Hammer, and Ulf-G. Meißner. "New Insights into the Nucleon's Electromagnetic Structure." In: *Physical Review Letters* 128.5 (Feb. 2022), p. 052002. DOI: 10.1103/PhysRevLett.128.052002. arXiv: 2109.12961 [hep-ph].

[33] Ina T. Lorenz *et al.* "Theoretical constraints and systematic effects in the determination of the proton form factors." In: *Physical Review D* 91.1 (Jan. 2015), p. 014023. DOI: 10.1103/PhysRevD.91.014023. arXiv: 1411.1704 [hep-ph].

[34] Jose Manuel Alarcón, Douglas W. Higinbotham, and Christian Weiss. "Precise determination of the proton magnetic radius from electron scattering data." In: *Physical Review C* 102.3 (Sept. 2020), p. 035203. DOI: 10.1103/PhysRevC.102.035203. arXiv: 2002.05167 [hep-ph].

[35] Silke Grieser *et al.* "A cryogenic supersonic jet target for electron scattering experiments at MAGIX@MESA and MAMI." In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 906 (Oct. 2018), pp. 120–126. DOI: 10.1016/j.nima.2018.07.076. arXiv: 1806.05409 [physics.ins-det].

[36] Ashot Gasparian *et al.* (PRad Collaboration). "PRad-II: A New Upgraded High Precision Measurement of the Proton Charge Radius." Jan. 2022. arXiv: 2009.10510 [nucl-ex].

[37] Toshimi Suda. "Low-energy electron scattering facilities in Japan." In: *Journal of Physics: Conference Series* 2391 (Dec. 2022), p. 012004. DOI: 10.1088/1742-6596/2391/1/012004.

[38] Ethan Cline *et al.* "MUSE: The MUon Scattering Experiment." In: *SciPost Physics Proceedings* 5 (Sept. 2021), p. 023. DOI: 10.21468/SciPostPhysProc.5.023.

[39] Catarina Quintans (AMBER collaboration). "The New AMBER Experiment at the CERN SPS." In: *Few-Body Systems* 63.4 (Sept. 2022), p. 72. DOI: 10.1007/s00601-022-01769-7.

[40] Gabriel Lee, John R. Arrington, and Richard J. Hill. "Extraction of the proton radius from electron-proton scattering data." In: *Physical Review D* 92.1 (July 2015), p. 013013. DOI: 10.1103/PhysRevD.92.013013. arXiv: 1505.01489 [hep-ph].

[41] Aldo Antognini, Franziska Hagelstein, and Vladimir Pascalutsa. "The proton structure in and out of muonic hydrogen." In: *Annual Review of Nuclear and Particle Science* 72.1 (Sept. 2022), pp. 389–418. DOI: 10.1146/annurev-nucl-101920-024709. arXiv: 2205.10076 [nucl-th].

[42] Ben Ledroit and Kurt Aulenbacher. "Collimation of target induced halo following MAGIX at MESA." In: *Journal of Physics: Conference Series* 1350 (Dec. 2019), p. 012138. DOI: 10.1088/1742-6596/1350/1/012138.

[43] Stephan Paul. "Measuring Hadron Charge Radii with AMBER." Mar. 2024. URL: https://indico.cern.ch/event/1367212/contributions/5839793/attachments/2822653/4929725/Paw24%202024%20ChargeRadii-AMBER%20Paul.pdf. Slides presented at PAW'24 - Physics at AMBER International Workshop 2024.

[44] Meinulf Göckeler *et al.* (QCDSF Collaboration). "Nucleon electromagnetic form factors on the lattice and in chiral effective field theory." In: *Physical Review D* 71.3 (Feb. 2005), p. 034508. DOI: 10.1103/PhysRevD.71.034508. arXiv: hep-lat/0303019.

[45] Takeshi Yamazaki *et al.* (RBC and UKQCD Collaborations). "Nucleon form factors with $2+1$ flavor dynamical domain-wall fermions." In: *Physical Review D* 79.11 (June 2009), p. 114505. DOI: 10.1103/PhysRevD.79.114505. arXiv: 0904.2039 [hep-lat].

[46] Sergey N. Syritsyn *et al.* (LHPC Collaboration). "Nucleon electromagnetic form factors from lattice QCD using $2+1$ flavor domain wall fermions on fine lattices and chiral perturbation theory." In: *Physical Review D* 81.3 (Feb. 2010), p. 034507. DOI: 10.1103/PhysRevD.81.034507. arXiv: 0907.4194 [hep-lat].

[47] Jonathan D. Bratt *et al.* (LHPC). "Nucleon structure from mixed action calculations using $2+1$ flavors of asqtad sea and domain wall valence fermions." In: *Physical Review D* 82.9 (Nov. 2010), p. 094502. DOI: 10.1103/PhysRevD.82.094502. arXiv: 1001.3620 [hep-lat].

[48] Constantia Alexandrou *et al.* "Nucleon form factors and moments of generalized parton distributions using N_f=2+1+1 twisted mass fermions." In: *Physical Review D* 88.1 (July 2013), p. 014509. DOI: 10.1103/PhysRevD.88.014509. arXiv: 1303.5979 [hep-lat].

[49] Tanmoy Bhattacharya *et al.* "Nucleon charges and electromagnetic form factors from $2+1+1$-flavor lattice QCD." In: *Physical Review D* 89.9 (May 2014), p. 094502. DOI: 10.1103/PhysRevD.89.094502. arXiv: 1306.5435 [hep-lat].

[50] Phiala E. Shanahan *et al.* (CSSM and QCDSF/UKQCD Collaborations). "Magnetic form factors of the octet baryons from lattice QCD and chiral extrapolation." In: *Physical Review D* 89.7 (Apr. 2014), p. 074511. DOI: 10.1103/PhysRevD.89.074511. arXiv: 1401.5862 [hep-lat].

[51] Phiala E. Shanahan *et al.* (CSSM and QCDSF/UKQCD Collaborations). "Electric form factors of the octet baryons from lattice QCD and chiral extrapolation." In: *Physical Review D* 90.3 (Aug. 2014), p. 034502. DOI: 10.1103/PhysRevD.90.034502. arXiv: 1403.1965 [hep-lat].

[52] Jeremy R. Green *et al.* "Nucleon electromagnetic form factors from lattice QCD using a nearly physical pion mass." In: *Physical Review D* 90.7 (Oct. 2014), p. 074507. DOI: 10.1103/PhysRevD.90.074507. arXiv: 1404.4029 [hep-lat].

[53] Stefano Capitani *et al.* "Nucleon electromagnetic form factors in two-flavor QCD." In: *Physical Review D* 92.5 (Sept. 2015), p. 054511. DOI: 10.1103/PhysRevD.92.054511. arXiv: 1504.04628 [hep-lat].

[54] Constantia Alexandrou *et al.* "Nucleon electromagnetic form factors using lattice simulations at the physical point." In: *Physical Review D* 96.3 (Aug. 2017), p. 034503. DOI: 10.1103/PhysRevD.96.034503. arXiv: 1706.00469 [hep-lat].

[55] Nesreen Hasan *et al.* "Computing the nucleon charge and axial radii directly at $Q^2 = 0$ in lattice QCD." In: *Physical Review D* 97.3 (Feb. 2018), p. 034504. DOI: 10.1103/PhysRevD.97.034504. arXiv: 1711.11385 [hep-lat].

[56] Ken-Ichi Ishikawa *et al.* (PACS Collaboration). "Nucleon form factors on a large volume lattice near the physical point in $2 + 1$ flavor QCD." In: *Physical Review D* 98.7 (Oct. 2018), p. 074510. DOI: 10.1103/PhysRevD.98.074510. arXiv: 1807.03974 [hep-lat].

[57] Eigo Shintani *et al.* (PACS Collaboration). "Nucleon form factors and root-mean-square radii on a $(10.8 \text{ fm})^4$ lattice at the physical point." In: *Physical Review D* 99.1 (Jan. 2019), p. 014510. DOI: 10.1103/PhysRevD.99.014510. arXiv: 1811.07292 [hep-lat]. Erratum ibid. [258].

[58] Constantia Alexandrou *et al.* "Proton and neutron electromagnetic form factors from lattice QCD." In: *Physical Review D* 100.1 (July 2019), p. 014509. DOI: 10.1103/PhysRevD.100.014509. arXiv: 1812.10311 [hep-lat].

