HPC-LEAP
EUROPEAN JOINT DOCTORATES

H2020 Framework Programme for Research and Innovation

University of Cyprus

BERGISCHE UNIVERSITÄT WUPPERTAL

Simulating maximally twisted fermions at the physical point with multigrid methods

Dissertation Submitted to the University of Cyprus in partial fulfillment of the requirements for the degree of Doctor of Philosophy and to the University of Wuppertal to obtain the academic degree of Dr. rer. nat.

Simone BACCHIO

May 2019

The PhD thesis can be quoted as follows:

Doctoral candidate:

Simone BACCHIO

Doctoral Thesis Title:

Simulating maximally twisted fermions
at the physical point with multigrid methods

The present Doctoral Dissertation was submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Department of Physics at the University of Cyprus and for the degree of Doctor in Rerum Naturalium at the Department of Mathematics and Natural Sciences of the University of Wuppertal. It was approved on the 31st of May 2019 by the members of the Examination Committee.

Examination Committee:

Prof. Dr. Constantia ALEXANDROU
University of Cyprus, Supervisor

Prof. Dr. Andreas FROMMER
University of Wuppertal, Supervisor

Prof. Dr. Luigi DEL DEBBIO
University of Edinburgh

Prof. Dr. Francesco KNECHTLI
University of Wuppertal

Prof. Dr. Haralambos PANAGOPOULOS
University of Cyprus

Prof. Dr. Mike PEARDON
Trinity College Dublin

Declaration of Doctoral Candidate

The present doctoral dissertation was submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Cyprus and for the degree of Doctor in Rerum Naturalium at the University of Wuppertal. It is a product of original work of my own, unless otherwise mentioned through references, notes, or any other statements.

Simone Bacchio
May 2019

Abstract

We develop multigrid methods for simulating lattice QCD with physical values of the quark masses within the twisted mass fermion formulation. We employ the developed multigrid method both for the calculation of the quark propagators needed for extracting the hadronic matrix elements and for accelerating the generation of gauge field ensembles. For the computation of the quark propagators we improve the performance by two orders of magnitudes as compared to conjugate gradient enabling to perform the analysis of key nucleon observables at physical values of the light quark mass. For the generation of gauge field ensembles with two degenerate flavors of the light quarks (N_f=2) an order of magnitude speedup is achieved. Extension of the multigrid approach is carried out to include in the simulation the dynamical strange and charm quarks. To accomplish this one needs the calculation of the square root of the non-degenerate twisted mass operator. We solve the square root with an optimal rational approximation and employ multigrid methods in the solution of the shifted linear system. In such a way we accelerate simulations with N_f=2+1+1 flavors of fermions all tuned at their physical value by an order of one magnitude. These methods are used for the production of four ensembles, two with N_f=2 and two with N_f=2+1+1, which are state-of-the-art worldwide. These ensembles are shared by all members of the Extended Twisted Mass Collaboration (ETMC) and are being used them for obtaining quantitative description of hadron properties including observables that can probe new physics beyond the standard model. In this thesis we focus on physical results on low-lying hadron masses, meson decay constants, and pion and nucleon electromagnetic form factors.[1]

Περίληψη

Αναπτύξαμε νέες μεθόδους για προσομοίωση πλέγματος ΚΧΔ με φυσικές τιμές των μαζών κουάρκς εντός του φορμαλισμού συνεστραμμένης μάζας φερμιονίων. Χρησιμοποιούμε τον αλγόριθμο μυλτιγριδ τόσο για τον υπολογισμό των διαδοτών κουάρκς που χρειάζονται για την εξαγωγή στοιχείων αδρονικής μήτρας καθώς και για την επιτάχυνση των προσομοιώσεων των γκλουανικών πεδίων. Για τον υπολογισμό των διαδοτών κουάρκ βελτιώνουμε την απόδοση κατά δύο τάξεις μεγέθους σε σύγκριση με τη συζευγμένη κλίση, επιτρέποντας την μελέτη των ιδιοτήτων των νουκλεονίων για δύο εκφυλισμένες γεύσεις των ελαφρών κουάρκ (N_f=2) με φυσική τιμή της μάζας τους. Για την προσομοίωση επιτεύχθηκε ταχύτητα επιτάχυνσης μιας τάξης μεγέθους. Επεκτείναμε τον αλγόριθμο μυλτιγριδ για να συμπεριλάβουμε και την προσομοίωση των παράξενων και των γοητευτικών κουάρκς όπου κάποιος χρειάζεται το υπολογισμό της τετραγωνικής ρίζας του μη εκφυλισμένου πίνακα Διρας. Με αυτό τον τρόπο επιταχύνουμε προσομοιώσεις με N_f=2+1+1 γεύσεις φερμιονίων όταν όλα έχουν την φυσική τους μάζα. Αυτές οι μέθοδοι χρησιμοποιούνται για την παραγωγή τεσσάρων συνόλων προσομοιώσεων που συγκαταλέγονται στα καλύτερα διεθνώς. Χρησιμοποιούμε αυτές τις προσομοιώσεις για την ποσοτική περιγραφή της δομής των αδρονίων και των ποσοτήτων που μπορούν να ανιχνεύσουν νέα φυσική πέρα από το μοντέλο πρότυπο. Φυσικά αποτελέσματα που παρουσιάζονται σε αυτή τη διατριβή περιλαμβάνουν μάζες αδρονίων, σταθερές αποσύνθεσης μεσονίων και την ηλεκτρομαγνητική μορφή πιονίου και νουκλεονίου. [1]

Zusammenfassung

In dieser Thesis werden Mehrgitterverfahren zur Simulation von Gittereichtherien mit physikalischen Werten der Quark-Massen innerhalb der Twisted-Mass-Fermion-Formulierung entwickelt. Diese entwickelten Mehrgitterverfahren werdenowohl für die Berechnung der Quark-Propagatoren, die für die Extraktion von hadronischen Matrixelementen benötigt werden, als auch für die Beschleunigung der Simulation von Eichfeld-Ensembles mit zwei degenerierten Light-Quarks (N_f=2) verwendet. Für die Berechnung der Quark-Propagatoren wird die erforderliche Lösungszeit um zwei Größenordnungen im Vergleich zur Conjugated Gradient Methode verbessert. Dies ermöglicht die Statistik für die Analyse wichtiger Nukleon-Observablen entscheident zu erhöhen. In der Simulation wurde die Lösungszeit um eine Größenordnung für Quarkmassen am physikalischen Punkt verringert. Des Weiteren wurde das Mehrgitterverfahren auf den nicht-degenerierten Twisted-Mass-Operators erweitert, was die Anwendung des Mehrgitterverfahren für die Simulation von dynamischer Strange und Charm Quarks ermöglicht. Dies erfordert die Berechnung der Quadratwurzel des Operators, welche mit einer optimalen rationalen Approximation und dem verwenden Mehrgitterverfahren zur Lösung des verschobenen Linearsystems berechnet wird. Auf diese Weise wurde der Rechenaufwand der Simulationen mit N_f=2+1+1 Fermionen deutlich reduziert, was Simulation mit Quarks ermöglicht, die alle auf ihren physikalischen Wert abgestimmt sind. Diese Ensembels, hier zwei mit N_f=2 und zwei mit N_f=2+1+1, sind weltweit auf dem neuesten Stand. Die Messungen auf diesen Ensembels werden verwenden für die quantitativen Beschreibung der Hadronenstruktur und können Hinweise auf neue Physik jenseits des Standardmodell geben. Die präsentierten physikalischen Ergebnisse beinhalten Hadronenmassen, Mesonzerfallskonstanten sowie elektromagnetische Formfaktoren von Pione und des Nukleon.[1]

---

[1]These zwischen der Universität von Zypern und der Bergischen Universität Wuppertal, welche ein Teil der Promotionsvoraussetzung für den Doppelabschluss im Rahmen des europäischen gemeinsamen Doktorandenprogramms HPC-LEAP ist.

Acknowledgments

Preface

In September 2015 I started my Ph.D. under the European Joint doctorate program High Performance Computing in Life sciences Engineering and Physics (HPC-LEAP) founded by the European Union's Horizon 2020 research and innovation program. As an early-stage researcher (ESR) I was given a joint research project between the University of Cyprus (UCY) and the University of Wuppertal (BUW). The aim of the project was the development and the application of multigrid methods in Lattice Quantum Chromodynamics (LQCD), which nicely combined the research interests of my two supervisors C. Alexandrou, Professor of theoretical physics at UCY, and A. Frommer, Professor of applied mathematics at BUW. Prof. Alexandrou has many years of experience in research in LQCD, she is the Acting Director of the Computation-based Science and Technology Research Center (CaSToRC) of The Cyprus Institute (CyI) and member of the Extended Twisted Mass (ETM) collaboration. This, as well as the HPC-LEAP network, gave me the opportunity to collaborate with many other researchers in lattice QCD outside of my two institutions. Prof. Frommer has pioneered algorithms that have become mainstream within LQCD community codes. I was thus exposed in the mathematical basis for the development of algorithms for large scale simulations and for the tools for achieving improved performance.

For the successful implementation of the project, I received a multidisciplinary training in numerical methods, computational physics and HPC. The workshops organized by the HPC-LEAP program contributed significantly in the rapid development of the required interdisciplinary knowledge and skills. During the first year of the project, we had an intensive training program divided in four workshops of three weeks each. The topic of the first workshop organized by the university of Wuppertal was "Numerical analysis and algorithms towards exascale". It was a perfect starting and gave me also the advantage to be embedded in the group of Prof. Frommer at the University of Wuppertal receiving an introduction to multigrid methods and to employ them in our LQCD projects. This collaboration resulted in our first paper entitled "Adaptive Aggregation-based Domain Decomposition Multigrid for Twisted Mass Fermions" [1]. Other works followed on the application of multigrid methods for twisted mass fermions [2, 3].

In collaboration with the LQCD group at the Cyprus Institute as well as with members of the ETM collaboration we carried out an intensive research program with the goal to improve simulation of gauge field ensembles and to enable studies of hadron structure. Using multigrid methods we succeeded in simulating a series of twisted mass ensembles with quark masses tuned to their physical value. To accomplish this we found an optimal setup for the multigrid algorithm and developed codes optimized for the biggest supercomputers, including several numerical studies that resulted in several publications [4–7]. This included the generation of the first ensemble of twisted mass fermions at the physical point with up, down, strange and charm quarks all tuned to their physical mass as reported in Ref. [8]. Two additional ensembles were also simulated and a fourth one is currently under production. We used our new ensembles for conducting several studies on hadronic observables as scattering processes, meson decay constants, nucleon parton distribution functions and the pion and nucleon electromagnetic form factors. They can be found in our publications [9–13].

Contents

List of Figures

List of Tables

1. Introduction

Research in lattice Quantum Chromodynamics (QCD) relies critically on algorithms and novel computer architectures. This thesis addresses the development of algorithms that deal with the critical slowing down as light quark mass approaches its physical value. The algorithm development is specifically oriented for twisted mass fermions.

Lattice QCD provides the only known framework to solve non-perturbatively the fundamental theory of the strong interactions starting directly from the QCD Lagrangian. It uses the same input parameters as those of the continuum theory, namely the bare quark masses and coupling constant. It is a very appropriate formalism for nuclear physics since it provides a rigorous framework to compute hadronic matrix elements and compare with experimental results. In lattice QCD one solves the QCD Lagrangian on a 4-dimensional Euclidean lattice, discretizing the space-time. This introduces several sources of systematical effects which need to be kept under control. Large lattice sizes, fine lattice spacing and dynamical quarks are required for obtaining results that can be compared with observations. Simulations that are performed with dynamical quarks with masses tuned to their physical value are referred to as *the physical point* simulations. Such simulations have been targeted for many years by the lattice QCD community. A lot of effort has been devoted to reach the technological, algorithmic and theoretical developments that nowadays allow to perform simulations at the physical point with large enough volume and fine enough lattice spacing. A number of lattice QCD collaborations are currently pursuing such simulations providing results that have never before been so reliable.

The thesis is structured as follows:

- In Chapter 1 we introduce the fundamental concepts of Quantum Chromodynamics starting from the standard model. We also discuss physical quantities studied in lattice QCD and open questions we aim to address.

- In Chapter 2 we introduce the concepts of lattice QCD focusing in particular on the twisted mass fermion formulation used in our studies.

- In Chapter 3 we present the iterative solvers for computing the inverse of the Dirac twisted mass operator. We begin with a qualitative description of standard iterative Krylov solvers. We derive the Conjugate Gradient algorithm and present the problem of the critical slowing down, which appears when we approach the physical pion mass. Then we introduce multigrid methods, which provide a solution to the critical slowing down and we review their application in lattice QCD. We explain first inexact deflation as introduced by M. Lüscher [14], we review the first algebraic multigrid approach in lattice QCD referred to as MG-GCR, and we conclude with the formal description of the multigrid algorithm we use, namely the Aggregation-based Domain Decomposition Multigrid (DD-αAMG) solver.

- In Chapter 4 we describe the application of the DD-αAMG solver to compute quark propagators of twisted mass fermions. We present several numerical results obtained on configurations at the physical point and we provide a detailed analysis of the multigrid parameters and their properties.

- In Chapter 5 we describe the simulation approach used in Lattice QCD for computing observables, i.e. the Monte Carlo integration. We start from pure-gauge theories employing the Metropolis algorithm, then we integrate the fermions using the pseudo-fermions representation and we solve the integral using hybrid MC (HMC) and rational HMC algorithms. We also present our approach for accelerating inversions in the rational approximation by providing an initial guess to the solver.

- In Chapter 6 we detail our procedure for the generation of the gauge ensembles at the physical point that include two $N_f = 2$ and two with $N_f = 2 + 1 + 1$ ensembles of twisted mass clover-improved fermions. We discuss the improvements achieved using the new algorithms, the tuning procedure and physical observables that test the accuracy of the simulation.

- In Chapter 7 we show selected results for meson and nucleon observables computed on the aforementioned physical ensembles and we compare them with older results using an $N_f = 2$ ensemble with lattice volume $V = 48^3 \times 96$ addressing finite size and heavy quark quenching effects on these observables.

1.1 Standard model
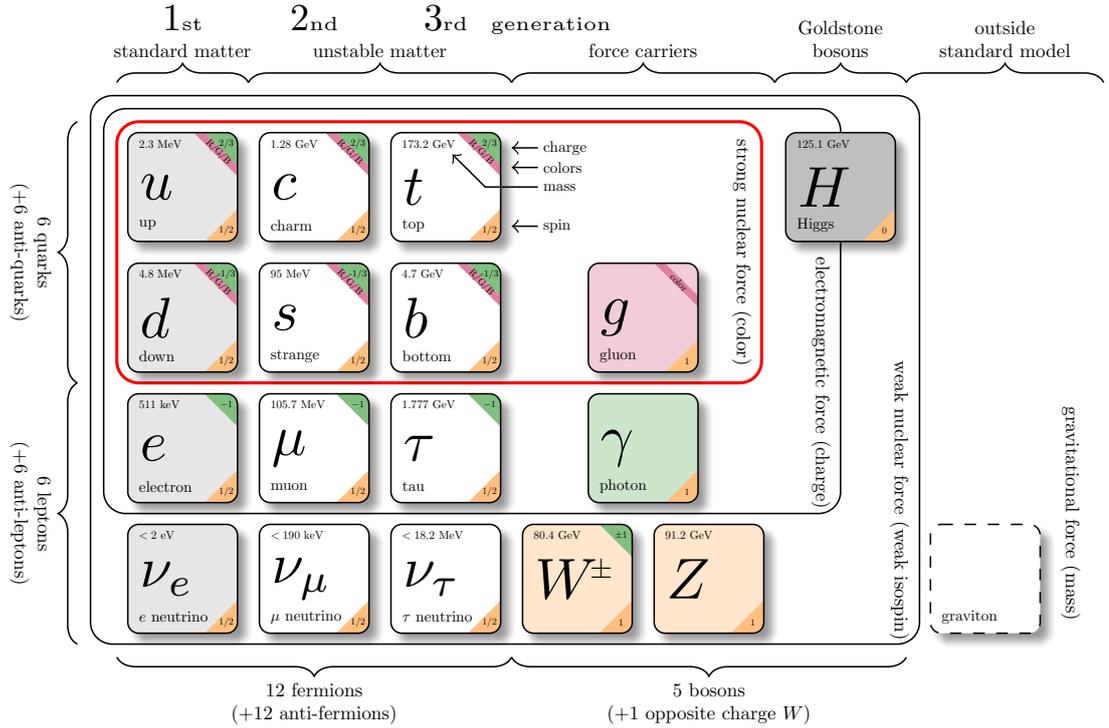
Figure 1.1: Schematic representation of the standard model. Within the red box we indicate the degrees of freedom of the theory of the strong interactions, QCD. *Image credit: CERN.*

Four fundamental forces are believed to govern the universe and they describe its dynamical evolution. Gravity is the most well-known force that is evident on the macroscopic level being a long distance interaction. The classical theory of gravity finds its most complete formulation in the general theory of relativity [15] by A. Einstein. The other three forces are electromagnetism, the weak and

strong forces. Electromagnetism is a long-distance force while the weak and strong forces contribute to short distance interactions among particles. The Standard Model (SM) is the quantum field theory formulation of the last three forces. The gravitational force does not have yet a consistent quantum field theory formulation. The effects of this force are very small on the microscopic level and can thus be neglected for the phenomena of interest here.