[59] Constantia Alexandrou *et al.* "Model-independent determination of the nucleon charge radius from lattice QCD." In: *Physical Review D* 101.11 (June 2020), p. 114504. DOI: 10.1103/PhysRevD.101.114504. arXiv: 2002.06984 [hep-lat].

[60] Dalibor Djukanovic *et al.* "Isovector electromagnetic form factors of the nucleon from lattice QCD and the proton radius puzzle." In: *Physical Review D* 103.9 (May 2021), p. 094522. DOI: 10.1103/PhysRevD.103.094522. arXiv: 2102.07460 [hep-lat].

[61] Ken-Ichi Ishikawa *et al.* (PACS Collaboration). "Calculation of derivative of nucleon form factors in $N_f = 2 + 1$ lattice QCD at $M_\pi = 138$ MeV on a $(5.5 \text{ fm})^3$ volume." In: *Physical Review D* 104.7 (Oct. 2021), p. 074514. DOI: 10.1103/PhysRevD.104.074514. arXiv: 2107.07085 [hep-lat].

[62] Steven Weinberg. *The quantum theory of fields: 1. Foundations*. Cambridge: Cambridge University Press, 1995. DOI: 10.1017/CBO9781139644167.

[63] Steven Weinberg. *The quantum theory of fields: 2. Modern applications*. Cambridge: Cambridge University Press, 1996. DOI: 10.1017/CBO9781139644174.

[64] Michael Edward Peskin and Daniel V. Schroeder. *An Introduction To Quantum Field Theory*. Boca Raton: CRC Press, 2018. DOI: 10.1201/9780429503559.

[65] Owe Philipsen. *Quantenfeldtheorie und das Standardmodell der Teilchenphysik: Eine Einführung*. Berlin, Heidelberg: Springer Spektrum, 2018. DOI: 10.1007/978-3-662-57820-9.

[66] Jean Zinn-Justin. *Quantum field theory and critical phenomena*. Vol. 171. International series of monographs on physics. Oxford: Oxford University Press, Apr. 2021. DOI: 10.1093/oso/9780198834625.001.0001.

[67] Jonas Wilhelm. "The strangeness form factors of the nucleon from lattice QCD." PhD thesis. Mainz: Johannes Gutenberg-Universität Mainz, 2019.

[68] Andreas Risch. "Isospin breaking effects in hadronic matrix elements on the lattice." PhD thesis. Mainz: Johannes Gutenberg-Universität Mainz, 2021. DOI: 10.25358/openscience-6314.

[69] Alexandre Deur, Stanley J. Brodsky, and Craig D. Roberts. "QCD running couplings and effective charges." In: *Progress in Particle and Nuclear Physics* 134 (Jan. 2024), p. 104081. DOI: 10.1016/j.ppnp.2023.104081. arXiv: 2303.00723 [hep-ph].

[70] David J. Gross and Frank Wilczek. "Ultraviolet Behavior of Non-Abelian Gauge Theories." In: *Physical Review Letters* 30.26 (June 1973), pp. 1343–1346. DOI: 10.1103/PhysRevLett.30.1343.

[71] H. David Politzer. "Reliable Perturbative Results for Strong Interactions?" In: *Physical Review Letters* 30.26 (June 1973), pp. 1346–1349. DOI: 10.1103/PhysRevLett.30.1346.

[72] Harald Fritzsch, Murray Gell-Mann, and Heinrich Leutwyler. "Advantages of the color octet gluon picture." In: *Physics Letters B* 47.4 (Nov. 1973), pp. 365–368. DOI: 10.1016/0370-2693(73)90625-4.

[73] Kenneth G. Wilson. "Confinement of quarks." In: *Physical Review D* 10.8 (Oct. 1974), pp. 2445–2459. DOI: 10.1103/PhysRevD.10.2445.

[74] Martin Lüscher. "Lattice QCD – from Quark Confinement to Asymptotic Freedom." In: *Annales Henri Poincaré* 4 (Dec. 2003), pp. 197–210. DOI: 10.1007/s00023-003-0916-z. arXiv: hep-ph/0211220.

[75] Hideo Suganuma *et al.* "Lattice QCD Study for Confinement in Hadrons." In: *AIP Conference Proceedings* 1388.1 (Oct. 2011), pp. 195–201. DOI: 10.1063/1.3647373. arXiv: 1103.4015 [hep-lat].

[76] Ron L. Workman *et al.* (Particle Data Group). "Review of Particle Physics." In: *Progress of Theoretical and Experimental Physics* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.

[77] Yoichiro Nambu and Giovanni Jona-Lasinio. "Dynamical Model of Elementary Particles Based on an Analogy with Superconductivity. I." In: *Physical Review* 122.1 (Apr. 1961), pp. 345–358. DOI: 10.1103/PhysRev.122.345.

[78] Jeffrey Goldstone. "Field theories with « Superconductor » solutions." In: *Il Nuovo Cimento* 19.1 (Jan. 1961), pp. 154–164. DOI: 10.1007/BF02812722.

[79] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. "Broken Symmetries." In: *Physical Review* 127.3 (Aug. 1962), pp. 965–970. DOI: 10.1103/PhysRev.127.965.

[80] István Montvay and Gernot Münster. *Quantum Fields on a Lattice*. Cambridge Monographs on Mathematical Physics. Cambridge: Cambridge University Press, 1994. DOI: 10.1017/CBO9780511470783.

[81] Konrad Osterwalder and Robert Schrader. "Axioms for Euclidean Green's functions." In: *Communications in Mathematical Physics* 31 (June 1973), pp. 83–112. DOI: 10.1007/BF01645738.

[82] Konrad Osterwalder and Robert Schrader. "Axioms for Euclidean Green's functions II." In: *Communications in Mathematical Physics* 42 (Oct. 1975), pp. 281–305. DOI: 10.1007/BF01608978.

[83] Christof Gattringer and Christian B. Lang. *Quantum Chromodynamics on the Lattice. An Introductory Presentation*. Vol. 788. Lecture Notes in Physics. Berlin, Heidelberg: Springer, 2010. DOI: 10.1007/978-3-642-01850-3.

[84] Benjamin Jäger. "Hadronic matrix elements in lattice QCD." PhD thesis. Mainz: Johannes Gutenberg-Universität Mainz, May 2014. DOI: 10.25358/openscience-4245.

[85] Alessandro Sciarra. "The QCD phase diagram at purely imaginary chemical potential from the lattice." PhD thesis. Frankfurt am Main: Johann Wolfgang Goethe-Universität Frankfurt am Main, 2017. URL: https://publikationen.ub.uni-frankfurt.de/frontdoor/index/index/docId/44416.

[86] Kurt Symanzik. "Continuum limit and improved action in lattice theories: (I). Principles and φ^4 theory." In: *Nuclear Physics B* 226.1 (Sept. 1983), pp. 187–204. DOI: 10.1016/0550-3213(83)90468-6.

[87] Kurt Symanzik. "Continuum limit and improved action in lattice theories: (II). O(N) non-linear sigma model in perturbation theory." In: *Nuclear Physics B* 226.1 (Sept. 1983), pp. 205–227. DOI: 10.1016/0550-3213(83)90469-8.

[88] Peter Weisz. "Continuum limit improved lattice action for pure Yang-Mills theory (I)." In: *Nuclear Physics B* 212.1 (Feb. 1983), pp. 1–17. DOI: 10.1016/0550-3213(83)90595-3.

[89] Martin Lüscher and Peter Weisz. "On-shell improved lattice gauge theories." In: *Communications in Mathematical Physics* 97.1 (Mar. 1985), pp. 59–77. DOI: 10.1007/BF01206178. Erratum ibid. [259].

[90] Mattia Bruno *et al.* "Simulation of QCD with $N_\mathrm{f} = 2 + 1$ flavors of non-perturbatively improved Wilson fermions." In: *Journal of High Energy Physics* 2015.2 (Feb. 2015), p. 43. DOI: 10.1007/JHEP02(2015)043. arXiv: 1411.3982 [hep-lat].

[91] Kenneth G. Wilson. "Quarks and Strings on a Lattice." In: *New Phenomena in Subnuclear Physics: Part A*. Ed. by Antonino Zichichi. Vol. 13. The Subnuclear Series. Boston, MA: Springer, 1977, pp. 69–142. DOI: 10.1007/978-1-4613-4208-3_6.

[92] H. B. Nielsen and M. Ninomiya. "Absence of neutrinos on a lattice: (I). Proof by homotopy theory." In: *Nuclear Physics B* 185.1 (July 1981), pp. 20–40. DOI: 10.1016/0550-3213(81)90361-8.

[93] H. B. Nielsen and M. Ninomiya. "Absence of neutrinos on a lattice: (II). Intuitive topological proof." In: *Nuclear Physics B* 193.1 (Dec. 1981), pp. 173–194. DOI: 10.1016/0550-3213(81)90524-1.

[94] H. B. Nielsen and M. Ninomiya. "A no-go theorem for regularizing chiral fermions." In: *Physics Letters B* 105.2 (Oct. 1981), pp. 219–223. DOI: 10.1016/0370-2693(81)91026-1.

[95] B. Sheikholeslami and R. Wohlert. "Improved continuum limit lattice action for QCD with Wilson fermions." In: *Nuclear Physics B* 259.4 (Sept. 1985), pp. 572–596. DOI: 10.1016/0550-3213(85)90002-1.

[96] Martin Lüscher *et al.* "Chiral symmetry and O(a) improvement in lattice QCD." In: *Nuclear Physics B* 478.1 (Oct. 1996), pp. 365–397. DOI: 10.1016/0550-3213(96)00378-1. arXiv: hep-lat/9605038.

[97] Martin Lüscher *et al.* "Non-perturbative O(a) improvement of lattice QCD." In: *Nuclear Physics B* 491.1 (Apr. 1997), pp. 323–343. DOI: 10.1016/S0550-3213(97)00080-1. arXiv: hep-lat/9609035.

[98] D. H. Weingarten and D. N. Petcher. "Monte Carlo integration for lattice gauge theories with fermions." In: *Physics Letters B* 99.4 (Feb. 1981), pp. 333–338. DOI: 10.1016/0370-2693(81)90112-X.

[99] Nicholas Metropolis *et al.* "Equation of State Calculations by Fast Computing Machines." In: *The Journal of Chemical Physics* 21.6 (June 1953), pp. 1087–1092. DOI: 10.1063/1.1699114.

[100] R. T. Scalettar, D. J. Scalapino, and R. L. Sugar. "New algorithm for the numerical simulation of fermions." In: *Physical Review B* 34.11 (Dec. 1986), pp. 1911–7917. DOI: 10.1103/PhysRevB.34.7911.

[101] Simon Duane and John B. Kogut. "Hybrid Stochastic Differential Equations Applied to Quantum Chromodynamics." In: *Physical Review Letters* 55.25 (Dec. 1985), pp. 2774–2777. DOI: 10.1103/PhysRevLett.55.2774.

[102] Simon Duane and John B. Kogut. "The theory of hybrid stochastic algorithms." In: *Nuclear Physics B* 275.3 (Nov. 1986), pp. 398–420. DOI: 10.1016/0550-3213(86)90606-1.

[103] Simon Duane *et al.* "Hybrid Monte Carlo." In: *Physics Letters B* 195.2 (1987), pp. 216–222. DOI: 10.1016/0370-2693(87)91197-X.

[104] Steven Gottlieb *et al.* "Hybrid-molecular-dynamics algorithms for the numerical simulation of quantum chromodynamics." In: *Physical Review D* 35.8 (Apr. 1987), pp. 2531–2542. DOI: 10.1103/PhysRevD.35.2531.

[105] Ivan Horváth, Anthony D. Kennedy, and Stefan Sint. "A new exact method for dynamical fermion computations with non-local actions." In: *Nuclear Physics B – Proceedings Supplements* 73.1 (Mar. 1999), pp. 834–836. DOI: 10.1016/S0920-5632(99)85217-7. arXiv: hep-lat/9809092.

[106] M. A. Clark and Anthony D. Kennedy. "Accelerating Dynamical-Fermion Computations Using the Rational Hybrid Monte Carlo Algorithm with Multiple Pseudofermion Fields." In: *Physical Review Letters* 98.5 (Jan. 2007), p. 051601. DOI: 10.1103/PhysRevLett.98.051601. arXiv: hep-lat/0608015.

[107] Alan M. Ferrenberg and Robert H. Swendsen. "New Monte Carlo technique for studying phase transitions." In: *Physical Review Letters* 61.23 (Dec. 1988), pp. 2635–2638. DOI: 10.1103/PhysRevLett.61.2635. Erratum ibid. [260].

[108] Alan M. Ferrenberg and Robert H. Swendsen. "Optimized Monte Carlo data analysis." In: *Physical Review Letters* 63.12 (Sept. 1989), pp. 1195–1198. DOI: 10.1103/PhysRevLett.63.1195.

[109] Daniel Mohler and Stefan Schaefer. "Remarks on strange-quark simulations with Wilson fermions." In: *Physical Review D* 102.7 (Oct. 2020), p. 074506. DOI: 10.1103/PhysRevD.102.074506. arXiv: 2003.13359 [hep-lat].

[110] Karl Jansen *et al.* "Non-perturbative renormalization of lattice QCD at all scales." In: *Physics Letters B* 372.3 (Apr. 1996), pp. 275–282. DOI: 10.1016/0370-2693(96)00075-5. arXiv: hep-lat/9512009.

[111] Tanmoy Bhattacharya *et al.* "Improved bilinears in lattice QCD with nondegenerate quarks." In: *Physical Review D* 73.3 (Feb. 2006), p. 034504. DOI: 10.1103/PhysRevD.73.034504. arXiv: hep-lat/0511014.

[112] Antoine Gérardin, Tim Harris, and Harvey B. Meyer. "Nonperturbative renormalization and O(a)-improvement of the nonsinglet vector current with $N_f = 2 + 1$ Wilson fermions and tree-level Symanzik improved gauge action." In: *Physical Review D* 99.1 (Jan. 2019), p. 014519. DOI: 10.1103/PhysRevD.99.014519. arXiv: 1811.08209 [hep-lat].

[113] Roberto Frezzotti, Stefan Sint, and Peter Weisz. "O(a) improved twisted mass lattice QCD." In: *Journal of High Energy Physics* 2001.07 (Aug. 2001), p. 048. DOI: 10.1088/1126-6708/2001/07/048. arXiv: hep-lat/0104014.

[114] Martin Lüscher. *openQCD*. URL: https://luscher.web.cern.ch/luscher/openQCD/.

[115] Mattia Bruno, Tomasz Korzec, and Stefan Schaefer. "Setting the scale for the CLS $2 + 1$ flavor ensembles." In: *Physical Review D* 95.7 (Apr. 2017), p. 074504. DOI: 10.1103/PhysRevD.95.074504. arXiv: 1608.08900 [hep-lat].

[116] John Bulava and Stefan Schaefer. "Improvement of $N_f = 3$ lattice QCD with Wilson fermions and tree-level improved gauge action." In: *Nuclear Physics B* 874.1 (Sept. 2013), pp. 188–197. DOI: 10.1016/j.nuclphysb.2013.05.019. arXiv: 1304.7093 [hep-lat].

[117] Stefan Schaefer, Rainer Sommer, and Francesco Virotta (ALPHA Collaboration). "Critical slowing down and error analysis in lattice QCD simulations." In: *Nuclear Physics B* 845.1 (Apr. 2011), pp. 93–119. DOI: 10.1016/j.nuclphysb.2010.11.020. arXiv: 1009.5228 [hep-lat].

[118] Martin Lüscher and Stefan Schaefer. "Lattice QCD without topology barriers." In: *Journal of High Energy Physics* 2011.7 (July 2011), p. 36. DOI: 10.1007/JHEP07(2011)036. arXiv: 1105.4749 [hep-lat].

[119] Martin Lüscher and Stefan Schaefer. "Lattice QCD with open boundary conditions and twisted-mass reweighting." In: *Computer Physics Communications* 184.3 (Mar. 2013), pp. 519–528. DOI: 10.1016/j.cpc.2012.10.003. arXiv: 1206.2809 [hep-lat].

[120] Thomas A. DeGrand. "A conditioning technique for matrix inversion for Wilson fermions." In: *Computer Physics Communications* 52.1 (Dec. 1988), pp. 161–164. DOI: 10.1016/0010-4655(88)90180-4.

[121] Thomas A. DeGrand and Carlton DeTar. *Lattice Methods For Quantum Chromodynamics*. Singapore: World Scientific Publishing Company, 2006. DOI: 10.1142/6065.

[122] Luigi Del Debbio *et al.* "Stability of lattice QCD simulations and the thermodynamic limit." In: *Journal of High Energy Physics* 2006.02 (Feb. 2006), p. 011. DOI: 10.1088/1126-6708/2006/02/011. arXiv: hep-lat/0512021.