Electromagnetism was the first to be quantized and it serves as a prototype quantum field theory. Originally, electricity and magnetism were considered as two separate forces until J. C. Maxwell unified electric and magnetic phenomena, which are now described by Maxwell's classical theory of electromagnetism. Maxwell provided a theoretical formulation where electric and magnetic fields are described by wave functions. Within the standard model, electromagnetism is quantized and it is known as Quantum Electrodynamics (QED). In QED [16] the interactions between charged particles – the fermions, namely electrons, muons and τs – are described through the exchange of photons, γ. Photons are massless particles and thus the electromagnetic forces have infinite range. Mathematically, QED is an abelian gauge theory with symmetry group $U(1)$. The gauge field, which mediates the interaction between the charged spin-1/2 fermions is the electromagnetic field. The QED Lagrangian is given in natural units by the real part of

$$\mathscr{L}_{QED} = \sum_{f=e,\mu,\tau} \bar{\psi}_f (i\gamma^\mu D_\mu - m_f)\psi_f - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{1.1}$$

where

- γ^μ are Dirac matrices[1];

- ψ_f is a spinor field and f denotes the different leptons, namely the electron, the muon and the tau; $\bar{\psi} \equiv \psi^\dagger \gamma^0$;

- $D_\mu \equiv \partial_\mu + ieA_\mu$ is the gauge covariant derivative;

- e is the electric charge;

- m_f is the mass of the different fermions;

- A_μ is the electromagnetic gauge field;

- $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field tensor describing the gauge field.

The weak force is responsible for β-decays and nuclear fission, it is short ranged and can change the flavor of a quark. The mediators of this force are the Z^0 and $W^\pm$ bosons, which have masses [17] $91.1876(21)$ GeV and $80.379(12)$ GeV respectively. The Standard Model unifies the electromagnetic and the weak interactions in what is referred to as the electroweak (EW) force (S. Glashow [18], A. Salam [19] and S. Weinberg [20] won the Nobel price in 1979 for the formulation of the SM). The EW lagrangian describes the interaction between leptons and quarks through a composite $U(1) \times SU(2)$ gauge-field theory:

$$\mathscr{L}_{EW} = \sum_{\substack{f=u,d,s,c,t,b, \\ e,\mu,\tau,\nu_e,\nu_\mu,\nu_\tau}} \bar{\psi}_f (i\gamma^\mu D_\mu^f - m_f)\psi_f - \frac{1}{4}W_a^{\mu\nu}W_{\mu\nu}^a - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + \mathscr{L}_h + \mathscr{L}_y \tag{1.2}$$

---

[1]The γ-matrices fulfill the Dirac algebra $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$ where $g^{\mu\nu} = \text{diag}(+1,-1,-1,-1)$ is the metric in Minkowski space.

where

- D_μ^f is the gauge covariant derivative for leptons when $f = e, \mu, \tau, \nu_e, \nu_\mu, \nu_\tau$ and for quarks when $f = u, d, s, c, t, b$;

- $W_a^{\mu\nu}$, with $a = 1, 2, 3$, ($SU(2)$) and $B^{\mu\nu}$ ($U(1)$) are the field strength tensors for the weak isospin and weak hypercharge gauge fields;

- $\mathscr{L}_h$ describes the Higgs field and its interactions;

- $\mathscr{L}_y$ describes the Yukawa interaction between the Higgs field and the fermions.

The Higgs field produces spontaneous symmetry breaking through a non-vanishing vacuum expectation value. From the Higgs mechanism QED emerges as an $U(1)$ gauge field theory, where the electromagnetic field is described by the photon, and an $SU(2)$ gauge field theory representing the weak interaction with bosons Z^0 and $W^\pm$. The new bosons are connected to W_a and B, cf. Eq. (1.2) by the relations

$$\begin{pmatrix} \gamma \\ Z^0 \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix} \quad \text{and} \quad W^\pm = \frac{1}{\sqrt{2}}(W_1 \mp iW_2), \tag{1.3}$$

where θ_W is the weak mixing angle. Moreover the bosons Z^0 and $W^\pm$ acquire masses, M_Z and M_W, respectively, connected by the relation $M_Z = \frac{M_W}{\cos\theta_W}$, while γ is still mass-less.

The strong interaction is the force that binds together protons and neutrons to form the nucleus of atoms [21]. For nuclear systems the characteristic strong scale is about 1 *fm* or about 1 *GeV*. The carrier of this force is the gluon, which mediates the interaction between quarks. The strong interaction introduces a new kind of charge, the so-called color, that can take three values (*r*, *g* or *b*). Combinations of three different colored charges or color charge and anti-charge (e.g. $\bar{r} + r$) give colorless states, which the only ones that can be observed. Therefore, the combinations $\varepsilon^{abc} q_a q_b q_c$ or $\bar{q} + q$ make a colorless baryon and meson respectively. From this follow the name assigned to the respective field theory, Quantum Chromodynamics (QCD). We describe the properties of QCD in the next section.

All the constituents of the standard model have been discovered when in 2012 CERN announced the discovery of a particle with all the quantum numbers of the Higgs boson and, in 2014, confirmed that this particle is indeed the Higgs boson [22]. Thus the standard model describing the electroweak and strong interactions is the current theory of particle physics. Major experiments at CERN are searching for physics beyond the standard model (BSM). Such theories may provide an explanation for dark matter making 24% of matter in the universe the nature of which is currently unknown. Indeed the purpose of lattice QCD is three-fold, namely to:

- provide reliable theoretical predictions solving non-perturbatively the QCD Lagrangian enabling comparison with the experimental results;

- to provide results for systems where experimental measurements are inaccessible, and

- to compute quantities that can probe BSM physics.

1.2 Quantum Chromodynamics (QCD)

In the 50s, after the invention of spark and bubble chambers, a large number of hadrons were discovered. These particles were categorized according to their properties, see Fig. 1.2. Following the idea of the periodic table of elements in chemistry, an underlying model was proposed to understand the properties of these baryons and mesons. In 1963 R. Feynman, M. Gell-Mann and G. Zweig [23] proposed that the structure of these particles can be explained if hadrons were not fundamental particles, but composed from three elementary particles called partons. However, the Δ^{++} presented a puzzle since it consists of three strange quarks with aligned spins. This state is prohibited by the Pauli principle, so the only case that such a particle could exist was if the three quarks carried an additional quantum number, the color charge. When first proposed, quarks were considered as a mathematical construction and not as physical particles. R. Feynman *et al.* [24] interpreted the results of experiments at high energy colliders as showing evidence that hadron were indeed not elementary particles and that the proton consisted from 3 quarks. This established the physical existence of quarks.
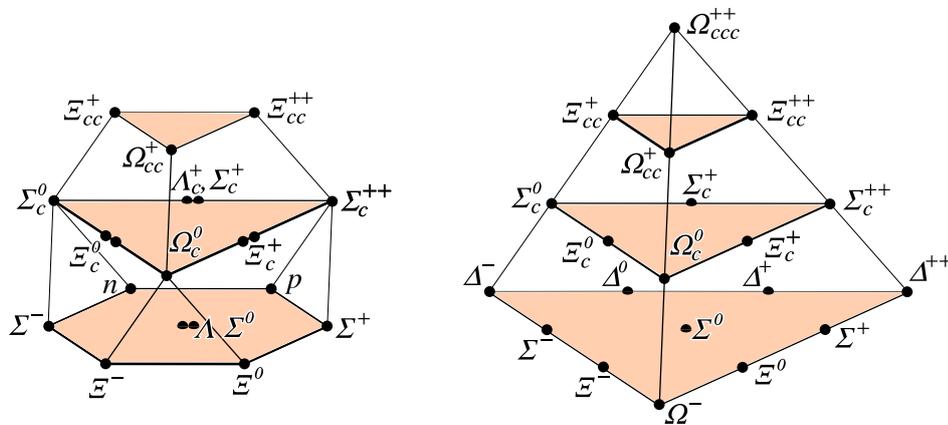
Figure 1.2: The baryon multiplets of up (u), down (d), strange (s) and charm (c) quarks obtained using the $SU(4)$ group. The left diagram shows the spin-1/2 20-plet and the right the spin-3/2 20′-plet.

In 1973 Fritzsch, Gell-Mann and Leutwyler [25] formulated a consistent theory for QCD, as a non-abelian $SU(3)$ gauge-field theory [26] where quarks carry three different values of color charges. This introduces a fundamental difference between QED and QCD. The gluons can interact with each other, as opposed to photons, which do not self-interact. Two main fundamental characteristics of the non-abelian nature of QCD follow:

- **Confinement**, which arises because the force between two quarks remains constant as the distance between them goes to infinity. As we try to separate two quarks the energy transferred to the system will permit the creation of a quark-antiquark pair. Thus, it is not possible to isolate a quark in order to study it experimentally. The mechanism of confinement is still not understood analytically but it has been demonstrated in lattice QCD [27].

- **Asymptotic freedom**, which implies that at high energy scales or equivalently when the distance between quarks becomes small, they interact very weakly. This prediction was confirmed experimentally in the 70s [28, 29], by measuring the strong coupling constant as a function of energy.

The Lagrangian of QCD can be written in a very elegant form following the notation of the QED

Lagrangian in Eq. (1.1), namely

$$\mathscr{L}_{QCD} = \sum_{f=u,d,s,c,t,b} \bar{\psi}_f(i\gamma^\mu D_\mu - m_f)\psi_f - \frac{1}{2}F^a_{\mu\nu}F^{\mu\nu}_a \quad . \tag{1.4}$$

Even though the mathematical equation for the QCD and QED Lagrangian look very similar, the physical properties of QCD differ significantly from those described by QED. In the QCD Lagrangian

- $\psi_f \equiv (\psi^a_\mu)_f$ are the quark color-spinor fields and they carry a color index a, additionally to the spin index μ; the same applies to the gluon gauge field A^a_μ;

- $D_\mu \equiv \partial_\mu + ig_s A^a_\mu$ is the gauge covariant derivative where the coupling "constant", g_s, increases logarithmically with decreasing energy, E,

$$g^2_s(E^2) \propto \ln\left(\frac{\Lambda^2_{QCD}}{E^2}\right), \tag{1.5}$$

  where Λ_{QCD} is the QCD renormalization group invariant scale and its value depends on the specific renormalization scheme;

- $F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g_s f^a_{bc} A^b_\mu A^c_\nu$ is the field tensor[2] describing the gauge field. The last term $g f^a_{bc} A^b_\mu A^c_\nu$ is responsible for the self-interaction among gluons not present in QED.

The particles resulting from QCD are "colorless" and are made up of valence quarks and/or anti-quarks confined inside mesons and baryons. The six flavors of quarks differ only by their mass term in the QCD Langragian. The quarks also carry an electric charge, which is $+\frac{2}{3}e$ for the up-, charm- and top-quarks and $-\frac{1}{3}e$ for the down-, strange- and bottom-quarks. The strange quark carries a strangeness quantum number, $S = (n_{\bar{s}} - n_s)$ where n_s and $n_{\bar{s}}$ represent the number of strange- and antistrange-quark, respectively. This number is conserved in strong and electromagnetic interactions while in weak interactions can vary between $-1, 0, +1$.

1.2.1 Solving QCD numerically on the lattice

The theory of strong interactions describes the complex hierarchy and relations among hadrons with a small number of input parameters, namely the mass of the quarks and the coupling constant. However, due to its non-perturbative nature it is a major challenge to solve it. K. G. Wilson [30] managed to regularize the theory by defining it on a discretized space-time. Rotating the discretized theory to Euclidean time provides a starting point for a numerical solution of QCD. Creutz performed the first SU(2) gauge simulation in 1980 [31]. The discretized Euclidean formulation of QCD is known as lattice QCD and it provides a robust *ab initio* method for solving QCD. After taking the continuum limit of lattice QCD, the continuum theory is recovered.

In its early days, due to limited computational resources and advanced algorithms, LQCD simulations were restricted to pure gauge. Quenched QCD simulations, i.e. without sea quarks were performed. A major success, was the numerical confirmation of the confinement via the computation of the linearly rising quark-antiquatk potential [27]. Full QCD simulations and fermionic observables started after 2000 with the advancement in theoretical and numerical methods as well as the availability of bigger computers. It is only in the last 10 years that reliable computation of properties of

---

[2]Note that $F^a_{\mu\nu}$ is the result of $[D_\mu, D_\nu] = F_{\mu\nu} = -igF^a_{\mu\nu}t_a$ where t_a are the generators of $SU(3)$ that fulfill $[t_b, t_c] = if^a_{bc}t_a$.

hadrons have emerged. A landmark was the computation of the low-lying hadrons masses [32]. After taking the continuum and infinite volume limits and extrapolating to physical light quark mass the low lying hadron spectrum was reproduced. A collection of results on the low-lying baryon masses computed by several LQCD collaborations is depicted in Fig. 1.3. A major success was the prediction of the masses of doubly charmed baryons within LQCD. These particle were subsequently discovered at LHCb at CERN [33] with mass consistent with the one predicted by lattice QCD, see e.g. the red bands in middle plot of Fig. 1.3.

Another benchmark quantity beyond hadron masses is the nucleon axial charge g_A. It is well-measured experimentally and it is the simplest hadronic matrix element that can be computed within LQCD. Many LQCD collaborations have computed g_A mostly using light quark masses larger than physical and consistently underestimated it. It is only in the past five years that one can extract g_A directly at the physical point. A number of systematic errors are still under scrutiny but we are beginning to see a resolution to the disagreement between lattice QCD results and experiment. With the progress in both algorithms and computers LQCD is targeting a wider range of observables including quantities that can probe physics BSM. For a comprehensive overview of recent lattice QCD results see the Flavour Lattice Averaging Group (FLAG) report[3] with the latest version given in Ref. [45]. The FLAG report aims to be the lattice equivalent of the PDG providing results from various lattice QCD groups and performing an average of state-of-the-art results that satisfy certain criteria.

The computations presented in this thesis are done within the twisted mass fermion discretization scheme. The twisted mass formulation [46] is a lattice regularization that allows automatic $\mathcal{O}(a)$ improvement by tuning only one parameter. This formulation is particularly appropriate to hadron structure studies since the renormalization of local operators is significantly simplified with respect to the standard Wilson regularization and no improvement of the currents is needed. In the continuum, the twisted mass formulation is equivalent to the standard QCD action in a different basis. The calculation of nucleon form factors is a major line of research in hadron structure. For instance, the proton electric and magnetic form factors of the nucleon, $G_E(Q^2)$ and $G_M(Q^2)$, encapsulate information on the distribution of charge and magnetization and are among the quantities that have been extensively studied experimentally. The slope of the electric and magnetic form factors at zero momentum yields the electric and magnetic root mean square radius, while the value of the form factors at zero momentum gives its electric charge and magnetic moment. Extensive electron scattering experiments have been carried out since the 50s for the precise determination of the nucleon form factors, including recent experiments at Jefferson Lab, MIT-Bates and Mainz. The proton radius can also be obtained spectroscopically, namely via the Lamb shifts of the hydrogen atom and of muonic hydrogen [47] and via transition frequencies of electronic and muonic deuterium. Despite the longer than fifty years of experimental activity, new precise experiments with muonic atoms hold surprises. As depicted in Fig. 1.4, a measurement of the charge radius of the proton at PSI [47] extracted a value that is smaller by almost seven standard deviations as compared to the value extracted from electron scattering experiments. This sparked a wealth of theoretical explanations some of which involve new physics. Whether new physics is responsible for this discrepancy, or errors in the theoretical or experimental analyses [48], a first principles calculation of the electromagnetic form factors of the nucleon can provide valuable insight. Although nucleon electromagnetic form factors have been extensively studied in lattice QCD, most of these studies have been carried out at higher than physical pion masses, requiring extrapolations to the physical point, which for the case of baryons carry a large systematic

---

[3] http://flag.unibe.ch/
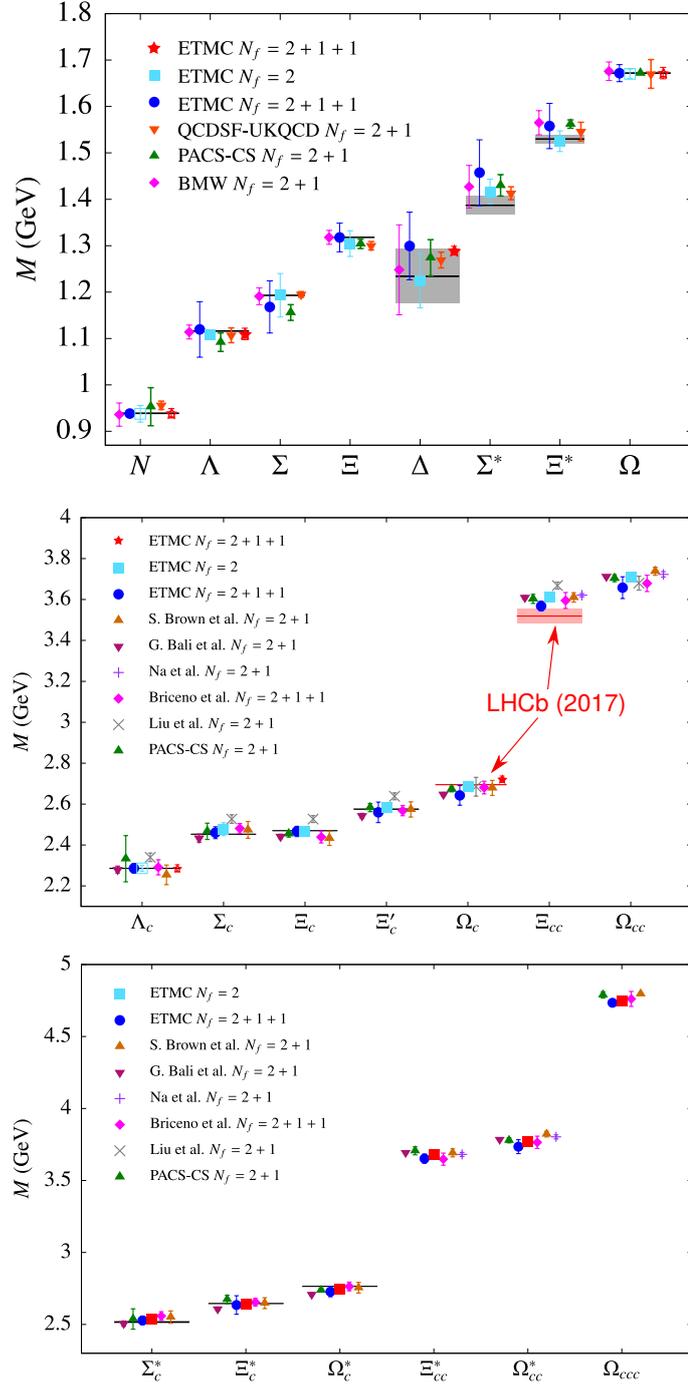
Figure 1.3: Low-laying baryon masses from LQCD and experiments. Points are the results produced by a number of LQCD collaborations [32, 34–43]. The bands are the results obtained by experiments [33, 44]. For most baryons the band is too small to be visible.
Top: The octet and decuplet baryon masses made of up, down and strange quarks.
Middle: The masses of spin-1/2 charm baryons.
Bottom: The masses of spin-3/2 charm baryons.
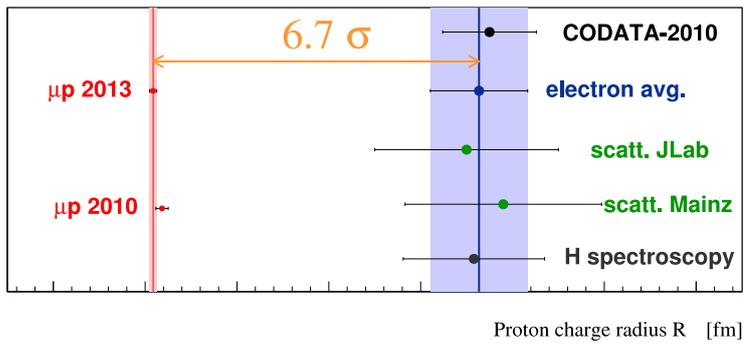
Figure 1.4: Left: Extractions of the proton charge radius from muonic hydrogen measurements [47, 49], hydrogen spectroscopy [50], electron scattering measurements at Mainz [51], and a global analysis of earlier world data [52]. The direct average shown is compared to the CODATA-2010 evaluation [50]. Image credit: *Randolf Pohl.* Right: Nature cove page, volume 466 issue number 7303, where Ref. [47] has been published.

uncertainty. One of our objectives is to obtain high accuracy results at the physical point for nucleon form factors.