[123] Filippo Palombi and Martin Lüscher. "Fluctuations and reweighting of the quark determinant on large lattices." In: *Proceedings of The XXVI International Symposium on Lattice Field Theory — PoS(LATTICE 2008)*. Vol. 066. June 2009, p. 049. DOI: 10.22323/1.066.0049. arXiv: 0810.0946 [hep-lat].

[124] Simon Kuberski. "Low-mode deflation for twisted-mass and RHMC reweighting in lattice QCD." In: *Computer Physics Communications* 300 (July 2024), p. 109173. DOI: 10.1016/j.cpc.2024.109173. arXiv: 2306.02385 [hep-lat].

[125] Martin Hasenbusch. "Speeding up the hybrid Monte Carlo algorithm for dynamical fermions." In: *Physics Letters B* 519.1 (Oct. 2001), pp. 177–182. DOI: 10.1016/S0370-2693(01)01102-9. arXiv: hep-lat/0107019.

[126] Naum I. Achieser. *Vorlesungen über Approximationstheorie*. Vol. 2. Mathematische Lehrbücher und Monographien. Abteilung 1, Mathematische Lehrbücher. Berlin: Akademie-Verlag, 1953.

[127] James C. Sexton and Donald H. Weingarten. "Hamiltonian evolution for the hybrid Monte Carlo algorithm." In: *Nuclear Physics B* 380.3 (Aug. 1992), pp. 665–677. DOI: 10.1016/0550-3213(92)90263-B.

[128] I. P. Omelyan, I. M. Mryglod, and R. Folk. "Symplectic analytically integrable decomposition algorithms: classification, derivation, and application to molecular dynamics, quantum and celestial mechanics simulations." In: *Computer Physics Communications* 151.3 (Apr. 2003), pp. 272–314. DOI: 10.1016/S0010-4655(02)00754-3.

[129] Martin Lüscher. "Local coherence and deflation of the low quark modes in lattice QCD." In: *Journal of High Energy Physics* 2007.07 (July 2007), p. 081. DOI: 10.1088/1126-6708/2007/07/081. arXiv: 0706.2298 [hep-lat].

[130] Martin Lüscher. "Deflation acceleration of lattice QCD simulations." In: *Journal of High Energy Physics* 2007.12 (Dec. 2007), p. 011. DOI: 10.1088/1126-6708/2007/12/011. arXiv: 0710.5417 [hep-lat].

[131] Martin Lüscher. "Lattice QCD and the Schwarz alternating procedure." In: *Journal of High Energy Physics* 2003.05 (June 2003), p. 052. DOI: 10.1088/1126-6708/2003/05/052. arXiv: hep-lat/0304007.

[132] Martin Lüscher. "Solution of the Dirac equation in lattice QCD using a domain decomposition method." In: *Computer Physics Communications* 156.3 (2004), pp. 209–220. DOI: 10.1016/S0010-4655(03)00486-7. arXiv: hep-lat/0310048.

[133] Hermann Amandus Schwarz. "Über einen Grenzübergang durch alternierendes Verfahren." In: *Vierteljahresschrift der Naturforschenden Gesellschaft in Zürich* 15.3 (Aug. 1870), pp. 272–286. URL: https://www.ngzh.ch/archiv/1870_15/15_3/15_19.pdf.

[134] Stanley C. Eisenstat, Howard C. Elman, and Martin H. Schultz. "Variational Iterative Methods for Nonsymmetric Systems of Linear Equations." In: *SIAM Journal on Numerical Analysis* 20.2 (Apr. 1983), pp. 345–357. DOI: 10.1137/0720023.

[135] Youcef Saad and Martin H. Schultz. "GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems." In: *SIAM Journal on Scientific and Statistical Computing* 7.3 (July 1986), pp. 856–869. DOI: 10.1137/0907058.

[136] Martin Lüscher. "Properties and uses of the Wilson flow in lattice QCD." In: *Journal of High Energy Physics* 2010.8 (Aug. 2010), p. 71. DOI: 10.1007/JHEP08(2010)071. arXiv: 1006.4518 [hep-lat]. Erratum ibid. [261].

[137] Sinya Aoki *et al.* (FLAG Working Group). "Review of lattice results concerning low-energy particle physics." In: *The European Physical Journal C* 74.9 (Sept. 2014), p. 2890. DOI: 10.1140/epjc/s10052-014-2890-7. arXiv: 1310.8555 [hep-lat].

[138] Yasumichi Aoki *et al.* (Flavour Lattice Averaging Group). "FLAG Review 2021." In: *The European Physical Journal C* 82.10 (Oct. 2022), p. 869. DOI: 10.1140/epjc/s10052-022-10536-1. arXiv: 2111.09849 [hep-lat].

[139] Stefan Sint and Peter Weisz. "Further results on O(a) improved lattice QCD to one-loop order of perturbation theory." In: *Nuclear Physics B* 502.1 (Sept. 1997), pp. 251–268. DOI: 10.1016/S0550-3213(97)00372-6. arXiv: hep-lat/9704001.

[140] Mattia Dalla Brida *et al.* "Heavy Wilson quarks and O(a) improvement: nonperturbative results for b_{g}." In: *Journal of High Energy Physics* 2024.01 (Jan. 2024), p. 188. DOI: 10.1007/JHEP01(2024)188. arXiv: 2401.00216 [hep-lat].

[141] Jürg Gasser and Heinrich Leutwyler. "Chiral perturbation theory: Expansions in the mass of the strange quark." In: *Nuclear Physics B* 250.1 (1985), pp. 465–516. DOI: 10.1016/0550-3213(85)90492-4.

[142] Oliver Bär and Maarten Golterman. "Chiral perturbation theory for gradient flow observables." In: *Physical Review D* 89.3 (Feb. 2014), p. 034505. DOI: 10.1103/PhysRevD.89.034505. arXiv: 1312.4999 [hep-lat]. Erratum ibid. [262].

[143] Piotr Korcyl. "CLS 2+1 flavor simulations." In: *Proceedings of The 32nd International Symposium on Lattice Field Theory — PoS(LATTICE2014)*. Vol. 214. May 2015, p. 029. DOI: 10.22323/1.214.0029. arXiv: 1501.00094 [hep-lat].

[144] Thomas Blum *et al.* (RBC and UKQCD Collaborations). "Domain wall QCD with physical quark masses." In: *Physical Review D* 93.7 (Apr. 2016), p. 074505. DOI: 10.1103/PhysRevD.93.074505. arXiv: 1411.7017 [hep-lat].

[145] Szabolcs Borsányi *et al.* (Budapest-Marseille-Wuppertal collaboration). "High-precision scale setting in lattice QCD." In: *Journal of High Energy Physics* 2012.09 (Sept. 2012), p. 10. DOI: 10.1007/JHEP09(2012)010. arXiv: 1203.4469 [hep-lat].

[146] Gunnar S. Bali *et al.* (RQCD Collaboration). "Lattice simulations with $N_f = 2 + 1$ improved Wilson fermions at a fixed strange quark mass." In: *Physical Review D* 94.7 (Oct. 2016), p. 074501. DOI: 10.1103/PhysRevD.94.074501. arXiv: 1606.09039 [hep-lat].

[147] Franz Gross *et al.* "50 Years of quantum chromodynamics." In: *The European Physical Journal C* 83.12 (Dec. 2023), p. 1125. DOI: 10.1140/epjc/s10052-023-11949-2. arXiv: 2212.11107 [hep-ph].

[148] Volker D. Burkert *et al.* "Colloquium: Gravitational form factors of the proton." In: *Reviews of Modern Physics* 95.4 (Dec. 2023), p. 041002. DOI: 10.1103/RevModPhys.95.041002. arXiv: 2303.08347 [hep-ph].

[149] Maxim V. Polyakov and Peter Schweitzer. "D-term, strong forces in the nucleon, and their applications." Jan. 2018. arXiv: 1801.05858 [hep-ph].

[150] Maxim V. Polyakov and Peter Schweitzer. "Forces inside hadrons: Pressure, surface tension, mechanical radius, and all that." In: *International Journal of Modern Physics A* 33.26 (Sept. 2018), p. 1830025. DOI: 10.1142/S0217751X18300259. arXiv: 1805.06596 [hep-ph].

[151] Daniel C. Hackett, Dimitra A. Pefkou, and Phiala E. Shanahan. "Gravitational form factors of the proton from lattice QCD." In: *Physical Review Letters* 132.25 (June 2024), p. 251904. DOI: 10.1103/PhysRevLett.132.251904. arXiv: 2310.08484 [hep-lat].

[152] G. Peter Lepage. "The Analysis of Algorithms for Lattice Field Theory." In: *From Actions to Answers – Proceedings of the 1989 Theoretical Advanced Study Institute in Elementary Particle Physics*. Ed. by Thomas DeGrand and Doug Toussaint. Singapore: World Scientific, 1990, pp. 97–120. DOI: 10.1142/9789814540988.

[153] Stephan Güsken. "A study of smearing techniques for hadron correlation functions." In: *Nuclear Physics B – Proceedings Supplements* 17 (Sept. 1990), pp. 361–364. DOI: 10.1016/0920-5632(90)90273-W.

[154] Georg von Hippel *et al.* "The shape of covariantly smeared sources in lattice QCD." In: *Journal of High Energy Physics* 2013.9 (Sept. 2013), p. 14. DOI: 10.1007/JHEP09(2013)014. arXiv: 1306.1440 [hep-lat].

[155] M. Albanese *et al.* (APE Collaboration). "Glueball masses and string tension in lattice QCD." In: *Physics Letters B* 192.1 (June 1987), pp. 163–169. DOI: 10.1016/0370-2693(87)91160-9.

[156] Gunnar S. Bali, Sara Collins, and Andreas Schäfer. "Effective noise reduction techniques for disconnected loops in Lattice QCD." In: *Computer Physics Communications* 181.9 (Sept. 2010), pp. 1570–1583. DOI: 10.1016/j.cpc.2010.05.008. arXiv: 0910.3970 [hep-lat].

[157] Thomas Blum, Taku Izubuchi, and Eigo Shintani. "New class of variance-reduction techniques using lattice symmetries." In: *Physical Review D* 88.9 (Nov. 2013), p. 094503. DOI: 10.1103/PhysRevD.88.094503. arXiv: 1208.4349 [hep-lat].

[158] Eigo Shintani *et al.* "Covariant approximation averaging." In: *Physical Review D* 91.11 (June 2015), p. 114511. DOI: 10.1103/PhysRevD.91.114511. arXiv: 1402.0244 [hep-lat].

[159] G. Martinelli and C. T. Sachrajda. "A lattice study of nucleon structure." In: *Nuclear Physics B* 316.2 (Apr. 1989), pp. 355–372. DOI: 10.1016/0550-3213(89)90035-7.

[160] Simon Dinter *et al.* "Precision study of excited state effects in nucleon matrix elements." In: *Physics Letters B* 704.1 (Oct. 2011), pp. 89–93. DOI: 10.1016/j.physletb.2011.09.002. arXiv: 1108.1076 [hep-lat].

[161] Khalil Bitar *et al.* "The QCD finite temperature transition and hybrid Monte Carlo." In: *Nuclear Physics B* 313.2 (Feb. 1989), pp. 348–376. DOI: 10.1016/0550-3213(89)90323-4.

[162] Leonardo Giusti *et al.* "Frequency-splitting estimators of single-propagator traces." In: *The European Physical Journal C* 79.7 (July 2019), p. 586. DOI: 10.1140/epjc/s10052-019-7049-0. arXiv: 1903.10447 [hep-lat].

[163] Shao-Jing Dong and Keh-Fei Liu. "Stochastic estimation with Z_2 noise." In: *Physics Letters B* 328.1 (May 1994), pp. 130–136. DOI: 10.1016/0370-2693(94)90440-5. arXiv: hep-lat/9308015.

[164] Marco Cè *et al.* "The hadronic running of the electromagnetic coupling and the electroweak mixing angle from lattice QCD." In: *Journal of High Energy Physics* 2022.8 (Aug. 2022), p. 220. DOI: 10.1007/JHEP08(2022)220. arXiv: 2203.08676 [hep-lat].

[165] Philippe Boucaud *et al.* (ETM Collaboration). "Dynamical twisted mass fermions with light quarks: simulation and analysis details." In: *Computer Physics Communications* 179.10 (Nov. 2008), pp. 695–715. DOI: 10.1016/j.cpc.2008.06.013. arXiv: 0803.0224 [hep-lat].

[166] Craige McNeile and Christopher Michael (UKQCD Collaboration). "Decay width of light quark hybrid meson from the lattice." In: *Physical Review D* 73.7 (Apr. 2006), p. 074506. DOI: 10.1103/PhysRevD.73.074506. arXiv: hep-lat/0603007.

[167] Christopher Thron *et al.* "Padé-Z_2 estimator of determinants." In: *Physical Review D* 57.3 (Feb. 1998), pp. 1642–1653. DOI: 10.1103/PhysRevD.57.1642. arXiv: hep-lat/9707001.

[168] Vera Gülpers, Georg von Hippel, and Hartmut Wittig. "Scalar pion form factor in two-flavor lattice QCD." In: *Physical Review D* 89.9 (May 2014), p. 094503. DOI: 10.1103/PhysRevD.89.094503. arXiv: 1309.2104 [hep-lat].

[169] Andreas Stathopoulos, Jesse Laeuchli, and Kostas Orginos. "Hierarchical Probing for Estimating the Trace of the Matrix Inverse on Toroidal Lattices." In: *SIAM Journal on Scientific Computing* 35.5 (Oct. 2013), S299–S322. DOI: 10.1137/120881452. arXiv: 1302.4018 [hep-lat].

[170] Colin Morningstar *et al.* "Improved stochastic estimation of quark propagation with Laplacian Heaviside smearing in lattice QCD." In: *Physical Review D* 83.11 (June 2011), p. 114505. DOI: 10.1103/PhysRevD.83.114505. arXiv: 1104.3870 [hep-lat].

[171] Ulli Wolff (ALPHA Collaboration). "Monte Carlo errors with less errors." In: *Computer Physics Communications* 156.2 (Jan. 2004), pp. 143–153. DOI: 10.1016/S0010-4655(03)00467-3. arXiv: hep-lat/0306017. Erratum ibid. [263].

[172] Bradley Efron. "Bootstrap Methods: Another Look at the Jackknife." In: *The Annals of Statistics* 7.1 (Jan. 1979), pp. 1–26. DOI: 10.1214/aos/1176344552.

[173] Bradley Efron and Robert J. Tibshirani. *An Introduction to the Bootstrap*. Vol. 57. Monographs on Statistics and Applied Probability. New York: Chapman & Hall, 1993. DOI: 10.1201/9780429246593.

[174] M. H. Quenouille. "Notes on Bias in Estimation." In: *Biometrika* 43.3/4 (Dec. 1956), pp. 353–360. DOI: 10.2307/2332914.

[175] John W. Tukey. "Bias and Confidence in Not-quite Large Samples." In: *The Annals of Mathematical Statistics* 29.2 (June 1958), p. 614. DOI: 10.1214/aoms/1177706647.

[176] Pauli Virtanen *et al.* "SciPy 1.0: fundamental algorithms for scientific computing in Python." In: *Nature Methods* 17 (Feb. 2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[177] Kenneth Levenberg. "A method for the solution of certain non-linear problems in least squares." In: *Quarterly of Applied Mathematics* 2.2 (1944), pp. 164–168. DOI: 10.1090/qam/10666.

[178] Donald W. Marquardt. "An Algorithm for Least-Squares Estimation of Nonlinear Parameters." In: *Journal of the Society for Industrial and Applied Mathematics* 11.2 (June 1963), pp. 431–441. DOI: 10.1137/0111030.

[179] Mary Ann Branch, Thomas F. Coleman, and Yuying Li. "A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems." In: *SIAM Journal on Scientific Computing* 21.1 (Jan. 1999), pp. 1–23. DOI: 10.1137/S1064827595289108.

[180] Mattia Bruno and Rainer Sommer. "On fits to correlated and auto-correlated data." In: *Computer Physics Communications* 285 (Apr. 2023), p. 108643. DOI: 10.1016/j.cpc.2022.108643. arXiv: 2209.14188 [hep-lat].

[181] Gene H. Golub and V. Pereyra. "The Differentiation of Pseudo-Inverses and Nonlinear Least Squares Problems Whose Variables Separate." In: *SIAM Journal on Numerical Analysis* 10.2 (Apr. 1973), pp. 413–432. DOI: 10.1137/0710036.