2. Overview of lattice QCD

In this chapter we review the Wilson's formulation of lattice gauge theories. We begin with the connection between statistical mechanics and the lattice regularization of Feynman's path integrals. We then introduce the Wilson proposal for regularizing QCD on the lattice regarding gauge and fermionic quantities. Finally, we introduce the fermionic discretization used in this thesis, twisted mass fermions, which is an extension of the Wilson discretization in the twisted mass base.

2.1 Formulation of a lattice gauge field theory

The lattice regularization of gauge field theories was proposed by K. G. Wilson [30]. We review here this formulation. Let us consider a group $\mathscr{G}$ with infinitesimal generators t_a, which fulfill $[t_b, t_c] = i f_{bc}^a t_a$ where f_{bc}^a are referred to as structure constants. The covariant derivative of the theory is

$$D_\mu = \partial_\mu + A_\mu^a t_a, \tag{2.1}$$

where $A_\mu^a(x)$ is an element of the algebra referred to as gauge field. The so-called field strength tensor is

$$F_{\mu\nu} = [D_\mu, D_\nu] = -ig F_{\mu\nu}^a t_a. \tag{2.2}$$

A Lagrangian of the form

$$\mathscr{L}_{\mathscr{G}} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{2} F_{\mu\nu}^a F_a^{\mu\nu} \tag{2.3}$$

is invariant under gauge transformations $G(x)$ under which

$$
\begin{aligned}
A_\mu(x) &\to A_\mu'(x) = G(x) A_\mu(x) G^{-1}(x) + i\left(\partial_\mu G(x)\right) G^{-1}(x), \\
\psi(x) &\to \psi'(x) = G(x)\psi(x), \\
\bar{\psi}(x) &\to \bar{\psi}'(x) = \bar{\psi}(x) G^{-1}(x).
\end{aligned}
\tag{2.4}
$$

In what follows we indicate the Lagrangian with $\mathscr{L}_{\mathscr{G}} \equiv \mathscr{L}_{\mathscr{G}}(A_\mu, \bar{\psi}, \psi, g, m)$ and the operators with $O \equiv O(A_\mu, \bar{\psi}, \psi)$ where the fields A_μ, ψ and $\bar{\psi}$ are elements of the algebra, g stands for the coupling parameter, and m_f for the quark masses. Note that the operators are functions of the fields only.

2.1.1 Feynman path integral formulation

The Feynman path integral is a useful formulation for quantum field theories and an alternative description of quantum mechanics [53, 54]. In quantum mechanics the quantum mechanical probability amplitude for a particle described with the Hamiltonian H to evolve from a point $\vec{x}_0$ to $\vec{x}_f$ of space in the time interval t is given by the multidimensional integral

$$\langle \vec{x}_f | e^{-iHt} | \vec{x}_0 \rangle = \int_{\vec{x}_0, 0}^{\vec{x}_f, t} \mathscr{D}\vec{x}(t) \, e^{iS(\vec{x}(t))} \tag{2.5}$$

where the action $S(\vec{x}(t)) = \int_{\vec{x}(t)} \mathscr{L}(x) d^4 x$ depends on the Lagrangian density $\mathscr{L}(x)$ and $\int_{\vec{x}_0, 0}^{\vec{x}_f, t} \mathscr{D}\vec{x}(t)$ is the normalized measure which integrates over every possible path in the space, $\vec{x}(t)$, which fulfills

$\vec{x}(0) = \vec{x}_0$ and $\vec{x}(t) = \vec{x}_f$.

Applying the path integral formulation to a gauge field theory described by a Lagrangian L, we can compute the vacuum expectation value of an operator O using

$$\langle O \rangle \equiv \langle \Omega | O | \Omega \rangle = \frac{1}{Z} \int \mathscr{D}A_\mu \mathscr{D}\bar{\psi} \mathscr{D}\psi \, O(A_\mu, \bar{\psi}, \psi) e^{iS(A_\mu, \bar{\psi}, \psi, g, m)} \qquad (2.6)$$

where now $\int DA_\mu \mathscr{D}\bar{\psi} \mathscr{D}\psi$ is a functional integral over all the possible fields. The normalization factor Z,

$$Z = \int \mathscr{D}A_\mu \mathscr{D}\bar{\psi} \mathscr{D}\psi e^{iS(A_\mu, \bar{\psi}, \psi, g, m)}, \qquad (2.7)$$

is the partition function.

2.1.2 Connection to statistical mechanics and lattice regularization

There is a close relation between Feynman path integral formulation and statistical mechanics. In statistical mechanics a system described by the Hamiltonian, $\mathscr{H}(s)$ is defined over a finite number of degrees of freedom, $s = s_i$, e.g. spin variables in the Heisenberg model. Then the probability of finding the system in a particular configuration s is given by

$$P(s) = \frac{1}{Z} e^{-\beta \mathscr{H}(s)} \qquad (2.8)$$

where the partition function

$$Z = \sum_s e^{-\beta \mathscr{H}(s)} \qquad (2.9)$$

has a sum over all the possible spin configurations. The parameter $\beta = 1/KT$ acts as a thermodynamic factor that scales the energy of the system, where T is the temperature. The expectation value of an operator is given by

$$\langle O \rangle = \frac{1}{Z} \sum_s O(s) e^{-\beta \mathscr{H}(s)}. \qquad (2.10)$$

Eqs. (2.6) and (2.10) express expectation values with a similar structure, i.e. a sum over all the possible configurations of the basic degrees of freedom of the system weighted by the exponential of the Hamiltonian or classical action of the system. In order to map a quantum field theory to a classical statistical theory the following steps are performed:

1. **Wick rotation:** In Eq. (2.6) we have a phase, which does not allow us to interpret it as a probability unlike Eq. (2.10), where one can use Monte Carlo methods introduced in Sec. 5.1 to evaluate the multidimensional integrals. If one performs a Wick rotation (see Ref. [55] for an overview) by letting the time $t_E = it$ then the action in Eq. (2.5) gets an imaginary factor, i.e.

$$S(\vec{p}(t)) = \int_0^t \mathscr{L}(\vec{p}(t')) dt' = i \int_0^t \mathscr{L}(\vec{p}(it_E')) dt_E' = iS_E(\vec{p}_E(t_E)) \qquad (2.11)$$

where S_E is the respective Euclidean action of S and $\vec{p}_E(t_E)$ is the respective path in Euclidean space-time of $\vec{p}(t)$ which is in Minkowski space-time. Therefore the observable expectation value becomes

$$\langle O \rangle = \frac{1}{Z} \int \mathscr{D}A_\mu \mathscr{D}\bar{\psi} \mathscr{D}\psi \, O(A_\mu, \bar{\psi}, \psi) \, e^{-S_E(A_\mu, \bar{\psi}, \psi, g, m)} \quad \text{with} \quad Z = \int \mathscr{D}A_\mu \mathscr{D}\bar{\psi} \mathscr{D}\psi \, e^{-S_E(A_\mu, \bar{\psi}, \psi, g, m)} \qquad (2.12)$$

2. **Lattice regularization:** In order to perform numerical simulations one discretizes the theory by defining it on a 4-dimensional Euclidean lattice. The purpose of this step is two-fold: it provides a non-perturbative regularization avoiding ultraviolet diverges and presents a formulation which is amenable to numerical simulations.

   The lattice Euclidean regularization

   - is defined on a four-dimensional hypercubic lattice of finite volume

   $$\mathscr{V} = \{x_i = (x_0, x_1, x_2, x_3), 1 \leq x_0 \leq N_T, 1 \leq x_1, x_2, x_3 \leq N_L\} \qquad (2.13)$$

   with N_T being the number of points in the temporal direction and N_L the number of points in the spatial directions x, y and z. Here and in what follows, the lattice spacing a is set to unity. The lattice volume is given by $V = T \cdot L^3$, where T is the temporal and L the spatial length.

   - defines discrete fields on the lattice sites $x_i \in \mathscr{V}$

   $$\psi(x) \to \psi_i = \psi(x_i), \quad \bar{\psi}(x) \to \bar{\psi}_i = \bar{\psi}(x_i) \quad \text{and} \quad A_\mu(x) \to A_{\mu,i} = A_\mu(x_i) \qquad (2.14)$$

3. **The β parameter:** as last, in analogy with the statistical approach in Eq. (2.10), one can introduce a β parameter that scales the action in the measure factor $e^{-S_E} \to e^{-\beta S_E}$. In a thermo-dynamical average β is related to the temperature T, $\beta = 1/kT$, and S_E is the internal energy of the system. In lattice gauge theory, instead, β is related to a coupling constant g, $\beta \propto 1/g^2$ and S_E is a rescaled action rather than a true energy. It is worth noting that g entering the definition of β is an unrenormalized coupling constant, which describes the interaction strength at the distance a. Contact with the physical world, is made via a suitable continuum limit as described later in the chapter.

   The introduction of the β parameter let us consider two expansion approaches one for high values of β, i.e. weak coupling, and one for small values of β, i.e. strong coupling:

   - The weak coupling expansion of lattice gauge theories applies standard perturbative approaches made, however, more complicated by the loss of Lorentz covariance that leads to additional terms in the perturbative expansion. Weak coupling expansions of the lattice theory become necessary when we need to compare with physical results.

   - The possibility of a performing strong coupling expansion is a tool provided by the lattice regularization that can lead to insights in the theory. Unfortunately, at intermediate values of β a phase transition occurs prohibiting the analytic continuation between weak and strong coupling. This phase transition is due to the lattice spacing becoming too rough [56] and it is named roughening phase transition.

2.1.3 Continuum limit

Lattice quantum field theories require taking a suitable continuum limit, i.e. $a \to 0$, at infinite lattice volume, i.e. $V \to \infty$. In analytic calculations the latter limit is taken as first and then the continuum limit is calculated. Namely,

$$\lim_{a \to 0} \lim_{V \to \infty} S_E(a, V, g, m) = S_E(g, m) \quad \text{and} \quad \lim_{a \to 0} \lim_{V \to \infty} O(a, V, g, m) = O(g, m). \qquad (2.15)$$

In numerical calculations, instead, one analyzes the dependence on a and V on the observables,

$$O(a,V,g,m) = \frac{1}{Z} \int d^V \phi_i \, O(A_{\mu,i}, \bar{\psi}_i, \psi_i, a) \, e^{-S_E(A_{\mu,i}, \bar{\psi}_i, \psi_i, a, g, m)} \tag{2.16}$$

and the values at continuum and infinite volume limits are extrapolated. The effects of finite a and V are known as *lattice artifacts* and they need to be kept under control for obtaining the correct physical result. They are classified into:

- **Discretization effects:** the continuum limit of the discretized quantities is studied through a Taylor expansion over small lattice spacing. For instance, an observable can be written as

$$O(a,g,m) = O(g,m) + \sum_{n=1}^{\infty} a^n O_n(g,m) \tag{2.17}$$

  where $O_n(g,m)$ are corrections of order $\mathscr{O}(a^n)$ of the continuum quantity $O(g,m)$. Thus one takes a lattice spacing small so that only leading order finite a corrections are important. *Improved* discretization schemes are those where the first order a correction vanishes. Such improved actions are nowadays by state-of-the-art simulations leading to discterization effects of order $\mathscr{O}(a^2)$ instead of $\mathscr{O}(a)$.

- **Finite volume effects:** These are a consequence of the finite volume used in numerical simulations. It is common to take periodic boundary conditions in space and anti-periodic in time and these are the ones considered in this work. However, other choices can also be considered.

  Finite volume effects depend on the lowest modes, which have longer wavelengths. In LQCD the lowest mode is the pion and thus the dimensionless quantity Lm_π is commonly used for characterizing finite size effects. Chiral perturbation theory can be used in particular in the meson sector to estimate volume effects, see e.g. Ref. [57, 58].

For completeness sake, other two well-known systematic effects in lattice simulation are

- **Quenching effects:** Simulations where dynamical quarks degrees of freedom are neglected introduce quenching effects. Such simulations are less computationally demanding and were used extensively in the 80's and 90's. Nowadays simulations commonly are done with dynamical light and strange quarks. In a number of simulations dynamical charm quarks are also included, as done in this thesis.

- **Pion mass dependence:** As last, simulating at physical pion mass makes LQCD computationally demanding. This is due to two reasons:

  1. as mentioned above, finite volume effects are characterized by the quantity Lm_π. Then, at fixed Lm_π the computational costs, which increase with the lattice volume $\propto L^4$, increase as $\propto 1/m_\pi^4$.
  2. Moreover, the time to solution of some iterative algorithms used in LQCD, as described in Sec. 3.1.2, has a dependence on the pion mass, namely $\propto 1/m_\pi^2$ in the referred case.

  Combining these two effects, simulation costs have a dependence on the pion mass $\propto 1/m_\pi^6$ which can significantly affect the feasibility of the calculation. Thus simulating at a pion mass larger than the physical one helps to reduce the computational costs and for this reason chiral

perturbation theory has been extensively used in LQCD for computing pion mass corrections to the physical quantities. In this thesis we consider, instead, only simulations performed at physical pion mass.

2.1.4 Wilson's proposal

In a gauge field theory, the action is invariant under gauge transformations, cf. Eq. (2.4). If the field $\psi(x)$ belongs to a definite representation of the gauge group, then a scalar product $\bar{\psi}(x+dx)\psi(x)$ would not be invariant under local gauge transformations,

$$\psi(x) \rightarrow G(x)\psi(x) \tag{2.18}$$

where $G(x)$ is a position dependent, finite element of the gauge group. To construct an invariant scalar product, the field must first be "transported" from x^μ to $x^\mu + dx^\mu$ and this is done by multiplying by the so-called transport operator

$$\bar{\psi}(x+dx)e^{igA_\mu^a(x)\tau_a dx^\mu}\psi(x), \tag{2.19}$$

where g is the gauge coupling constant. This quantity is invariant under the local gauge transformation in Eq. (2.18). Thus, the gauge acts as "transport operators" and can be used to compare orientations at neighboring points x^μ and $x^\mu + dx^\mu$.

This suggests that on the lattice the gauge dynamical variables are associated with oriented links between neighboring vertices x_i and $x_i + a\hat{\mu}$ rather vertices themselves. Thus, in the lattice regularization we take as basic dynamical variables finite group elements $U_\mu(x_i) \in \mathscr{G}$ defined on oriented links. The gauge-link U_μ is a finite element because the displacement is now over a finite length a given by

$$U_\mu(x_i) = e^{igA_\mu^a(x_i)\tau_a a} \simeq 1 + igA_\mu^a(x_i)\tau_a a. \tag{2.20}$$

Then the equivalent of Eq.(2.19) on the lattice is

$$\bar{\psi}(x_i + a\hat{\mu})U_\mu(x_i)\psi(x_i) \tag{2.21}$$

which is invariant under the gauge transformation $\psi(x_i) \rightarrow G(x_i)\psi(x_i)$.

Thus, putting the theory on the lattice, we have to consider not only the sites, but also the links. For clarity sake, if we consider a hyper-cubical lattice of size L in d dimensions, we will have L^d sites containing as many matter fields and dL^d links containing gauge elements $U_\mu(x_i)$. It is worth noting that fermions have spin-$1/2$ and they do not have a peculiar representation on the lattice as we will see in the next section they are integrated directly in either the action or the observables using properties of the Grassmann algebra.

2.1.5 Grassmann variables and integration over fermions

Fermion fields ψ satisfy anti-commuting relations given by

$$\{\psi_i, \psi_j\} = \psi_i\psi_j + \psi_j\psi_i = 0 \tag{2.22}$$

and thus $\psi_i\psi_j = -\psi_j\psi_i$ and $\psi_i\psi_i = 0$. Such property is described by the Grassmann algebra, which has the following properties:

- Any function $f(\psi_1, \cdots, \psi_N)$ with ψ_j a Grassmann variable, is represented by the following Taylor-series

$$f(\psi_1, \cdots, \psi_N) = a + \sum_i^N b_i \psi_i + \sum_{i<j}^N c_{ij} \psi_i \psi_j + \cdots + z \psi_1 \cdots \psi_N \tag{2.23}$$

since any other term with ψ_i^k would vanish for $k > 1$.