[182] G. Peter Lepage *et al.* "Constrained curve fitting." In: *Nuclear Physics B – Proceedings Supplements* 106–107 (Mar. 2002), pp. 12–20. DOI: 10.1016/S0920-5632(01)01638-3. arXiv: hep-lat/0110175.

[183] Andria Agadjanov *et al.* "Nucleon Sigma Terms with $N_f = 2 + 1$ Flavors of $\mathcal{O}(a)$-Improved Wilson Fermions." In: *Physical Review Letters* 131.26 (Dec. 2023), p. 261902. DOI: 10.1103/PhysRevLett.131.261902. arXiv: 2303.08741 [hep-lat].

[184] Gunnar S. Bali *et al.* (RQCD Collaboration). "Octet baryon isovector charges from $N_f = 2 + 1$ lattice QCD." In: *Physical Review D* 108.3 (Aug. 2023), p. 034512. DOI: 10.1103/PhysRevD.108.034512. arXiv: 2305.04717 [hep-lat].

[185] Hirotugu Akaike. "Information Theory and an Extension of the Maximum Likelihood Principle." In: *Proceedings of the 2nd International Symposium on Information Theory*. Ed. by Boris Nikolaevich Petrov and Frigyes Csaki. Budapest: Akadémiai Kiadó, 1973, pp. 267–281.

[186] Hirotugu Akaike. "A new look at the statistical model identification." In: *IEEE Transactions on Automatic Control* 19.6 (1974), pp. 716–723. DOI: 10.1109/TAC.1974.1100705.

[187] Tomasz Korzec *et al.* (European Twisted Mass Collaboration). "Nucleon form factors with dynamical twisted mass fermions." In: *Proceedings of The XXVI International Symposium on Lattice Field Theory — PoS(LATTICE 2008)*. Vol. 066. June 2009, p. 139. DOI: 10.22323/1.066.0139. arXiv: 0811.0724 [hep-lat].

[188] Dalibor Djukanovic *et al.* "Strange Electromagnetic Form Factors of the Nucleon with $N_f = 2 + 1$ $\mathcal{O}(a)$-Improved Wilson Fermions." In: *Physical Review Letters* 123.21 (Nov. 2019), p. 212001. DOI: 10.1103/PhysRevLett.123.212001. arXiv: 1903.12566 [hep-lat].

[189] Konstantin Ottnad. "Excited states in nucleon structure calculations." In: *The European Physical Journal A* 57.2 (Feb. 2021), p. 50. DOI: 10.1140/epja/s10050-021-00355-5. arXiv: 2011.12471 [hep-lat].

[190] Alexander M. Segner, Andreas Risch, and Hartmut Wittig. "Precision Determination of Baryon Masses including Isospin-breaking." In: *Proceedings of The 40th International Symposium on Lattice Field Theory — PoS(LATTICE2023)*. Vol. 453. Dec. 2023, p. 044. DOI: 10.22323/1.453.0044. arXiv: 2312.09065 [hep-lat].

[191] L. Maiani *et al.* "Scalar densities and baryon mass differences in lattice QCD with Wilson fermions." In: *Nuclear Physics B* 293 (1987), pp. 420–444. DOI: 10.1016/0550-3213(87)90078-2.

[192] Takumi Doi *et al.* (χQCD Collaboration). "Nucleon strangeness form factors from $N_f = 2 + 1$ clover fermion lattice QCD." In: *Physical Review D* 80.9 (Nov. 2009), p. 094503. DOI: 10.1103/PhysRevD.80.094503. arXiv: 0903.3232 [hep-ph].

[193] Konstantin Ottnad *et al.* "Improved analysis of nucleon isovector charges and twist-2 matrix elements on CLS $N_f = 2+1$ ensembles." In: *Proceedings of The 38th International Symposium on Lattice Field Theory — PoS(LATTICE2021)*. Vol. 396. July 2022, p. 343. DOI: 10.22323/1.396.0343. arXiv: 2110.10500 [hep-lat].

[194] Konstantin Ottnad *et al.* "Mass and isovector matrix elements of the nucleon at zero-momentum transfer." In: *Proceedings of The 39th International Symposium on Lattice Field Theory — PoS(LATTICE2022)*. Vol. 430. Apr. 2023, p. 117. DOI: 10.22323/1.430.0117. arXiv: 2212.09940 [hep-lat].

[195] Dalibor Djukanovic *et al.* "Improved analysis of isovector nucleon matrix elements with $N_f = 2 + 1$ flavors of $\mathcal{O}(a)$ improved Wilson fermions." In: *Physical Review D* 109.7 (Apr. 2024), p. 074507. DOI: 10.1103/PhysRevD.109.074507. arXiv: 2402.03024 [hep-lat].

[196] Dalibor Djukanovic *et al.* "Isovector axial form factor of the nucleon from lattice QCD." In: *Physical Review D* 106.7 (Oct. 2022), p. 074503. DOI: 10.1103/PhysRevD.106.074503. arXiv: 2207.03440 [hep-lat].

[197] Thomas Bauer, Jan C. Bernauer, and Stefan Scherer. "Electromagnetic form factors of the nucleon in effective field theory." In: *Physical Review C* 86.6 (Dec. 2012), p. 065206. DOI: 10.1103/PhysRevC.86.065206. arXiv: 1209.3872 [nucl-th].

[198] Bastian Kubis and Ulf-G. Meißner. "Low-energy analysis of the nucleon electromagnetic form factors." In: *Nuclear Physics A* 679.3 (Jan. 2001), pp. 698–734. DOI: 10.1016/S0375-9474(00)00378-X. arXiv: hep-ph/0007056.

[199] Matthias R. Schindler, Jambul Gegelia, and Stefan Scherer. "Electromagnetic form factors of the nucleon in chiral perturbation theory including vector mesons." In: *The European Physical Journal A* 26.1 (Oct. 2005), pp. 1–5. DOI: 10.1140/epja/i2005-10145-8. arXiv: nucl-th/0509005.

[200] Ken Kawarabayashi and Mahiko Suzuki. "Partially Conserved Axial-Vector Current and the Decays of Vector Mesons." In: *Physical Review Letters* 16.6 (Feb. 1966), pp. 255–257. DOI: 10.1103/PhysRevLett.16.255. Erratum ibid. [264].

[201] Riazuddin and Fayyazuddin. "Algebra of Current Components and Decay Widths of ρ and K^* Mesons." In: *Physical Review* 147.4 (July 1966), pp. 1071–1073. DOI: 10.1103/PhysRev.147.1071.

[202] Dalibor Djukanovic *et al.* "Universality of the ρ Meson Coupling in Effective Field Theory." In: *Physical Review Letters* 93.12 (Sept. 2004), p. 122002. DOI: 10.1103/PhysRevLett.93.122002. arXiv: hep-ph/0407239.

[203] T. Hahn and M. Pérez-Victoria. "Automated one-loop calculations in four and D dimensions." In: *Computer Physics Communications* 118.2 (May 1999), pp. 153–165. DOI: 10.1016/S0010-4655(98)00173-8. arXiv: hep-ph/9807565.

[204] G. J. van Oldenborgh and J. A. M. Vermaseren. "New algorithms for one-loop integrals." In: *Zeitschrift für Physik C – Particles and Fields* 46.3 (Sept. 1990), pp. 425–437. DOI: 10.1007/BF01621031.

[205] Torsten Bauer. "Elektromagnetische Formfaktoren des Nukleons in effektiver Feldtheorie." Diploma Thesis. Mainz: Johannes Gutenberg-Universität Mainz, Apr. 2009. URL: https://portal.kph.uni-mainz.de/T/pub/diploma/Dipl_Th_Bauer.pdf.

[206] Silas R. Beane. "Nucleon masses and magnetic moments in a finite volume." In: *Physical Review D* 70.3 (Aug. 2004), p. 034507. DOI: 10.1103/PhysRevD.70.034507. arXiv: hep-lat/0403015.

[207] L. I. Schiff. "Interpretation of Electron Scattering Experiments." In: *Physical Review* 92.4 (Nov. 1953), pp. 988–993. DOI: 10.1103/PhysRev.92.988. Erratum ibid. [265].

[208] Charles F. Perdrisat, Vina Punjabi, and Marc Vanderhaeghen. "Nucleon electromagnetic form factors." In: *Progress in Particle and Nuclear Physics* 59.2 (Oct. 2007), pp. 694–764. DOI: 10.1016/j.ppnp.2007.05.001. arXiv: hep-ph/0612014.