- The derivative of Grassmann variables are given by

$$\frac{\partial}{\partial \psi_j} \psi_i = \delta_{ij} \quad \text{and} \quad \frac{\partial}{\partial \psi_k} \psi_i \psi_j = \delta_{ik} \psi_j - \delta_{jk} \psi_i, \tag{2.24}$$

where δ_{ij} is the Kronecker delta.

- The integration rules for Grassmann variables are

$$\int d\psi_j \, \psi_i = \delta_{ij}, \quad \int d\psi_i = 0 \quad \text{and} \quad \int d\psi_k \, \psi_i \psi_j = \delta_{jk} \psi_i - \delta_{ik} \psi_j. \tag{2.25}$$

- For physical gauge theories we encounter Gaussian integrals of the form

$$\int (d\bar{\psi} d\psi)^N e^{-\psi^\dagger M \psi} = \det(M), \tag{2.26}$$

where M is a matrix of size $N \times N$ and $(d\bar{\psi} d\psi)^N = \prod_i^N d\bar{\psi}_i d\psi_i$, and of the form

$$\int (d\bar{\psi} d\psi)^N \psi_i \bar{\psi}_j e^{-\bar{\psi} M \psi} = M_{ij}^{-1}. \tag{2.27}$$

The fermions can be then integrated over exactly since they appear quadratically in the action to yield

$$Z_f = \int (d\bar{\psi} d\psi)^N e^{-\bar{\psi}(\gamma^\mu D_\mu - m)\psi} = \det(\gamma^\mu D_\mu - m) \tag{2.28}$$

and

$$\langle \psi_i \bar{\psi}_j \rangle = \frac{1}{Z_f} \int (d\bar{\psi} d\psi)^N \psi_i \bar{\psi}_j e^{-\bar{\psi}(\gamma^\mu D_\mu - m)\psi} = (\gamma^\mu D_\mu - m)_{ij}^{-1} \tag{2.29}$$

where we have considered the free Dirac equation.

2.1.6 Observables

Only expectation values of gauge invariant quantities yield a non-zero value as for example operators depicted in Fig. 2.1.

Wilson loop

Any product of gauge-links on a path Γ on the lattice, i.e. given by a sequence of neighboring sites $x_1, x_2 = x_1 + \hat{\mu}, \ldots, x_{i_N}$,

$$U_\Gamma(x_N, x_1) = U_\lambda(x_N - \hat{\lambda}) \ldots U_\nu(x_1 + \hat{\mu}) U_\mu(x_1) \tag{2.30}$$

transforms as

$$U_\Gamma(x_N, x_1) \to G(x_N) U_\Gamma(x_N, x_1) G^{-1}(x_1). \tag{2.31}$$
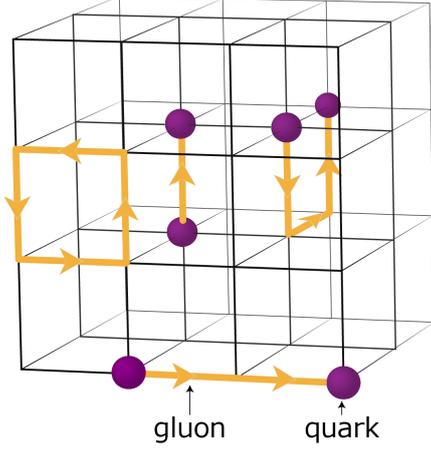
Figure 2.1: Schematic representation of a 3D lattice with gauge invariant quantities.

Thus the trace of a closed loop Λ, as depicted in Fig.2.1, which starts from a point x_i and returns to that point, is a gauge invariant quantity

$$U_\Lambda(x_i) \to G(x_i)U_\Lambda(x_i)G^{-1}(x_i) \quad \text{and} \quad W_\Lambda(x_i) = \text{Tr}\, U_\Lambda(x_i) \to W_\Lambda(x_i). \tag{2.32}$$

The quantity W_Λ is called Wilson loop and it is the most common pure gauge observable on the lattice.

Then, we can define the observable

$$W(\Lambda) = \left\langle \sum_{x_i} W_\Lambda(x_i) \right\rangle = \frac{1}{Z} \int \left[\prod_{\mu,x_i} dU_\mu(x_i) \right] \left[\sum_{x_i} \text{Tr}\, U_\Lambda(x_i) \right] e^{-S(U,g,m)}, \tag{2.33}$$

known as Wilson loop. More generally one may consider

$$W(\Lambda_1,\Lambda_2,\ldots,\Lambda_N) = \langle W_{\Lambda_1} W_{\Lambda_2} \ldots W_{\Lambda_N} \rangle \tag{2.34}$$

the correlation between several Wilson's loop factors; or less generally the rectangular Wilson loop

$$W(R,T) = \langle W_{\Lambda_{R,T}} \rangle \tag{2.35}$$

where R and T are the width and the height of a rectangular loop $\Lambda_{R,T}$, respectively.

Fermionic observables

Another gauge invariant quantity is a Wilson line joining a quark and an anti-quark field introduced in Eq. 2.21. In general, a path Γ starting from a matter field and ending at an anti-matter field is a gauge invariant quantity,

$$\Psi_\Gamma^{\lambda\rho}(x_N,x_1) = \bar{\psi}^\lambda(x_N)U_\Gamma(x_N,x_1)\psi^\rho(x_1), \tag{2.36}$$

where color indexes have been contracted. A good observable is then the propagation of field for a distance $\Delta x = x_N - x_1$ along all the possible paths Γ

$$\Psi^{\lambda\rho}(\Delta x) = \left\langle \sum_\Gamma \sum_{x_i} \Psi_\Gamma^{\lambda\rho}(x_i + \Delta x, x_i) \right\rangle. \tag{2.37}$$

For fermions we obtain from Eq. (2.29)

$$\Psi^{\lambda\rho}(\Delta x) = \frac{1}{Z} \int \left[\prod_{\mu,x_i}^{4V} dU_\mu(x_i) \right] \left[\sum_\Gamma \sum_{x_i} U_\Gamma(x_i + \Delta x, x_i)_{b,a} (\gamma^\mu D_\mu(U) - m)^{-1}_{x_i,x_i+\Delta x,a,b,\rho,\lambda} \right] e^{-S(U,g,m)}.$$

(2.38)

Levi-Civita contracted fields in special groups

A gauge transformation contracted by the completely anti symmetric Levi-Civita tensor gives

$$G^{j_1}_{i_1} G^{j_2}_{i_2} \ldots G^{j_N}_{i_N} \varepsilon^{i_1,i_2,\ldots,i_N} = \varepsilon^{j_1,j_2,\ldots,j_N} \det G = \varepsilon^{j_1,j_2,\ldots,j_N}$$

(2.39)

where G is an element of the group $SU(N)$, i and j are matrix indexes and $\det G \equiv 1$. It follows then that N matter fields or gauge-links contracted by the Levi-Civita tensor are invariant:

$$\psi^{j_1}_{i_1}(x_i) \ldots \psi^{j_N}_{i_N}(x_i) \varepsilon^{i_1,\ldots,i_N} \quad \rightarrow \quad G^{i_1}_{k_1} \ldots G^{i_N}_{k_N} \psi^{j_1}_{i_1}(x_i) \ldots \psi^{j_N}_{i_N}(x_i) \varepsilon^{k_1,\ldots,k_N} =$$
$$\psi^{j_1}_{i_1}(x_i) \ldots \psi^{j_N}_{i_N}(x_i) \varepsilon^{i_1,\ldots,i_N} \underbrace{\det G(x_i)}_{1}.$$

(2.40)

and similarly for the gauge-links. For a baryon propagator we thus have

$$B(\Delta x) = \langle \sum_{x_i} \varepsilon^{i_1,\ldots,i_N} (\bar{\psi})^{j_1}_{i_1}(x_i + \Delta x) \ldots (\bar{\psi})^{j_N}_{i_N}(x_i + \Delta x) \psi^{k_1}_{j_1}(x_i) \psi^{k_N}_{j_N}(x_i) \varepsilon_{k_1,\ldots,k_N} \rangle.$$

(2.41)

2.2   Wilson gauge action for $SU(N)$ gauge theories

The Yang-Mills action for $SU(N)$ gauge field theories can be written as

$$S_{YM} = \frac{1}{4} \int d^4x F^a_{\mu\nu} F^{\mu\nu}_a = \frac{1}{4} \lim_{a\to 0} \sum_x a^4 \text{Tr}[F^2_{\mu\nu}]$$

(2.42)

where the minimal lattice equivalent of $F_{\mu\nu}$, i.e. a gauge invariant quantity with the same dimensionality, is the trace of the plaquette

$$\text{Tr}\left[U_{\mu\nu}(x_i) \right] = \text{Tr}\left[U_\mu(x_i)\, U_\nu(x_i + a\hat{\mu})\, U_{-\mu}(x_i + a\hat{\mu} + a\hat{\nu})\, U_{-\nu}(x_i + a\hat{\nu}) \right],$$

(2.43)

as depicted in Fig. 2.2. We will use the gauge invariant $U_{\mu\nu}(x_i)$ for the construction of the lattice gauge action.
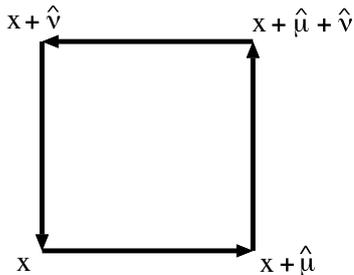
Figure 2.2: Schematic representation of the plaquette, $U_{\mu\nu}(x)$.

Taking the Taylor expansion of the shifted gauge-links, i.e.

$$U_\nu(x_i + a\hat{\mu}) = \left(1 + a\partial_\mu + O(a^2) \right) U_\nu(x_i) = \left(1 + iga^2 \partial_\mu A_\nu(x_i) + O(a^3) \right) U_\nu(x_i)$$

(2.44)

17

where $U_\nu(x_i) = e^{igaA_\nu(x_i)}$, one can prove that

$$S = \sum_{x, \mu < \nu} \frac{2N}{g^2} \left(1 - \frac{1}{N} \text{Re Tr } U_{\mu\nu}(x) \right) \tag{2.45}$$

gives at $\mathcal{O}(a)$ the Yang-Mills discretized action of Eq. (2.42). The action in Eq. (2.45) is commonly referred to as the Wilson lattice gauge action. Several remarks are in order here.

- we note that, in analogy to the continuum case, S_P takes its lowest value, namely 0, when the gauge field is a free field having $U = 1$, while it grows as U_P deviates from 1.

- the coupling constant determines the strength of the plaquette excitations: large g suppresses effects of small variations; the opposite happens for small g. It is common to define

$$\beta = \frac{2N}{g^2} \tag{2.46}$$

and thus g plays the same role of the temperature in a thermodynamic average where $\beta = 1/(kT)$.

- the Wilson gauge action produces the Yang-Mills action plus $\mathcal{O}(a)$ corrections. Better discretization schemes have been developed to cancel these corrections and produce actions free of $\mathcal{O}(a)$ effects. These are referred to as improved actions. In our simulations we use the Iwasaki improved gauge action [59], which reads as

$$S_g = \beta \left(1 - \frac{1}{6N} \sum_{x=1}^{N} \left(b_0 \sum_{\substack{\mu,\nu=1 \\ \mu<\nu}}^{4} \text{Re Tr} U_{\mu\nu}^{1\times1}(x) + b_1 \sum_{\substack{\mu,\nu=1 \\ \mu\neq\nu}}^{4} \text{Re Tr} U_{\mu\nu}^{1\times2}(x) \right) \right), \tag{2.47}$$

with the bare inverse gauge coupling $\beta = 6/g^2$, $b_1 = -0.331$ and $b_0 = 1 - 8b_1$. Compared to the Wilson gauge action of Eq. (2.45), the Iwasaki action has $\mathcal{O}(a)$-improvement and employs additionally to the plaquette $U_{\mu\nu}^{1\times1}(x)$ a rectangular Wilson loop, $U_{\mu\nu}^{1\times2}(x)$, of size 1×2 lattice units to correct $\mathcal{O}(a)$ effects.

2.3 Renormalization and lattice spacing

We discuss in this section the renormalization. Suppose that a physical quantity q_i bearing dimension d_i in units of the lattice spacing is calculated in the lattice regularized version of the theory. Adopting, as we do, natural units, it has a dimension of $[q_i] = [a]^{d_i} = [mass]^{-d_i}$. The result of the calculation will take the form

$$q_i = a^{d_i} f_i(g), \tag{2.48}$$

where the dependence on the lattice spacing is trivial and the whole content of the theory is expressed by the mathematical functions f_i of the dimensionless coupling constant g. If some of our observables, say q_l, are for instance correlation lengths, they will be given by

$$q_l = a \, f_l(g), \tag{2.49}$$

where $f_l(g)$ measures the correlation length in units of the lattice spacing.

It is now obvious that merely letting $a \to 0$ will not produce a meaningful continuum limit. The coupling constant g must be changed at the same time as the limit $a \to 0$ is taken, in such a way that the observables approach a well-defined limit when one recovers the continuum field theory. The fact that this is possible and, therefore, that the lattice theory admits a continuum limit is *a priori* not evident. Indeed, the existence of a continuum limit puts very stringent requirements on the theory. Let us first discuss this by examining the correlation length.

To make q_l approach a finite limit by suitably changing g as $a \to 0$, there must be a critical value g_c of g such that $f_l(g) \to \infty$ as $g \to g_c$. Values of the coupling constant, for which the correlation length goes to infinity are known as critical points. Moreover, as $g \to g_c$, all the functions $f_i(g)$ must tend either to infinity or zero (according to whether d_i is negative or positive). For g approaching g_c, $a \to 0$ in such a way so all the observables q_i tend to constant values. If this happens then we say that the critical point defined by $g = g_c$ is a scaling critical point.

From the above discussion it is clear that the passage to the continuum limit requires a definite functional relationship between a and g; in the literature it is generally known as Callan-Symanzik equation, because the inverse of a is proportional to the characteristic momentum scale at which the lattice coupling g is defined. This could be obtained by demanding that any of the observables remains strictly constant throughout the process of renormalization. Then, from Eq. (2.48) we would infer

$$a = a(g) = q_i^{1/d_i} f_i^{-1/d_i}(g). \tag{2.50}$$

It is useful, however, to allow more freedom in relating a and g. We therefore express the functional relationship between a and g in the form

$$a = \Lambda_l \, F(g), \tag{2.51}$$

where F is a dimensionless function embodying the correct scaling behavior of a as $g \to g_c$ and Λ_l is a scale parameter with length dimension. Λ_l does not have a direct physical significance and in general it will depend on the scheme of renormalization.

The scaling properties of the critical point now manifest themselves in the fact that all the functions $f_i(g)$ must behave as

$$f_i(g) \sim c_i \, F^{-d_i}(g) \tag{2.52}$$

for $g \to g_c$, with c_i being constants. Substituting Eqs. (2.51) and (2.52) into Eq. (2.48), we find that the continuum values of the observables q_i are given by

$$q_i = c_i \, \Lambda_l^{d_i}, \tag{2.53}$$

i.e. all physical observables are expressed in terms of the scale parameter Λ_l. Any observable might then be used to establish the scale and, after replacing Λ_l, all other observables are then expressed in terms of this one observable.

If $g \neq 0$, non-perturbative methods will be required to determine the functional relationship of Eq. (2.51). However, for non-Abelian quantum gauge theories it has been shown [60, 61] that $g_c = 0$ has the properties of a scaling critical point. Note that $g_c = 0$ is the value of the coupling constant required by asymptotic freedom. The functional form of Eq. (2.51) can then be determined by pertur-

bative arguments and this leads to [62, 63]

$$F(g) = \left(g^2 b_0\right)^{-\frac{b_1}{2b_0^2}} e^{-\frac{1}{2b_0 g^2}} \left(1 + \mathcal{O}(g^2)\right) \tag{2.54}$$

where

$$b_0 = \frac{11}{3} \frac{N}{16\pi^2}, \qquad b_1 = \frac{34}{3} \left[\frac{N}{16\pi^2}\right]^2 \tag{2.55}$$

for a pure $SU(N)$ theory. Using Eqs. (2.46), (2.51) and (2.54), we can write

$$a(\beta) \simeq \Lambda_l \left(\frac{\beta}{2Nb_0}\right)^{\frac{b_1}{2b_0^2}} e^{-\frac{\beta}{4Nb_0}}. \tag{2.56}$$

From this discussion it emerges that, in the case of QCD, we are interested in $\beta \to \infty$. Therefore, the strong coupling results obtained for small values of β must be extrapolated to at least a value of β large enough that scaling towards the continuum limit is seen to take place.

2.4 Wilson fermions

We introduce now the discretization of the Dirac fermionic action

$$\mathcal{L}_f = \bar{\psi} \left(\gamma^\mu \partial_\mu + m_q\right) \psi. \tag{2.57}$$

A simple discretization approach given by

$$\partial_\mu \psi(x) = \frac{U_\mu(x_i)\psi(x_i + a\hat{\mu}) - \psi(x)}{a} \tag{2.58}$$

suffers from the fermion doubling problem [64] i.e. in the continuum limit, it gives rise to two modes per each dimension. The simplest way to remove the unwanted modes is to give them an extra mass by adding to the action a term irrelevant in the continuum limit. Such a term is $a\bar{\psi}\partial^2\psi$ [65], which leads to

$$S_{WF} = m_q \sum_n \bar{\psi}(x_i)\psi(x_i) + \frac{1}{2a}\sum_{n,\mu} \bar{\psi}(x_i)\gamma^\mu \left[U_\mu(x_i)\psi(x_i + a\hat{\mu}) - U_\mu^\dagger(x_i - a\hat{\mu})\psi(x_i - a\hat{\mu})\right] +$$
$$+ \frac{r}{2a}\sum_{n,\mu} \bar{\psi}(x_i) \left[U_\mu(x_i)\psi(x_i + a\hat{\mu}) - 2\psi(x_i) - U_\mu^\dagger(x_i - a\hat{\mu})\psi(x_i - a\hat{\mu})\right]. \tag{2.59}$$

In this formulation, it is clear that the added term acts as a mass term that is non-zero even when the "bare mass" is set to zero. The coefficient of the mass term $\bar{\psi}(x_i)\psi(x_i)$ is $m_q + 4r/a$. In the continuum the doublers have a large mass $m_q + 4r/a$, and decouple; however in an interacting theory the value of r is a function of β and needs to be tuned.