[209] Richard J. Hill and Gil Paz. "Model-independent extraction of the proton charge radius from electron scattering." In: *Physical Review D* 82.11 (Dec. 2010), p. 113005. DOI: 10.1103/PhysRevD.82.113005. arXiv: 1008.4619 [hep-ph].

[210] P. Federbush, M. L. Goldberger, and S. B. Treiman. "Electromagnetic Structure of the Nucleon." In: *Physical Review* 112.2 (Oct. 1958), pp. 642–665. DOI: 10.1103/PhysRev.112.642.

[211] Véronique Bernard *et al.* "The form factors of the nucleon at small momentum transfer." In: *Nuclear Physics A* 635.1 (May 1998), pp. 121–145. DOI: 10.1016/S0375-9474(98)00175-4. arXiv: hep-ph/9801297. Erratum ibid. [266].

[212] Vina Punjabi *et al.* "The structure of the nucleon: Elastic electromagnetic form factors." In: *The European Physical Journal A* 51.7 (July 2015), p. 79. DOI: 10.1140/epja/i2015-15079-x. arXiv: 1503.01452 [nucl-ex].

[213] M. Eric Christy *et al.* "Form Factors and Two-Photon Exchange in High-Energy Elastic Electron-Proton Scattering." In: *Physical Review Letters* 128.10 (Mar. 2022), p. 102002. DOI: 10.1103/PhysRevLett.128.102002. arXiv: 2103.01842 [nucl-ex].

[214] Krzysztof Pachucki *et al.* "Comprehensive theory of the Lamb shift in light muonic atoms." In: *Reviews of Modern Physics* 96.1 (Jan. 2024), p. 015001. DOI: 10.1103/RevModPhys.96.015001. arXiv: 2212.13782 [physics.atom-ph].

[215] James Lewis Friar. "Nuclear finite-size effects in light muonic atoms." In: *Annals of Physics* 122.1 (Sept. 1979), pp. 151–196. DOI: 10.1016/0003-4916(79)90300-2.

[216] Krzysztof Pachucki. "Theory of the Lamb shift in muonic hydrogen." In: *Physical Review A* 53.4 (Apr. 1996), pp. 2092–2100. DOI: 10.1103/PhysRevA.53.2092.

[217] Krzysztof Pachucki. "Proton structure effects in muonic hydrogen." In: *Physical Review A* 60.5 (Nov. 1999), pp. 3593–3598. DOI: 10.1103/PhysRevA.60.3593. arXiv: physics/9906002.

[218] Alvaro De Rújula. "QED is not endangered by the proton's size." In: *Physics Letters B* 693.5 (Oct. 2010), pp. 555–558. DOI: 10.1016/j.physletb.2010.08.074. arXiv: 1008.3861 [hep-ph].

[219] Michael O. Distler, Jan C. Bernauer, and Thomas Walcher. "The RMS charge radius of the proton and Zemach moments." In: *Physics Letters B* 696.4 (Feb. 2011), pp. 343–347. DOI: 10.1016/j.physletb.2010.12.067. arXiv: 1011.1861 [nucl-th].

[220] Franziska Hagelstein and Vladimir Pascalutsa. "Breakdown of the expansion of finite-size corrections to the hydrogen Lamb shift in moments of charge distribution." In: *Physical Review A* 91.4 (Apr. 2015), p. 040502. DOI: 10.1103/PhysRevA.91.040502. arXiv: 1502.03721 [hep-ph].

[221] Masaharu Sato *et al.* "Laser spectroscopy of the hyperfine splitting energy in the ground state of muonic hydrogen." In: *20th International Conference on Particles and Nuclei*. Sept. 2014, pp. 460–463. DOI: 10.3204/DESY-PROC-2014-04/67.

[222] C. Pizzolotto *et al.* "The FAMU experiment: muonic hydrogen high precision spectroscopy studies." In: *The European Physical Journal A* 56.7 (July 2020), p. 185. DOI: 10.1140/epja/s10050-020-00195-9.

[223] C. Pizzolotto *et al.* "Measurement of the muon transfer rate from muonic hydrogen to oxygen in the range 70-336 K." In: *Physics Letters A* 403 (July 2021), p. 127401. DOI: 10.1016/j.physleta.2021.127401. arXiv: 2105.06701 [physics.atom-ph].

[224] Pedro Amaro *et al.* "Laser excitation of the 1s-hyperfine transition in muonic hydrogen." In: *SciPost Physics* 13.2 (Aug. 2022), p. 020. DOI: 10.21468/SciPostPhys.13.2.020. arXiv: 2112.00138 [physics.atom-ph].

[225] A. C. Zemach. "Proton Structure and the Hyperfine Shift in Hydrogen." In: *Physical Review* 104.6 (Dec. 1956), pp. 1771–1781. DOI: 10.1103/PhysRev.104.1771.

[226] Franziska Hagelstein, Vadim Lensky, and Vladimir Pascalutsa. "Chiral perturbation theory of the hyperfine splitting in (muonic) hydrogen." In: *The European Physical Journal C* 83.8 (Aug. 2023), p. 762. DOI: 10.1140/epjc/s10052-023-11866-4. arXiv: 2305.09633 [nucl-th].

[227] G. Peter Lepage and Stanley J. Brodsky. "Exclusive processes in perturbative quantum chromodynamics." In: *Physical Review D* 22.9 (Nov. 1980), pp. 2157–2198. DOI: 10.1103/PhysRevD.22.2157.

[228] Solomon Kullback and R. A. Leibler. "On Information and Sufficiency." In: *Annals of Mathematical Statistics* 22.1 (Mar. 1951), pp. 79–86. DOI: 10.1214/aoms/1177729694.

[229] Ethan T. Neil and Jacob W. Sitison. "Improved information criteria for Bayesian model averaging in lattice field theory." In: *Physical Review D* 109.1 (Jan. 2024), p. 014510. DOI: 10.1103/PhysRevD.109.014510. arXiv: 2208.14983 [stat.ME].

[230] K. Takeuchi. "Distribution of information statistics and criteria for adequacy of models." In: *Mathematical Science* 153 (1976), pp. 12–18.

[231] Halbert White. "Maximum Likelihood Estimation of Misspecified Models." In: *Econometrica* 50.1 (Jan. 1982), pp. 1–25. DOI: 10.2307/1912526.

[232] Ritei Shibata. "An optimal selection of regression variables." In: *Biometrika* 68.1 (Apr. 1981), pp. 45–54. DOI: 10.1093/biomet/68.1.45.

[233] William I. Jay and Ethan T. Neil. "Bayesian model averaging for analysis of lattice field theory results." In: *Physical Review D* 103.11 (June 2021), p. 114502. DOI: 10.1103/PhysRevD.103.114502. arXiv: 2008.01069 [stat.ME].

[234] Kenneth P. Burnham and David R. Anderson. "Multimodel inference: Understanding AIC and BIC in model selection." In: *Sociological Methods & Research* 33.2 (Nov. 2004), pp. 261–304. DOI: 10.1177/0049124104268644.

[235] Szabolcs Borsányi *et al.* "Ab initio calculation of the neutron-proton mass difference." In: *Science* 347.6229 (Mar. 2015), pp. 1452–1455. DOI: 10.1126/science.1257050. arXiv: 1406.4088 [hep-lat].

[236] Nuria Carrasco *et al.* (The European Twisted Mass Collaboration). "Up, down, strange and charm quark masses with $N_f = 2 + 1 + 1$ twisted mass lattice QCD." In: *Nuclear Physics B* 887 (Oct. 2014), pp. 19–68. DOI: 10.1016/j.nuclphysb.2014.07.025. arXiv: 1403.4504 [hep-lat].

[237] Zhihong Ye *et al.* "Proton and neutron electromagnetic form factors and uncertainties." In: *Physics Letters B* 777 (Feb. 2018), pp. 8–15. DOI: 10.1016/j.physletb.2017.11.023. arXiv: 1707.09063 [nucl-ex].

[238] Ryutaro Tsuji *et al.* (PACS Collaboration). "Nucleon form factors in $N_f = 2 + 1$ lattice QCD at the physical point: Finite lattice spacing effect on the root-mean-square radii." In: *Physical Review D* 109.9 (May 2024), p. 094505. DOI: 10.1103/PhysRevD.109.094505. arXiv: 2311.10345 [hep-lat].