Usually the Wilson fermion action is rewritten in the form

$$S_{WF}(U, \bar{\psi}, \psi) = \sum_{i,j} \bar{\psi}(x_i) D_{ij}(U) \psi(x_j), \tag{2.60}$$

where

$$D_{ij} = \delta_{ij} - \kappa \sum_\mu \left[(r - \gamma_\mu)U_\mu(x_i)\delta_{x_i + a\hat{\mu}, x_j} + (r + \gamma_\mu)U_\mu^\dagger(x_i - a\hat{\mu})\delta_{x_i - a\hat{\mu}, x_j}\right]. \tag{2.61}$$

is the Wilson Dirac operator where we have defined

$$\kappa = \frac{1}{2a} \frac{1}{m_q + 4r/a} \qquad\qquad \psi \to \frac{\psi}{\sqrt{2\kappa}}. \tag{2.62}$$

where κ is known as the hopping parameter since it determines the strength of a fermion coupling to an adjacent site. The case $r = 1$ is special as the factor $(1 \pm \gamma_\mu)$ is a projection operator

$$\left(\frac{1 \pm \gamma_\mu}{2} \right)^2 = \frac{1 \pm \gamma_\mu}{2} \qquad\qquad \mathrm{Tr}\left(\frac{1 \pm \gamma_\mu}{2} \right) = 2. \tag{2.63}$$

The hopping parameter κ is related to the inverse of bare quark mass by

$$m_q a = \frac{1}{2}\left(\frac{1}{\kappa} - 8r \right) = \frac{1}{2}\left(\frac{1}{\kappa} - \frac{1}{\kappa_c} \right), \tag{2.64}$$

where the critical hopping parameter has been defined as $k_c = 1/8r$. We now have a new parameter, κ, beside β, on which the dynamics of the lattice depends and a parameter, κ_c, which we have to tune in order to find the bare mass of the quarks. Additionally, κ has to be smaller than κ_c in order to keep the mass positive.

One of the disadvantages of Wilson fermions is the explicit breaking of chiral symmetry at $\mathscr{O}(a)$. Wilson fermions can be improved by including a clover term $Q_{\mu\nu}$, given by

$$\begin{aligned}
Q_{\mu\nu}(x) =\, & U_\mu(x)\, U_\nu(x+\hat\mu)\, U_\mu(x+\hat\nu)^\dagger\, U_\nu(x)^\dagger +\\
& U_\nu(x)\, U_\mu(x-\hat\mu+\hat\nu)^\dagger\, U_\nu(x-\hat\mu)^\dagger\, U_\mu(x-\hat\mu) +\\
& U_\mu(x-\hat\mu)^\dagger\, U_\nu(x-\hat\mu-\hat\nu)^\dagger\, U_\mu(x-\hat\mu-\hat\nu)\, U_\nu(x-\hat\nu) +\\
& U_\nu(x-\hat\nu)^\dagger\, U_\mu(x-\hat\nu)\, U_\nu(x-\hat\nu+\hat\mu)\, U_\mu(x)^\dagger\,.
\end{aligned} \tag{2.65}$$

The so-called clover improved Wilson operator is

$$\begin{aligned}
D\psi(x) =\, & \left((m+4)I_{\mathscr{S}} \otimes I_{\mathscr{C}} - \frac{c_{sw}}{32} \sum_{\mu,\nu=0}^{3} (\gamma_\mu \gamma_\nu) \otimes \big(Q_{\mu\nu}(x) - Q_{\nu\mu}(x) \big) \right) \psi(x)\\
& - \frac{1}{2}\sum_{\mu=0}^{3} \big((I_{\mathscr{S}} - \gamma_\mu) \otimes U_\mu(x) \big)\, \psi(x+\hat\mu)\\
& - \frac{1}{2}\sum_{\mu=0}^{3} \big((I_{\mathscr{S}} + \gamma_\mu) \otimes U_\mu^\dagger(x-\hat\mu) \big)\, \psi(x-\hat\mu)\,,
\end{aligned} \tag{2.66}$$

where appropriate values of the Sheikoleslami-Wohlert improvement coefficient c_{sw} [66] reduce the discretization error from $\mathscr{O}(a)$ to $\mathscr{O}(a^2)$. This defines the Wilson Dirac operator that will be used in what follows. Compared to Eq. (2.61) we have taken $r = 1$, $m = 1/(2\kappa) - 4$, and we have explicitly indicated the Kronecker products between spin and color space. For clarity, the vector ψ is defined in Dirac space

$$\mathscr{D} = \mathscr{V} \otimes \mathscr{S} \otimes \mathscr{C}, \tag{2.67}$$

where $\mathscr{V}$ is the lattice volume defined in Eq. (2.13), $\mathscr{S}$ denotes spin space and $\mathscr{C}$ color space.

The Wilson Dirac operator satisfies the relation

$$\Gamma_5 D = D^\dagger \Gamma_5, \tag{2.68}$$

referred to as Γ_5-hermiticity, where

$$\Gamma_5 = I_{\mathscr{V}} \otimes \gamma_5 \otimes I_{\mathscr{C}} \tag{2.69}$$

acts on space $\mathscr{V}_s$ defined in Eq. (2.67) as a linear transformation of the spin degrees of freedom at each lattice site. In this thesis we use $\gamma_5 = \gamma_0 \gamma_1 \gamma_2 \gamma_3$, which is diagonal in spin space

$$\gamma_5 = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}. \tag{2.70}$$

Thanks to Γ_5-hermiticity, the symmetrized Wilson Dirac operator $H = \Gamma_5 D$ is hermitian (and indefinite), such that we have

$$H = \Gamma_5 D = V \Lambda V^\dagger, \tag{2.71}$$

where the diagonal matrix Λ contains the eigenvalues λ_j of H_W (which are all real) and the unitary matrix V the corresponding eigenvectors.

2.5   Quark mass determination

After introducing the Wilson fermion discretization, it is important to understand what is the mass of the quarks. The quark mass obtained in Eq. (2.64) is a bare mass and the calculation of the renormalized quark masses is an active subject of investigation in lattice QCD. An accurate determination of these parameters is in fact extremely important, for both phenomenological and theoretical applications. The values of quark masses cannot be directly measured in experiments because quarks are confined inside the hadrons and they are not physical states. The quark masses can be determined by for example comparing the mass of a hadron with the corresponding experimental value.

2.5.1   Ward identities

Non-perturbative definitions of quark masses are provided by the chiral Ward identities (WI) of QCD [67, 68]. These definitions allow us to express the renormalized quark mass, in a given scheme and at a given renormalization scale, in terms of lattice renormalization constants and bare quantities. Before we address the determination of quark masses, let us introduce the meaning of Ward identities. In the path integral expression for the partition function Eq. (2.12), the field variables, ϕ, are integrated. Then, a transformation of these variables in general gives the same result,

$$\langle O \rangle = \frac{1}{Z} \int \mathscr{D}\phi \; O(\phi) \; e^{-S_E(\phi,g,m)} = \frac{1}{Z'} \int \mathscr{D}\phi' \; O(\phi') \; e^{-S_E(\phi',g,m)} \quad \text{with} \quad Z' = \int \mathscr{D}\phi' \; e^{-S_E(\phi',g,m)}. \tag{2.72}$$

Exceptions are anomalous transformations, which affect the integration measure. Ward identities express the invariance of the partition function and of expectation values under a transformation of the field variables. Studying the effect of a local, unitary, infinitesimal symmetry transformation of the fermion fields, $\phi \to \phi + \delta\phi$, the result has to be invariant and we obtain Ward identities of the form

$$\langle \delta O \rangle = \langle O \delta S \rangle, \tag{2.73}$$

where δO and δS denote the linear change of the operator O and the action S under the transformation. The simplest case with $O = \mathbf{1}$ gives

$$\langle \delta S \rangle = 0, \tag{2.74}$$

which leads to relations analogous to the classical Noether conservation laws. The Ward identity, which is most of interest in calculation of quark mass is the one referred to as axial Ward identity (AWI). We consider the transformation,

$$\psi(x) \quad \longrightarrow \quad \psi'(x) = (\mathbf{1} + i\varepsilon(x)\gamma_5)\,\psi(x) \tag{2.75}$$

$$\bar{\psi}(x) \quad \longrightarrow \quad \bar{\psi}'(x) = \bar{\psi}(x)\,(\mathbf{1} + i\varepsilon(x)\gamma_5), \tag{2.76}$$

where $\varepsilon(x)$ is a function that vanishes smoothly outside some bounded region. In the continuum, we are led to

$$\langle (\partial_\mu A_\mu^a(x))\, O \rangle = 2m \langle P^a(x)\, O \rangle, \tag{2.77}$$

where m is the quark mass. In this equation, $A_\mu^a(x)$ and $P^a(x)$ are the isovector axial current and the pseudo-scalar density. The corresponding lattice operators are

$$A_\mu^a(x_i) \quad = \quad \frac{1}{2}\bar{\psi}(x_i)\gamma_\mu\gamma_5\tau^a\psi(x_i) \tag{2.78}$$

$$P^a(x_i) \quad = \quad \frac{1}{2}\bar{\psi}(x_i)\gamma_5\tau^a\psi(x_i). \tag{2.79}$$

The observable O is an operator chosen such that $\langle P^a(x)\, O \rangle$ does not vanish, with the restriction that its support does not include the point x where $A_\mu^a(x)$ and $P^a(x)$ are located. Eq. (2.77) implies that the isovector axial current is conserved for a chiral theory where $m = 0$. Thus, it is referred to as the partially conserved axial current relation (PCAC).

2.5.2   Renormalized quark mass

On the lattice we expect that after renormalization, the lattice expectation values approach the continuum relation (2.77), such that

$$\langle (\partial_\mu A_\mu^{(r)a}(x_i))\, O \rangle = 2m^{(r)} \langle P^{(r)a}(x_i)\, O \rangle + \mathcal{O}(a), \tag{2.80}$$

where the superscript (r) expresses the fact that renormalized quantities are being used. The renormalized operators are

$$A_\mu^{(r)a}(x_i) \quad = \quad Z_A A_\mu^a(x_i) + \mathcal{O}(a), \tag{2.81}$$

$$P_\mu^{(r)a}(x_i) \quad = \quad Z_P P^a(x_i) + \mathcal{O}(a) \tag{2.82}$$

and defining the unrenormalized quark mass

$$m_{\mathrm{PCAC}} = \frac{\langle (\partial_\mu A_\mu^a(x_i))\, O \rangle}{2\langle P^a(x_i)\, O \rangle}, \tag{2.83}$$

referred to as PCAC mass, we obtain

$$m^{(r)} = \frac{Z_A}{Z_P}\, m_{\mathrm{PCAC}} + \mathcal{O}(a). \tag{2.84}$$

23

where the $\mathcal{O}(a)$ effects are vanishing in case of improved actions having $\mathcal{O}(a^2)$ corrections. Now we can fix the renormalization factor Z_A/Z_P evaluating the PCAC mass using different operators O. These operators probing the PCAC relation need to have the quantum numbers of P^a and to leave the renormalization factor unchanged. Usually, P^a itself is used as operator in conjunction with other interpolator coupling to the pion. The operative definition of PCAC mass that we use is

$$m_{\text{PCAC}} = \frac{\langle (\partial_t A_0^1(\vec{x},t)) \, P^1(\vec{0},0) \rangle}{2\langle P^1(\vec{x},t) \, P^1(\vec{0},0) \rangle} = \frac{\partial_t \langle \text{Tr}(\bar{\psi}(\vec{x},t)\gamma_0\gamma_5\tau_1 \psi(\vec{x},t) \, \bar{\psi}(\vec{0},0)\gamma_5\tau_1\psi(\vec{0},0)) \rangle}{2\langle \text{Tr}(\bar{\psi}(\vec{x},t)\gamma_5\tau_1\psi(\vec{x},t) \bar{\psi}(\vec{0},0)\gamma_5\tau_1\psi(\vec{0},0)) \rangle}, \tag{2.85}$$

where the expectation value at the denominator is equivalent to the pion propagation from a point $(\vec{0},0)$ to a point $(\vec{x},t)$ in the space time.

2.6 Twisted mass fermions

The continuum QCD action is invariant under chiral symmetry. Two degenerate quark flavors of mass m_q, are invariant under the chiral rotation

$$\psi_2 = \begin{bmatrix} \psi_u \\ \psi_d \end{bmatrix} \rightarrow \psi_2' = e^{i\frac{\omega}{2}\Gamma_5 \otimes \tau_3} \begin{bmatrix} \psi_u \\ \psi_d \end{bmatrix} \quad \text{when} \quad m_q \rightarrow m_q e^{i\frac{\omega}{2}\Gamma_5 \otimes \tau_3} \tag{2.86}$$

where $\tau_3 = \text{diag}(1,-1)$ is the third Pauli matrix and acts in flavor space $\mathscr{F}$. Under this transformation the continuum Dirac operator in Eq. (2.57) takes the form

$$\begin{aligned} \mathscr{L}_f &= \bar{\psi}_2' \left(\gamma^\mu \partial_\mu + m_q' + i\mu\gamma_5\tau_3 \right) \psi_2' \\ &\equiv \bar{\psi}_2' \left(\gamma^\mu \otimes \partial_\mu \otimes I_{\mathscr{F}} + m_q' I_{\mathscr{S}} \otimes I_{\mathscr{C}} \otimes I_{\mathscr{F}} + i\mu\gamma_5 \otimes I_{\mathscr{C}} \otimes \tau_3 \right) \psi_2' \end{aligned} \tag{2.87}$$

where we used the identifications

$$m_q' = m_q\cos\omega\,, \quad \mu = m_q\sin\omega \quad \text{and} \quad m_q = \sqrt{m_q'^2 + \mu^2}. \tag{2.88}$$

This form is completely equivalent to the standard one, except that certain discrete symmetries look more complicated. The mass parameter μ is referred to as the *twisted mass*, m_q' as the quark mass and m_q as the *polar mass*. ψ_2' fields are referred to as the *twisted basis*. The relationship between the rotation angle and the mass terms is

$$\omega = \arctan\frac{\mu}{m_q'}. \tag{2.89}$$

In addition to the standard case $\omega = 0$, a special value $\omega = \pi/2$ can be identified for which the quark mass is given entirely by the twisted mass, this particular value is referred to as *maximal twist*.

2.6.1 Degenerate twisted mass operator

The twisted mass QCD becomes non-trivial on the lattice since Wilson fermions explicitly break chiral symmetry. The lattice discretization of the twisted mass Dirac operator in Eq. (2.87) is straightforward. We denote with D the Wilson operator introduced in Eq. (2.66) and with $\mu \in \mathbb{R}$ the twisted mass parameter. We use $\tau_3 = \text{diag}(1,-1)$ diagonal in flavor space. The degenerate twisted mass

operator discretized on the lattice, $D_D(\mu)$, is then given by

$$D_D(\mu) = (D \otimes I_2) + i\bar{\mu}(\Gamma_5 \otimes \tau_3) \quad \longrightarrow \quad D_D(\mu)\psi_2 = \begin{bmatrix} D_{TM}(\mu) & 0 \\ 0 & D_{TM}(-\mu) \end{bmatrix} \begin{bmatrix} \psi_u \\ \psi_d \end{bmatrix}$$
$$= \begin{bmatrix} D\psi_u \\ D\psi_d \end{bmatrix} + \begin{bmatrix} +i\mu\Gamma_5\psi_u \\ -i\mu\Gamma_5\psi_d \end{bmatrix}, \qquad (2.90)$$

where

$$D_{TM}(\mu) = D + i\mu\Gamma_5 \qquad (2.91)$$

is what we refer to as the twisted mass Wilson operator, which is restricted to the space $\mathscr{D}$, cf. Eq. (2.67). Properties of the twisted mass fermions discretization have been studied in Refs. [69–71] and fully reviewed for the degenerate twisted mass operator in Ref. [46]. An important property of twisted mass fermions is the automatic $\mathscr{O}(a)$-improvement for simulations at maximal twist, i.e. when the twist angle ω, cf. Eq. (2.89), is equal to $\pi/2$.

2.6.2 Isospin symmetry breaking

In the degenerate twisted mass operator the twist is applied linearly in the flavor space of the up (u-) and down (d-) quark. Then the twisted mass term acts with a positive shift given by $i\mu\Gamma_5$ on the u-quark operator and with a negative shift given by $-i\mu\Gamma_5$ on the d-quark operator. The twisted mass term does not change the up and down quark masses, cf. Eq. (2.88), but breaks explicitly the isospin symmetry between the u- and the d-quark, which was preserved by the Wilson operator. The isospin symmetry breaking is an *unwanted* lattice artifact, which is removed in the continuum limit. An important consequence of the isospin symmetry breaking is that the neutral pion is lighter than the charged pion. This slows down or even prohibits simulations for light quark masses close to the physical value for certain values of the lattice spacing.

The neutral to charge pion mass splitting can be estimated in chiral perturbation theory. To first order one finds [72]

$$m_{\pi^0}^2 - m_{\pi^\pm}^2 \simeq 4c_2 a^2 \qquad (2.92)$$

where c_2 depends on low energy constants and it is found to be negative for twisted mass fermions [73]. This means that the neutral pion mass may vanish for finite light quark masses triggering a phase transition into a non-physical phase. As a consequence, simulations at the physical point of twisted mass fermion without a clover term would require a very fine lattice spacing, i.e. $a < 0.062$, and therefore imply a very large volume, as depicted in Fig. 2.3.

The employment of a clover term into the action comes in our rescue. As shown by the purple points in Fig. 2.3, the critical twisted mass value for the light quarks is significantly reduced as compared to simulations without this term. By introducing a clover term the quadratic lattice artifacts in the isospin symmetry breaking are suppressed, e.g. in case of the neutral pion mass as shown in Refs.[6, 74–76]. The clover parameter is usually set by using an estimate from 1–loop [77] tadpole boosted perturbation theory given by

$$c_{SW} \cong 1 + 0.113(3)\frac{g_0^2}{P}, \qquad (2.93)$$

with P being the plaquette expectation value. This allows simulations at the physical point with a

value of the lattice spacing around $a = 0.1$ fm or even larger [76]. Thus, differently from the Wilson operator, the clover term is employed in reducing the isospin breaking effects since the discretization effects are already $\mathcal{O}(a^2)$ improved being at maximal twist.
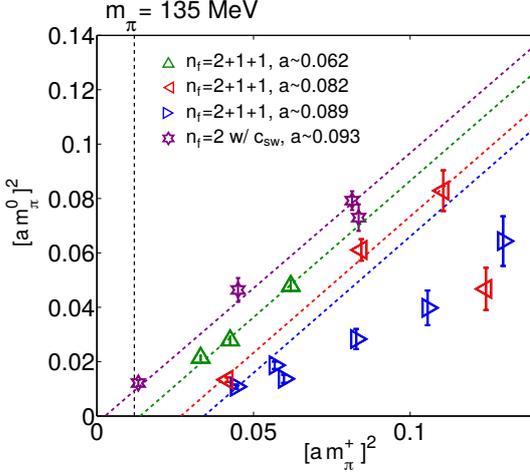
Figure 2.3: The squared neutral pion mass is plotted against the squared charged pion mass in units of the lattice spacing a. We show results for N_f=2+1+1 ETMC ensembles with no clover term i.e. taking $c_{SW} = 0$ and for the N_f=2 ensemble with clover term from Ref. [76]. We use a linear fit in $a^2 m_{\pi^{\pm}}^2$ excluding ensembles with $m_{\pi^{\pm}} > 400$ MeV. The vertical line corresponds to the physical value of the charged pion mass.

2.6.3 Quark currents and renormalized quark mass

After applying the chiral transformation, the Ward quark currents introduced in Eq. (2.78) and (2.79) transform as

$$A_\mu^{\prime a} = \frac{1}{2}\bar{\psi}_2' \gamma_\mu \gamma_5 \tau^a \psi_2' \tag{2.94}$$

$$P^{\prime a} = \frac{1}{2}\bar{\psi}_2' \gamma_5 \tau^a \psi_2' \tag{2.95}$$

$$V_i^{\prime a} = \frac{1}{2}\bar{\psi}_2' \gamma_i \tau^a \psi_2'. \tag{2.96}$$

where we also include the vector current V_i^a with $i = 1, 2, 3$ being the spatial indexes. As we see in the following the vector current plays an important role in twisted mass fermions. The PCAC and the PCVC relations are [70]

$$\langle (\partial_\mu A_\mu^{\prime a}) O \rangle = 2m_q' \langle P^{\prime a} O \rangle \quad \text{with} \quad a = 1, 2 \tag{2.97}$$

$$\langle (\partial_\mu A_\mu^{\prime 3}) O \rangle = 2m_q' \langle P^{\prime 3} O \rangle + i\mu \langle \bar{\psi}_2' O \psi_2' \rangle \tag{2.98}$$

$$\langle (\partial_i V_i^{\prime a}) O \rangle = -2\mu \varepsilon_b^a \langle P^{\prime b} O \rangle \quad \text{with} \quad a, b = 1, 2 \tag{2.99}$$

$$\langle (\partial_i V_i^{\prime 3}) O \rangle = 0 \tag{2.100}$$

At zero twist then the vector current is conserved and the axial current is the same as in Eq. (2.77). At maximal twist, when $m_q' = 0$, the components $a = 1, 2$ of the axial current are conserved. As shown in Ref. [78, 79] this gives the most suitable and theoretically-sound condition for the desired automatic $\mathcal{O}(a)$-improvement. It is achieved by demanding a vanishing value for the partially conserved axial current (PCAC) quark mass, cf. Eq. (2.85). The renormalized quark mass at maximal twist is then

$$m_q^{(r)} = \frac{1}{Z_P}|\mu| + \mathcal{O}(a^2). \tag{2.101}$$

which is directly proportional to the twisted mass parameter μ with Z_P being the pseudoscalar renormalization constant.

2.6.4 Properties of the degenerate twisted mass operator

Due to the close connection to the Wilson operator, the twisted mass operator has similar properties. The equivalent to the Γ_5-hermiticity, cf. Eq. (2.68), is

$$D_{\text{TM}}(\mu)\Gamma_5 = \Gamma_5 D_{\text{TM}}^\dagger(-\mu) \tag{2.102}$$

which connects the "up", $D_{\text{TM}}(\mu)$, and the "down", $D_{\text{TM}}(-\mu)$, flavor of the operator, cf. Eq. (2.90). Since D_{TM} is non-normal, the left and right eigenvectors differ. For the twisted mass operator, the left eigenvectors of the u-quarks are connected to the right eigenvectors of the d-quarks. Let $\varphi_{j,L}^{u(d)}$ and $\varphi_{j,R}^{u(d)}$ be left and right eigenvectors of $D_{\text{TM}}((-)\mu)$, respectively, with corresponding eigenvalue $\lambda_j^{u(d)}$, then due to Γ_5-hermiticity, cf. Eq. (2.102) we have

$$\lambda_j^u = \varphi_{j,L}^{u}{}^\dagger D_{\text{TM}}(\mu)\varphi_{j,R}^{u} = \left(\varphi_{j,R}^{u}{}^\dagger (D_{\text{TM}}(\mu))^\dagger \varphi_{j,L}^{u}\right)^\dagger = \left(\varphi_{j,R}^{u}{}^\dagger \Gamma_5 D_{\text{TM}}(-\mu)\Gamma_5 \varphi_{j,L}^{u}\right)^\dagger$$
$$= \left(\varphi_{j,L}^{d}{}^\dagger D_{\text{TM}}(-\mu)\varphi_{j,R}^{d}\right)^\dagger = (\lambda_j^d)^\star. \tag{2.103}$$

Thus, eigenpairs of $D_{\text{TM}}(\mu)$ are connected to the eigenpairs of $D_{\text{TM}}(-\mu)$ with the relations

$$\lambda_j^u = (\lambda_j^d)^\star, \quad \varphi_{j,L}^u = \Gamma_5 \varphi_{j,R}^d \quad \text{and} \quad \varphi_{j,R}^u = \Gamma_5 \varphi_{j,L}^d. \tag{2.104}$$

The equivalent of the symmetrized Wilson operator, cf. Eq. (2.71), is

$$H_{\text{TM}}(\mu) = \Gamma_5 D_{\text{TM}}(\mu) = H + i\mu = V\Lambda V^\dagger + i\mu = Ve^{i\Theta}\sqrt{\Lambda^2 + \mu^2}\, V^\dagger, \tag{2.105}$$

where the diagonal matrix $e^{i\Theta}$ contains the complex phases θ_j of the eigenvalues $\lambda_j + i\mu$ and $\sqrt{\Lambda^2 + \mu^2}$ is their absolute values. For the non-symmetrized twisted mass operator $D(\mu)$, analogously to the Wilson case [80], this gives the singular value decomposition

$$D_{\text{TM}}(\mu) = \Gamma_5 Ve^{i\Theta}\sqrt{\Lambda^2 + \mu^2}\, V^\dagger = U\sqrt{\Lambda^2 + \mu^2}\, V^\dagger \tag{2.106}$$

with $U = \Gamma_5 Ve^{i\Theta}$ and V being unitary. The smallest singular value $\sqrt{\lambda_i^2 + \mu^2}$ is thus not smaller than μ, which shows that a non-zero value of μ protects the twisted mass operator $D(\mu)$ from being singular, unlike the Wilson operator where this can happen for quark mass m close to its critical value.

Similarly, for the squared twisted mass operator

$$D_{\text{TM}}^\dagger(\mu)D_{\text{TM}}(\mu) = (D^\dagger - i\mu\Gamma_5)(D + i\mu\Gamma_5) = D^\dagger D + \mu^2 = H^2 + \mu^2 \tag{2.107}$$

we have

$$H^2 + \mu^2 = V(\Lambda^2 + \mu^2)V^\dagger, \tag{2.108}$$

the eigenvalues of which are bounded from below by μ^2.

A special property of the twisted mass operator is that at maximal twist the region just above μ^2 is densely populated with the eigenvalues of the squared operator, cf. Eq. (2.107). We illustrate this in Figure 2.4, which displays a histogram of the (scaled) ensemble averaged moduli of the eigenvalues of the non-squared symmetrized even-odd reduced (or preconditioned) twisted mass operator, $\hat{H}_{\text{TM}}$. The eigenvalues of the operator are measured on an ensemble simulated at a physical value of the light quark mass doublet, which we will refer to as the *physical ensemble*. $\hat{H}_{\text{TM}}$ is obtained using an

even-odd ordering of the lattice sites such that

$$D_{\mathrm{TM}} = \begin{bmatrix} D_{oo} & D_{oe} \\ D_{eo} & D_{ee} \end{bmatrix} \tag{2.109}$$

where D_{oo} and D_{ee} are diagonal in spinor space since only first neighbor interactions are found in the Wilson operator. $\hat{H}_{\mathrm{TM}}$ is then given as $\hat{H}_{\mathrm{TM}} = \Gamma_5 \hat{D}_{\mathrm{TM}}$ with

$$\hat{D}_{\mathrm{TM}} = (D_{ee} - D_{eo}D_{oo}^{-1}D_{oe}). \tag{2.110}$$

Figure 2.4: The density of the approximated rescaled eigenvalues of the hermitian even-odd reduced twisted mass Dirac operator measured on gauge configurations of the physical ensemble $cA2.48$ (see Ref. [76]). The quark mass is given by $m_q = 3.89$ MeV in the $\overline{\mathrm{MS}}$-scheme.
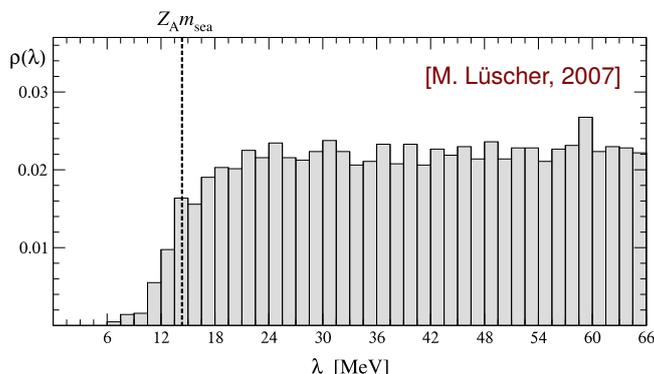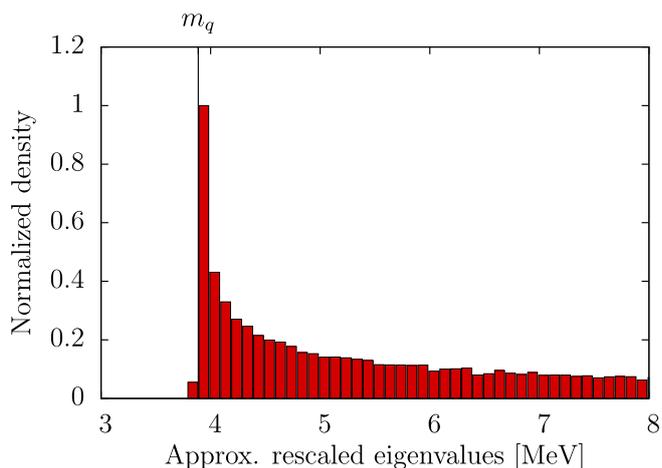
Figure 2.5: The density of small eigenvalues in the Wilson operator. Figure taken from Ref. [14], reported here for comparison to Fig. 2.4

For $c_{sw} = 0$, i.e. without the clover term, the spectrum of $\hat{D}_{\mathrm{TM}}$ is directly connected to the spectrum of the full operator D_{TM}, e.g. in the case of the small eigenvalues we have

$$\frac{\lambda_i(D_{\mathrm{TM}})}{m+4} = 1 - \sqrt{\frac{1 - \lambda_i(\hat{D}_{\mathrm{TM}})}{m+4}} \approx \frac{1}{2}\frac{\lambda_i(\hat{D}_{\mathrm{TM}})}{m+4} \tag{2.111}$$

where $\lambda_i(D_{\mathrm{TM}})$ and $\lambda_i(\hat{D}_{\mathrm{TM}})$ are respectively eigenvalues of D and $\hat{D}$. Although this relation does not hold exactly for the hermitian even-odd reduced twisted mass operator, we found that numerically this relation still holds approximately for the eigenvalues close to μ. We find that the largest relative deviation of the smallest eigenvalue to the approximated cut-off 2μ is given by $|\lambda_{min} - 2\mu|/2\mu < 0.0005$ MeV. Thus we rescale the spectrum by a factor two and relate it to the energy scale of the $\overline{\mathrm{MS}}$-scheme defined at 2 GeV. The eigenvalue density is shown in Fig. 2.4. In contrast to the spectrum of the Wilson Dirac operator, shown in Refs. [14, 81] and reported in Fig. 2.5, the density of the eigenvalues increases close to the physical quark mass.

2.6.5 Non-degenerate operator

A generalization of the degenerate twisted mass operator, cf. Eq. (2.90), is the so-called non-degenerate (ND) twisted mass operator

$$D_{\mathrm{ND}}(\bar{\mu}, \bar{\varepsilon}) = (D \otimes I_2) + i\bar{\mu} (\Gamma_5 \otimes \tau_3) - \bar{\varepsilon}(I \otimes \tau_1) = \begin{bmatrix} D_{\mathrm{TM}}(\bar{\mu}) & -\bar{\varepsilon}I \\ -\bar{\varepsilon}I & D_{\mathrm{TM}}(-\bar{\mu}) \end{bmatrix} \tag{2.112}$$

with the bare mass parameters $\bar{\mu}$ and $\bar{\varepsilon} \in \mathbb{R}$. Clearly with $\bar{\varepsilon} = 0$ and $\bar{\mu} = \mu$ we retrieve the degenerate twisted mass operator. In what follows, we consider the ND twisted mass operator for $\bar{\varepsilon} > 0$, the properties of which have been studied in Ref. [82]. Also the non-degenerate twisted mass operator is free of $\mathcal{O}(a)$ effects if the light quark sector is tuned at maximal twist [83, 84]. Due to the flavor mixing, introduced by the off-diagonal term proportional to $\bar{\varepsilon}$, the non-degenerate operator generates two renormalized quark masses connected to the bare twisted mass parameters by

$$m_s = \frac{1}{Z_P}\mu_s = \frac{1}{Z_P}\bar{\mu} - \frac{1}{Z_S}\bar{\varepsilon} \qquad \text{and} \qquad m_c = \frac{1}{Z_P}\mu_c = \frac{1}{Z_P}\bar{\mu} + \frac{1}{Z_S}\bar{\varepsilon} \tag{2.113}$$

where Z_S and Z_P are respectively the scalar and the pseudoscalar renormalization constants. We use the names m_s and m_c since we will use this operator for the strange and charm quarks doublet. The two flavors are although mixed and we cannot directly retrieve the strange and charm quark propagators. This operator is then mostly used in simulations adding to the partition function the determinant of D_{ND}.

The ND twisted mass operator is $(\Gamma_5 \otimes \tau_1)$-hermitian

$$H_{\mathrm{ND}}(\bar{\mu}, \bar{\varepsilon}) = (\Gamma_5 \otimes \tau_1) D_{\mathrm{ND}}(\bar{\mu}, \bar{\varepsilon}) = D_{\mathrm{ND}}^{\dagger}(\bar{\mu}, \bar{\varepsilon}) (\Gamma_5 \otimes \tau_1) = H_{\mathrm{ND}}^{\dagger}(\bar{\mu}, \bar{\varepsilon}) \tag{2.114}$$

thanks to the Γ_5-hermiticity of the Wilson operator. For the determinant of the ND twisted mass operator, it follows that

$$\det[D_{\mathrm{ND}}(\bar{\mu}, \bar{\varepsilon})] \in \mathbb{R} \tag{2.115}$$

thanks to the $(\Gamma_5 \otimes \tau_1)$-hermiticity. This implies the reality of the determinant but not its positiveness. The positiveness of the operator is proved for $\bar{\mu} > \bar{\varepsilon}$ [46]. In case of the heavy quark doublet, taking the physical values of the strange and charm quark masses, this is achieved if the ratio of the renormalization constants is $Z_P/Z_S > 0.85$. However, on the lattice we typically have $Z_P/Z_S \simeq 0.8$. This does not imply that the determinant is negative, but its positiveness is not ensured. Numerically, we find that for $Z_P/Z_S \simeq 0.8$ the determinant is positive with no exceptions observed.