[239] Arseniy A. Filin *et al.* "High-accuracy calculation of the deuteron charge and quadrupole form factors in chiral effective field theory." In: *Physical Review C* 103.2 (Feb. 2021), p. 024313. DOI: 10.1103/PhysRevC.103.024313. arXiv: 2009.08911 [nucl-th].

[240] U. D. Jentschura *et al.* "Hydrogen-deuterium isotope shift: From the $1S - 2S$-transition frequency to the proton-deuteron charge-radius difference." In: *Physical Review A* 83.4 (Apr. 2011), p. 042505. DOI: 10.1103/PhysRevA.83.042505.

[241] Helmut Hellwig *et al.* "Measurement of the Unperturbed Hydrogen Hyperfine Transition Frequency." In: *IEEE Transactions on Instrumentation and Measurement* 19.4 (Nov. 1970), pp. 200–209. DOI: 10.1109/TIM.1970.4313902.

[242] Rudolf N. Faustov and Aleksey P. Martynenko. "Proton-polarizability contribution to the hyperfine splitting in hydrogen." In: *Physics of Atomic Nuclei* 65 (Feb. 2002), pp. 265–270. DOI: 10.1134/1.1451940. arXiv: hep-ph/0007044.

[243] E. V. Cherednikova, Rudolf N. Faustov, and Aleksey P. Martynenko. "Proton polarizability contribution to the hyperfine splitting in muonic hydrogen." In: *Nuclear Physics A* 703.1 (May 2002), pp. 365–377. DOI: 10.1016/S0375-9474(01)01461-0. arXiv: hep-ph/0108044.

[244] Carl E. Carlson, Vahagn Nazaryan, and Keith Griffioen. "Proton-structure corrections to hyperfine splitting in muonic hydrogen." In: *Physical Review A* 83.4 (2011), p. 042509. DOI: 10.1103/PhysRevA.83.042509. arXiv: 1101.3239 [physics.atom-ph].

[245] A. V. Volotka *et al.* "Zemach and magnetic radius of the proton from the hyperfine splitting in hydrogen." In: *The European Physical Journal D* 33.1 (Feb. 2005), pp. 23–27. DOI: 10.1140/epjd/e2005-00025-9. arXiv: physics/0405118.

[246] Kaushik Borah *et al.* "Parametrization and applications of the low-Q^2 nucleon vector form factors." In: *Physical Review D* 102.7 (Oct. 2020), p. 074012. DOI: 10.1103/PhysRevD.102.074012. arXiv: 2003.13640 [hep-ph].

[247] James Lewis Friar and Ingo Sick. "Comment on 'Constraints on Proton Structure from Precision Atomic-Physics Measurements'." In: *Physical Review Letters* 95.4 (July 2005), p. 049101. DOI: 10.1103/PhysRevLett.95.049101. arXiv: nucl-th/0503020.

[248] Clara Peset and Antonio Pineda. "The two-photon exchange contribution to muonic hydrogen from chiral perturbation theory." In: *Nuclear Physics B* 887 (Oct. 2014), pp. 69–111. DOI: 10.1016/j.nuclphysb.2014.07.027. arXiv: 1406.4524 [hep-ph].

[249] Hans-Werner Hammer and Ulf-G. Meißner. "The proton radius: from a puzzle to precision." In: *Science Bulletin* 65.4 (Feb. 2020), pp. 257–258. DOI: 10.1016/j.scib.2019.12.012. arXiv: 1912.03881 [hep-ph].

[250] Eite Tiesinga *et al.* "CODATA recommended values of the fundamental physical constants: 2018." In: *Review of Modern Physics* 93.2 (June 2021), p. 025010. DOI: 10.1103/RevModPhys.93.025010.

[251] Lorenzo Barca, Gunnar Bali, and Sara Collins. "Toward N to $N\pi$ matrix elements from lattice QCD." In: *Physical Review D* 107.5 (Mar. 2023), p. L051505. DOI: 10.1103/PhysRevD.107.L051505. arXiv: 2211.12278 [hep-lat].

[252] Lorenzo Barca, Gunnar Bali, and Sara Collins. "Progress on nucleon transition matrix elements with a lattice QCD variational analysis." In: *Proceedings of European network for Particle physics, Lattice field theory and Extreme computing — PoS(EuroPLEx2023)*. Vol. 451. June 2024, p. 002. DOI: 10.22323/1.451.0002. arXiv: 2405.20875 [hep-lat].

[253] Ko Abe *et al.* (Hyper-Kamiokande Proto-Collaboration). "Physics potential of a long-baseline neutrino oscillation experiment using a J-PARC neutrino beam and Hyper-Kamiokande." In: *Progress of Theoretical and Experimental Physics* 2015.5 (May 2015), p. 053C02. DOI: 10.1093/ptep/ptv061. arXiv: 1502.05199 [hep-ex].

[254] R. Acciarri *et al.* (DUNE Collaboration). "Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE): Conceptual Design Report, Volume 2: The Physics Program for DUNE at LBNF." Dec. 2015. arXiv: 1512.06148 [physics.ins-det].

[255] Martin Hoferichter, Javier Menéndez, and Achim Schwenk. "Coherent elastic neutrino-nucleus scattering: EFT analysis and nuclear responses." In: *Physical Review D* 102.7 (Oct. 2020), p. 074018. DOI: 10.1103/PhysRevD.102.074018. arXiv: 2007.08529 [hep-ph].

[256] Daniel Becker, S. Baunack, and Frank E. Maas. "P2—a new measurement of the weak charge of the proton." In: *Hyperfine Interactions*. Ed. by H. W. Wilschut and E. Pallante. Vol. 214. 1–3. Jan. 2013, pp. 141–148. DOI: 10.1007/s10751-013-0775-z.

[257] Darko Androić *et al.* (The Jefferson Lab Qweak Collaboration). "Precision measurement of the weak charge of the proton." In: *Nature* 557 (May 2018), pp. 207–211. DOI: 10.1038/s41586-018-0096-0. arXiv: 1905.08283 [nucl-ex].

[258] Eigo Shintani *et al.* (PACS Collaboration). "Erratum: Nucleon form factors and root-mean-square radii on a (10.8 fm)^4 lattice at the physical point [Phys. Rev. D 99, 014510 (2019)]." In: *Physical Review D* 102.1 (July 2020), p. 019902. DOI: 10.1103/PhysRevD.102.019902.

[259] Martin Lüscher and Peter Weisz. "On-shell improved lattice gauge theories." In: *Communications in Mathematical Physics* 98.3 (Sept. 1985), p. 433. DOI: 10.1007/BF01205792. Erratum.

[260] Alan M. Ferrenberg and Robert H. Swendsen. "New Monte Carlo technique for studying phase transitions." In: *Physical Review Letters* 63.15 (Oct. 1989), p. 1658. DOI: 10.1103/PhysRevLett.63.1658.2. Erratum.

[261] Martin Lüscher. "Erratum: Properties and uses of the Wilson flow in lattice QCD." In: *Journal of High Energy Physics* 2014.3 (Mar. 2014), p. 92. DOI: 10.1007/JHEP03(2014)092.

[262] Oliver Bär and Maarten Golterman. "Erratum: Chiral perturbation theory for gradient flow observables [Phys. Rev. D 89, 034505 (2014)]." In: *Physical Review D* 89.9 (May 2014), p. 099905. DOI: 10.1103/PhysRevD.89.099905.

[263] Ulli Wolff (ALPHA Collaboration). "Erratum to 'Monte Carlo errors with less errors' [Comput. Phys. Comm. 156 (2004) 143–153]." In: *Computer Physics Communications* 176.5 (Mar. 2007), p. 383. DOI: 10.1016/j.cpc.2006.12.001.

[264] Ken Kawarabayashi and Mahiko Suzuki. "Partially Conserved Axial-Vector Current and the Decays of Vector Mesons." In: *Physical Review Letters* 16.9 (Feb. 1966), p. 384. DOI: 10.1103/PhysRevLett.16.384.3. Erratum.

[265] L. I. Schiff. "Interpretation of Electron Scattering Experiments." In: *Physical Review* 93.6 (Mar. 1954), p. 1434. DOI: 10.1103/PhysRev.93.1434. Erratum.

[266] Véronique Bernard *et al.* "Erratum to 'The form factors of the nucleon at small momentum transfer'." In: *Nuclear Physics A* 642.3 (Nov. 1998), p. 563. DOI: 10.1016/S0375-9474(98)00566-1.

List of Figures

List of Tables