3. Inverting the Dirac operator on the lattice

For lattice QCD calculations involving fermions we need to compute the inverse of the Dirac operator D. It is a very frequent operation, which makes LQCD simulations computationally demanding. The Dirac operator discretized on a lattice is represented by a *sparse matrix*. A sparse matrix is a matrix where the number of non-zero components is proportional to the matrix size. The stencil of the matrix, i.e. how the non-zero components are distributed, changes accordingly to the fermions discretization scheme and the boundary conditions. For Wilson fermions, as well as for twisted mass fermions, we have local and first-neighbor contributions and commonly periodic boundary conditions are applied. We depict a schematic representation of the Dirac operator in Fig. 3.1.
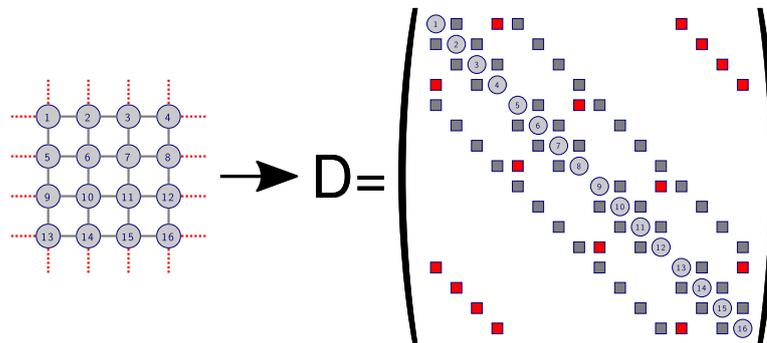
Figure 3.1: Schematic representation of the Dirac operator. On the left a 4×4 lattice with sites numbered in row-major order from 1 to 16. The gray links show inner connection between the sites, while red links show connections through the boundaries. On the right hand side a schematic matrix representation of the Dirac operator where the non-zero contributions are depicted with symbols. Each square or circle represent a 12×12 block matrix being the color-spin components per lattice site. On the diagonal of the matrix we show the local contributions for each lattice site. The gray squares are the first neighbor contributions through the inner links, while the red squares are the first neighbor contributions through the boundaries.

Even though D is a sparse matrix, its inverse is not. D^{-1} is a full matrix with size proportional to the lattice volume squared. Considering simulations performed currently, where we typically have lattices of size $V = 64^3 \times 128$, it would be impossible to store in memory the inverse of such a matrix, which would have size $(2 \times N_c \times N_s \times V)^2 \simeq 6 \cdot 10^{17} \simeq 600$ petabytes (PB). The fact that D is sparse can be utilized in that we always compute the inverse of the matrix applied to a vector b by solving the linear system

$$Dx = b \quad \longrightarrow \quad x = D^{-1}b. \tag{3.1}$$

In this chapter, we introduce iterative approaches for solving sparse linear systems. We will refer to x as the *solution* and to b as *right hand side* (rhs). The matrix D here indicates any non-singular matrix, even though our methods target the Dirac operator.

3.1 Iterative Krylov solvers

Iterative Krylov solvers are iterative algorithms for computing an approximate solution of the inverse of a matrix. Many methods exist and for a complete survey we refer to the book [85]. Some important features that characterize these methods are:

- The *residual*, r, and the *error*, e, of the *approximate* solution, x, are defined as

$$r = b - Dx \quad \text{and} \quad e = D^{-1}b - x = D^{-1}r. \tag{3.2}$$

While e is the *real* error of x, the residual r is the quantity we compute to *approximately* estimate the error since it does not require the knowledge of the *exact* solution $D^{-1}b$.

- The solution x is searched in the so-called Krylov subspace, $\mathscr{K}_n$,

$$x \in \mathscr{K}_n(D,b) = \left\{ b, Db, D^2b, \ldots, D^{n-1}b \right\}, \tag{3.3}$$

where n indicates the number of iterations. It can be proven that for non-singular matrices of size $N \times N$ the exact solution belongs to

$$D^{-1}b \in \mathscr{K}_N(D,b) = \left\{ b, Db, D^2b, \ldots, D^{N-1}b \right\}. \tag{3.4}$$

Therefore N is the maximum number of iterations we need to exactly solve a linear system. We need only an approximate but precise solution aiming to a number of iteration $n \ll N$.

- An initial guess x_0 can be given to the solver. The correction is then given by

$$x = x_0 + D^{-1}b - x_0 = x_0 + D^{-1}r_0 \quad \text{with} \quad r_0 = b - Dx_0. \tag{3.5}$$

The solution to $D^{-1}r_0$ is searched in the Krylov susbspaces $\mathscr{K}_i(D,r_0)$.

- The Krylov subspace is constructed iteratively by applying the operator D

$$\mathscr{K}_{i+1} = \mathscr{K}_i \cup D\mathscr{K}_i[i], \tag{3.6}$$

where $\mathscr{K}_i[i]$ is the last, i-th, element of $\mathscr{K}_i$. Also the solution is constructed iteratively as $x_{i+1} = x_i + \Delta x_{i+1}$, where $x_i \in \mathscr{K}_i$ and $\Delta x_{i+1} \in \mathscr{K}_{i+1}$.

- Most of the Krylov solvers aim to find the solution $x_i \in \mathscr{K}_i$ which minimizes the norm of the residual in the current Krylov subspace $\mathscr{K}_i$. The approaches differ by the definitions of the norm used. For instance, a solver like Minimal Residual (MR) minimizes the Euclidean norm of the residual, i.e. $\|r\|_2 = \sqrt{r^\dagger r}$, while a solver like Conjugate Gradient (CG) minimizes the D-norm of the error, i.e. $\|e\|_D = \sqrt{e^\dagger De}$, where D must be an hermitian matrix, resulting in the minimization of the D^{-1}-norm of the residual.

- The norm of the residual is often used as a stopping criterium. The number of iterations n is such that
$$\|r_n\| = \|b - Dx_n\| < \rho, \tag{3.7}$$

where ρ is the stopping criterium. Using the Euclidean norm we have

$$\|r\|_2 = \|De\|_2 \leq \|D\|_2 \|e\|_2 = |\lambda_{\max}| \|e\|_2 \quad \text{and}$$

$$\|e\|_2 = \|D^{-1}r\|_2 \leq \|D^{-1}\|_2 \|r\|_2 = \frac{1}{|\lambda_{\min}|} \|r\|_2, \tag{3.8}$$

with $|\lambda_{\min}| = \sqrt{\lambda_{\min}^\star \lambda_{\min}}$ and $|\lambda_{\max}| = \sqrt{\lambda_{\max}^\star \lambda_{\max}}$ where $\lambda_{\min}$ and $\lambda_{\max}$ are respectively the smallest and the largest eigenvalues of D. Therefore the error in the solution is bounded by

$$\frac{1}{|\lambda_{\max}|}\rho \sim \frac{1}{|\lambda_{\max}|}\|r\|_2 \leq \|e\|_2 \leq \frac{1}{|\lambda_{\min}|}\|r\|_2 < \frac{1}{|\lambda_{\min}|}\rho. \tag{3.9}$$

It follows that the smallest eigenvalue plays an important role since the error is bounded from above by $\rho/|\lambda_{\min}|$, which can take very large values for small $\lambda_{\min}$. It is then important to keep $\rho/|\lambda_{\min}|$ sufficiently small.

- From these considerations, it follows that the convergence of Krylov solver to the exact solution is a function of the condition number of the operator. The condition number K of a matrix D is defined as

$$K(D) = \|D\| \, \|D^{-1}\| = \frac{|\lambda_{\max}|}{|\lambda_{\min}|}, \tag{3.10}$$

where the right hand side has been obtained using the Euclidean norm. Following the same approach as above we find that

$$\frac{\|e\|}{\|x\|} \leq K(D)\frac{\|r\|}{\|b\|}. \tag{3.11}$$

- We distinguish the Krylov methods between short and long recurrence algorithms. In a *short recurrence* algorithm the update of the iterated solution requires the storage of a constant number of vectors; in a *long recurrence* algorithm the update of the solution requires a number of vectors which increases with the number of iterations. For instance the MR solver is a long recurrence algorithm where all the vectors belonging to the Krylov subspace are stored and the solution is searched in the full Krylov subspace in order to find the absolute minimum to the norm of the residual.

- The storage required by the long recurrence solver can be reduced by restarting the Krylov solver, i.e. using as initial guess the last iterated solution. This though does not guarantee the convergence of the solver.

Several Krylov solvers have been analyzed for the twisted mass operator in Ref. [79]. It has been found that the best performing approach is the conjugate gradient (CG). In the following we introduce this approach, which has been extensively used for twisted mass fermions until the employment of multigrid methods.

3.1.1 Conjugate gradient

The conjugate gradient (CG) algorithm is one of the best known iterative techniques for solving sparse symmetric positive definite linear systems. It can be employed in LQCD for inverting the squared Dirac operator, $H^2 = DD^\dagger \rightarrow H^2 x = b$, and, when needed, the solution of the Dirac operator is retrieved as

$$D^{-1}b = D^\dagger H^{-2} b = D^\dagger x. \tag{3.12}$$

Inverting the square of the Dirac operator has the disadvantage that it increases the condition number of the matrix, having $\kappa(H^2) = \kappa(D)^2$. For this reason, in case of Wilson fermions other Krylov solvers, which can invert directly D are faster and therefore preferred. For twisted mass fermions it has been showed that CG is the best performing solver [79]. This has been observed for twisted mass fermions at maximal twist where the mass of the Wilson operator is tuned to its critical value. Thus,

the smallest eigenvalue of $\Gamma_5 D_{\text{TM}}(\mu)$ is $\lambda_{\min} = i\mu$, as seen from Eq. (2.105), which is purely imaginary and makes the operator a complex shifted maximally indefinite matrix. The squared operator $D^\dagger D + \mu^2$ is instead well conditioned and CG is the best algorithm for the inversion.

The starting point of the CG approach is to consider a set of residuals, which are orthogonal to each other:

$$r_j^\dagger r_i = \delta_{ij} \|r_i\|. \tag{3.13}$$

Supposing that any residual r_i belongs to the Krylov subspace $\mathscr{K}_{i+1}(A, r_0)$, as should be the case of Krylov solvers, then

$$r_i \in \mathscr{K}_{i+1}(A, r_0) \quad \Longrightarrow \quad r_i \in \mathscr{K}_i^\perp(A, r_0) \quad \text{since} \quad r_i \perp r_{i-1} \in \mathscr{K}_i(A, r_0), \tag{3.14}$$

where $\mathscr{K}_i^\perp(A, r_0)$ is a orthogonal subspace to $\mathscr{K}_i(A, r_0)$. By definition of Krylov solver, also the iterated solution x_i belongs to $\mathscr{K}_i(A, r_0)$ and thus

$$x_i^\dagger r_j = 0 \quad \forall i \leq j. \tag{3.15}$$

The difference between two consecutive iterations is

$$(x_i - x_{i-1})^\dagger (r_{j+1} - r_j) = \Delta x_i^\dagger A \Delta x_{j+1} = 0 \quad \forall i \leq j, \tag{3.16}$$

which show that the updates of the solution $\Delta x_i = x_i - x_{i-1}$ are A-orthogonal to each other. Thus, the exact solution should be written as

$$A^{-1} b = x_N = x_0 + \sum_{i=0}^{N} a_i p_i, \tag{3.17}$$

where N is the size of the matrix, x_0 is the initial guess and p_i are a set of A-orthogonal vectors, i.e.

$$p_j^\dagger A p_i = \delta_{ij} \|p_i\|_A \quad \text{with} \quad \|p_i\|_A = p_i^\dagger A p_i. \tag{3.18}$$

This relation is satisfied for any i and j if A is symmetric (or hermitian), otherwise needs to be satisfied only for $j \leq i$, as follows from Eq. (3.16). The CG algorithm considers the first case, but the algorithm can be generalized to any matrix A as in the formulation of the generalized conjugate residual (GCR) [85]. Multiplying Eq. (3.17) on the left by $p_j^\dagger A$ we obtain

$$p_j^\dagger A A^{-1} b = p_j^\dagger b = p_j^\dagger A x_0 + \sum a_i p_j^\dagger A p_i = p_j^\dagger A x_0 + a_j \|p_j\|_A \quad \Longrightarrow \quad a_j = \frac{p_j^\dagger r_0}{\|p_j\|_A}. \tag{3.19}$$

This gives the coefficients a_j, which can be computed without the knowledge of the exact solution x_N. The iterated solution and residual are

$$x_k = x_0 + \sum_{i=1}^{k-1} a_i p_i = x_{k-1} + a_{k-1} p_{k-1} \quad \text{and} \quad r_k = r_0 - \sum_{i=1}^{k-1} a_i A p_i = r_{k-1} - a_{k-1} A p_{k-1}. \tag{3.20}$$

Since the updated vector for the next solution, p_k, needs to belong to the subspace $\mathscr{K}_{k+1}(A, r_0)$ it can be written as

$$p_k = r_k + \beta_k p_{k-1}, \tag{3.21}$$

where $r_k \in \mathcal{K}_{k+1}(A, r_0)$ and $p_{k-1} \in \mathcal{K}_k(A, r_0)$. Using this we obtain

$$p_{k-1}^{\dagger} A r_k = p_{k-1}^{\dagger} A(p_k - \beta_k p_{k-1}) = -\beta_k \|p_{k-1}\|_A \quad \rightarrow \quad \beta_k = -\frac{p_{k-1}^{\dagger} A r_k}{\|p_{k-1}\|_A}. \tag{3.22}$$

Different definitions of α_k and β_k can be derived from the properties of x_k, r_k and p_k. The one that is most commonly used is

$$\alpha_k = \frac{\|r_k\|}{\|p_k\|_A} \quad \text{and} \quad \beta_k = \frac{\|r_{k+1}\|}{\|r_k\|} \tag{3.23}$$

since only two norms need to be computed every iteration, namely $\|r_{k+1}\|$ and $\|p_k\|_A$. An implementation of the CG algorithm is listed in Alg. 1. From the computational side, it requires per iteration one application of A (line 6) and two global reductions, $(p^{\dagger} p')$ and $\|r\|$ (lines 7, 10). These components in a parallel implementation require communication, while the remaining of the algorithm is computed in-node. The convergence rate of CG is [85],

---

Algorithm 1: $x = CG(A, b, x_0)$

Input: b, x_0
Output: x
Parameters: $n_{\max}$, $\rho_{\min}$

1   $x = x_0$
2   $r = b - Ax$
3   $p = r$
4   $\rho = \|r\|$
5   **for** $i = 1$ **to** $n_{\max}$ **or until** $\rho > \rho_{\min}$ **do**
6      $p' = Ap$
7      $\alpha = \rho / (p^{\dagger} p')$
8      $x \leftarrow x + \alpha p$
9      $r \leftarrow x - \alpha p'$
10     $\beta = \|r\| / \rho$
11     $p \leftarrow r + \beta p$
12     $\rho \leftarrow \beta \rho$
13 **return** x

$$\rho = \frac{\|r\|}{\|b\|} \leq 2 \left(\frac{\sqrt{K(A)} - 1}{\sqrt{K(A)} + 1} \right)^n, \tag{3.24}$$

where n is the number of iterations. It follows that the number of iterations required for obtaining the desired relative residual ρ is

$$n \leq \frac{\ln(2/\rho)}{\ln\left(\sqrt{K(A)} + 1\right) - \ln\left(\sqrt{K(A)} - 1\right)}. \tag{3.25}$$

Considering a constant time per iteration, t_{iter}, we can parameterize the time to solution of the CG solver at fixed ρ as function of the condition number κ as

$$t_{\text{CG}}(K) \leq t_{\text{init}} + t_{\text{iter}} \frac{\ln(2/\rho)}{\ln\left(\sqrt{K} + 1\right) - \ln\left(\sqrt{K} - 1\right)}, \tag{3.26}$$

where t_{init} is the initialization cost of CG.

3.1.2 Critical slowing down

When one approaches the physical value of the light quark mass, the efficiency of the CG solver suffers due to the increase of the condition number. Any discretization of the Dirac operator suffers from this so-called critical slowing down. In case of CG with twisted mass fermions, the time to solution increases with $1/\mu^2$. We show this phenomenon in Fig. 3.2, where we depict numerical results computed on an ensemble with light quark mass tuned to its physical value (physical ensemble) with a lattice volume of $V = 48^3 \times 96$. The characteristic quantities of this ensemble are listed in Tab. 3.1.

Name	Ref.	N_f	V	β	c_{SW}	κ	μ	a [fm]	m_π [MeV]
$cA2.48$	[76]	2	$48^3 \times 96$	2.1	1.57551	0.13729	0.0009	0.0913(2)	131.9(3)

Table 3.1: We give the simulation parameters (N_f, V, β, c_{SW}, κ and μ) and some characterizing quantities (a and m_π) for the $cA2.48$ ensemble studied in this section.
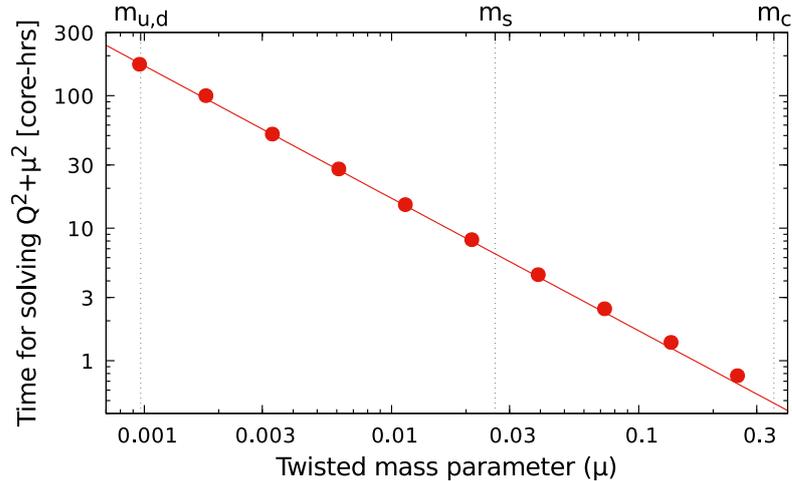
Figure 3.2: Time to solution for computing the twisted mass fermion propagator at different quark mass using the CG solver with the odd-even reduced twisted mass operator. The value of the light, strange and charm quark masses are shown by the vertical lines with the labels $m_{u,d}$, m_s and m_c, respectively.

We fit the time to solution using Eq. (3.26), with initialization time $t_{init} \simeq 0$ core-seconds, iteration time $t_{iter} \simeq 35$ core-seconds and stopping criterion $\rho = 10^{-9}$. Having $\lambda_{\min}(Q^2 + \mu^2) \simeq \mu^2$, we find $\lambda_{\max}(Q^2 + \mu^2) \simeq 0.26$. The fitted data are depicted in Fig. 3.2. From this data set one understands easily how prohibitive simulations at the physical point are. This has limited simulations to larger than physical values of the pion mass for a number of years. In order to extract physical results one then resorted to chiral perturbation theory. Chiral extrapolations typically work well for mesons but for baryon such an extrapolation is problematic and brings in and uncontrolled error. Therefore, for baryon observables one benefits from having simulations performed directly at the physical point.

3.1.3 CG with exact deflation

Before the employment of multigrid methods, one used CG with exact deflation [86, 87]. In this approach, the eigenvectors with small eigenvalue are computed exactly and the deflated operator is

inverted. This improves the time to solution reducing the condition number of the Dirac matrix. Let us consider a set of the low-mode eigenvectors of $D^\dagger D + \mu^2$ collected in the matrix V, then

$$V^\dagger(D^\dagger D + \mu^2)V = \Lambda \quad \text{and} \quad H_V = (1 - VV^\dagger)^\dagger(D^\dagger D + \mu^2)(1 - VV^\dagger), \tag{3.27}$$

where Λ is a diagonal matrix of the respective eigenvalues and H_V is the deflated operator. The solution of the linear system $(D^\dagger D + \mu^2)x = b$ is computed as

$$x = (D^\dagger D + \mu^2)^{-1}b = H_V^{-1}b + V\Lambda^{-1}V^\dagger b, \tag{3.28}$$

where $H_V^{-1}b$ is solved with CG. The approach just described is referred to as CG with exact deflation (CG-eDe) and it can be improved by the algorithm outlined in Ref. [88] known as eigCG. In the latter, the eigenvectors of the operator are computed progressively while computing solutions with CG. At every inversion a small set of eigenvectors would be computed and then used in the next inversion for accelerating the convergence. An alternative approach used in this work, is to compute all low-lying eigenvectors at once using the implicitly restarted Arnoldi method implemented in ARPACK [89].

For a well deflated CG solver, most of the smaller eigenvalues need to be computed in order to effectively deflate the operator. For the cA2.48 this requires the calculation of 1600 eigenvectors. Note that the number of eigenvectors needed grows with the volume since the spectral density increases with the volume. In Fig. 3.3 we present a comparison of the inversion times of the CG solver and the CG solver with exact deflation (CG-eDe). For CG and CG-eDe we use the publicly available software package *tmLQCD* [90] that is commonly used for simulations with the twisted mass operator.

Solver	Setup time [core-hrs]	Inversion time [core-hrs]	Iteration count
CG	–	338.6	34 790
CG-eDe	6 941.1	9.8	695
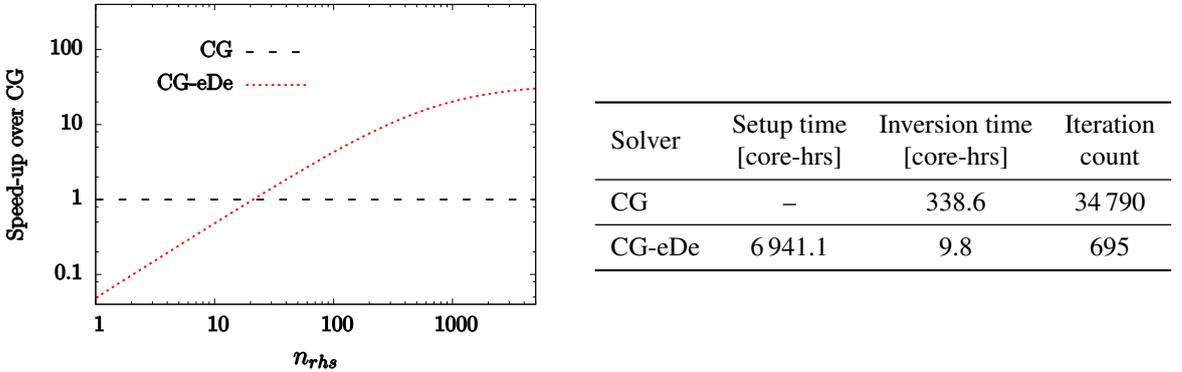
Figure 3.3: Comparison of CG (tmLQCD), CG-eDe (tmLQCD+ARPACK) with 1600 eigenvectors. The results have been computed for the cA2.48 ensemble, Tab. 3.1, with $m_\pi = 0.131$ GeV. We depict on the left the speed-up over CG when inverting multiple right hand side, n_{rhs}, using the same setup. The reference timings are reported in the table on the right. The

CG-eDe achieves a speed-up compared to CG which is of a factor greater than twenty. On the other hand, the setup time is expensive and more than one thousand rhs need to be inverted in order to take advantage of the full speed-up as depicted in Fig. 3.3. Since the eigenvectors are configuration dependent and they need to be computed exactly in order to make the approach work, this approach works if we have a high number of inversions. CG-eDe improves the inversion time for measurements but does not help in the simulations where only a few inversions are computed per configuration before updating the links. The introduction of multigrid solvers has thus played a very important role for accessing the physical point without a prohibitive computational cost.

3.2 Multigrid methods

The main idea of multigrid methods is to accelerate the convergence of standard iterative solvers by coarsening the degrees of freedom of the linear system reducing its complexity. The coarsen problem, constructed on a so-called *coarse-grid*, restricts long-wavelength information of the target, namely the *fine-grid*, linear system. The coarse linear operator is significantly cheaper to solve and is used to compute a *coarse-grid correction* of the iterated fine-grid solution. In multigrid methods this idea is extended recursively, i.e. by additionally coarsening the coarse problem, defining a hierarchy of coarse-grids *levels*, until a small enough linear system is reached.
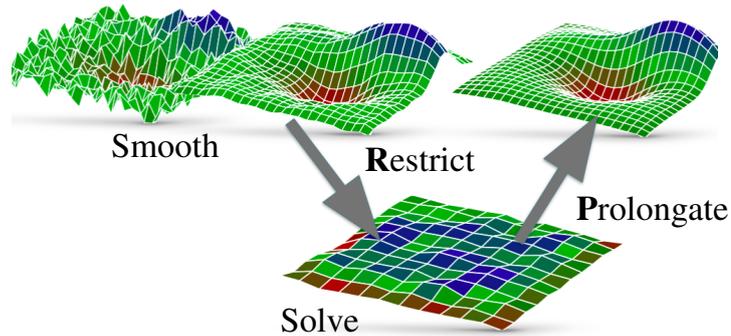
Figure 3.4: A two-grid cycle showing the main multigrid steps. *Image credit: Luke Olson.*

Multigrid methods can be used either as preconditioners or stand-alone solvers and many variations are available in the literature [91]. Their common features are:

- **A smoother,** a method applied on either the fine- or coarse-grid for reducing the short-wavelength fluctuations. The smoother is applied horizontally in the multigrid hierarchy without changing level. Typically, the smoother is a cheap and simple iterative method, such as a relaxation schemes like Jacobi, Gauss-Seidel or their block variants as well as few iterations of a Krylov subspace method;

- **A restriction operator,** the mapping that allows to move from a finer to a coarser grid in the multigrid hierarchy;

- **A prolongation operator,** the mapping that allows to move from a coarser to a finer grid in the multigrid hierarchy.

Restriction and prolongation operators are the most critical components for a successful multigrid approach and they are critically discussed in this chapter.

3.2.1 Two-level multigrid operators

In a two-level multigrid approach, restriction and prolongation operators, R and P respectively, are used to define the coarse-grid operator, D_c, from the fine-grid operator, D, as follows

$$D_c = RDP. \tag{3.29}$$

D is the Wilson Dirac operator but the approach can be applied to a generic linear operator. R and P are then matrices of size (m,n) and (n,m), respectively, where (m,m) is the size of D and (n,n) is the

size of D_c with $n < m$. Commonly, R and P also satisfy

$$RIP = I_c, \tag{3.30}$$

such that the fine-grid identity, I, is mapped into the coarse-grid identity, I_c.

In a standard multigrid (MG) approach, R and P are constant sparse matrix constructed with a geometric definition. For instance, in a *block-aggregation* of neighboring sites, demonstrated in Fig.3.5, R averages between sites of a block-aggregate, $\mathscr{A}_i$, and then P is constructed to satisfy the property in Eq.(3.30). A standard MG approach is usually applied for solving partial differential equations (PDEs) discretized on a grid, where the geometrical meaning of the linear operator is clear. The Dirac operator in lattice QCD is part of this category and such approach has been investigated at the beginning of the '90s [92–95] without satisfactory results. The reason is due to the random nature of the gauge fields entering in the lattice Dirac operator that obfuscates the geometrical interpretation of the underlying PDE. The advent of algebraic multigrid methods (AMG) and the pioneering studies in Refs. [96–98], opened the way for a successful application of multigrid methods in LQCD. These works are discussed later in this chapter.
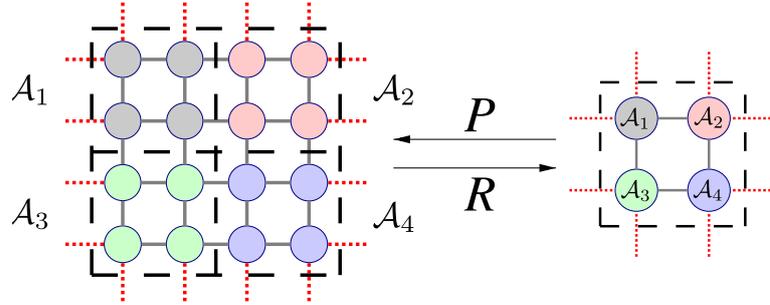
Figure 3.5: Block aggregation of neighboring sites and symbolic representation of the movement from the fine-grid to the coarse-grid and vice-versa.

From the coarse-grid operator in Eq. (3.29) the following error propagation is defined

$$\varepsilon \leftarrow \left(I - PD_c^{-1}RD\right)\varepsilon. \tag{3.31}$$

From this we obtain the *coarse-grid correction* of the iterated solution z

$$z \leftarrow z + PD_c^{-1}R\left(b - Dz\right), \tag{3.32}$$

where b is the right hand side. Such correction involves the solution of the coarse-grid linear system

$$D_c e_c = r_c := R\left(b - Dz\right). \tag{3.33}$$

It is then straightforward to extend recursively the multigrid approach by employing an additional coarse-grid correction for solving the coarse-grid linear system given in Eq. (3.33). A generic two-level *multigrid step* is then described via the error propagation

$$\varepsilon \leftarrow (I - MD)^k \left(I - PD_c^{-1}RD\right)(I - MD)^j \varepsilon, \tag{3.34}$$

where M denotes the smoother and j and k are, respectively, the number of pre- and post-smoothing

iterations, applied before and after the coarse-grid correction.

3.2.2 Multi-level approach

For a rigorous extension to a multi-level approach, we use N_ℓ to indicate the number of levels, with the first level being the fine-grid and the N_ℓth level the coarsest-grid. For each level ℓ, except the coarsest, $\ell \in [1, N_\ell)$, we construct R_ℓ, P_ℓ, M_ℓ and D_ℓ, where

$$D_{\ell+1} = R_\ell D_\ell P_\ell \quad \text{and} \quad D_1 = D. \tag{3.35}$$

On the coarsest-grid we only need D_{N_ℓ} since no additional coarsening is applied. From a straight-forward recursive application of the error propagation of Eq. (3.34) we obtain the algorithm listed in Alg. 2, commonly referred to as V-cycle. Other well-known algorithms are the W- and K-cycles. These approaches are schematically depicted in Fig. 3.6. In a W-Cycle, smoothing and coarse-grid correction are arbitrarily repeated multiple times for each level (see Alg. 3). In a K-cycle, instead, the repetition is not arbitrarily and a Krylov solver preconditioned by a next-level K-Cycle is used for computing the coarse-grid correction, (see Alg. 4). Note that the "V" and "W" in the name of the respective cycles refer to the shape of the diagram as depicted in Fig. 3.6. On the other hand, the "K" of K-cycle refers to Krylov solver, despite we also depict the K-cycle in the figure with a diagram of shape "K".
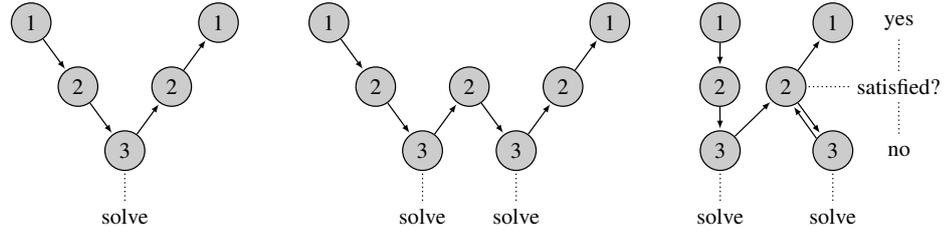
Figure 3.6: Schematic representation of V-, W- and K-Cycles for a three-level multigrid approach.

Algorithm 2: $z_\ell = \textbf{V-Cycle}(\ell, b_\ell)$

Input: b_ℓ, $\ell < N_\ell$
Output: z_ℓ
Parameters: N_ℓ, ρ, j, k

1   $z_\ell = 0$
2   **for** $i = 1$ **to** j **do**
3     $z_\ell \leftarrow z_\ell + M_\ell(b_\ell - D_\ell z_\ell)$ } Pre-smoothing
4   $r_{\ell+1} \leftarrow R_\ell(b_\ell - D_\ell z_\ell)$
5   **if** $\ell + 1 = N_\ell$ **then**
6     $z_{\ell+1} \leftarrow \textbf{solve}(D_{\ell+1} z_{\ell+1} = r_{\ell+1},\ \text{residual} = \rho)$
7   **else**
8     $z_{\ell+1} \leftarrow \textbf{V-Cycle}(\ell+1, r_{\ell+1})$ Coarse-grid correction
9   $z_\ell \leftarrow z_\ell + P_\ell z_{\ell+1}$
10   **for** $i = 1$ **to** k **do**
11     $z_\ell \leftarrow z_\ell + M_\ell(b_\ell - D_\ell z_\ell)$ } Post-smoothing
12   **return** z_ℓ

3.2.3 Algebraic multigrid (AMG)

Algebraic multigrid (AMG) is an important extension of multigrid methods. The main difference with standard, i.e. geometrical, MG approaches regards the construction of the restriction and prolongation operators which in a AMG method do not necessary have a geometrical meaning. They are indeed non-constant matrices constructed upon the fine-grid operator itself, i.e. $R \equiv R(D)$ and $P \equiv P(D)$. The purpose of R and P is still to define a coarse operator D_c as in Eq. (3.29) with long wavelength information of the fine grid operator D. Multilevel approaches can be used as done for standard MG.

Exact deflation could be an example of AMG algorithm, but we will explain why in fact it is not. We have introduced CG with exact deflation in Sec. 3.1.3 and here we explain the same concept using the multigrid operators. Let v_i be right-eigenvectors of D and $\hat{v}_i$ be the respective left-eigenvectors, then they satisfy

$$\hat{v}_i^{\mathrm{T}} D v_i = \lambda_i \, \hat{v}_i \cdot v_i = \lambda_i. \tag{3.36}$$

If we construct R and P from a set of N_v left- and right- eigenvectors, respectively,

$$P_{eDe} = (v_1 \mid \ldots \mid v_{N_v}) \quad \text{and} \quad R_{eDe} = (\hat{v}_1 \mid \ldots \mid \hat{v}_{N_v})^{\mathrm{T}}, \tag{3.37}$$

where the notation $(\|)$ means that the vectors are used as columns of the matrix, we obtain then a coarse-grid operator, as defined in (3.29), which is the diagonal matrix of the respective eigenvalues,

$$D_{eDe} = R_{eDe} D P_{eDe} = \mathrm{diag}(\lambda_1, \ldots, \lambda_{N_v}), \tag{3.38}$$

equivalent to Λ in Eq. (3.27). In case of the Wilson Dirac operator D, left- and right-eigenvectors are

---

Algorithm 3: $z_\ell = \textbf{W-Cycle}(\ell, b_\ell)$

Input: $b_\ell, \ell < N_\ell$
Output: z_ℓ
Parameters: $N_\ell, \rho, j, k, n_\ell$
Comment: Same as Algorithm 2 with a for-loop wrapping the lines from 2 to 11 and replacing V-Cycle with W-Cycle.

   1   $z_\ell = 0$
   2   **for** $i = 1$ **to** n_ℓ **do**
 3-8 | Lines 2 - 7 of Algorithm 2
   9 | $z_{\ell+1} \leftarrow \textbf{W-Cycle}(\ell+1, r_{\ell+1})$
10-12 | Lines 9 - 11 of Algorithm 2
 13   **return** z_ℓ

Algorithm 4: $z_\ell = \textbf{K-Cycle}(\ell, b_\ell)$

Input: $b_\ell, \ell < N_\ell$
Output: z_ℓ
Parameters: N_ℓ, ρ_ℓ, j, k
Comment: Same as Algorithm 2 replacing line 8 as follow.

 1-7   Lines 1 - 7 of Algorithm 2 using $\rho = \rho_\ell$
   8 | $z_{\ell+1} \leftarrow \textbf{solve}(D_{\ell+1} z_{\ell+1} = r_{\ell+1}, \text{ residual} = \rho_\ell, \text{ preconditioner} = \textbf{K-Cycle}(\ell+1, \_\_))$
 9-12   Lines 9 - 12 of Algorithm 2

---

connected by the property

$$\hat{v}_i^{\mathrm{T}} = (\Gamma_5 v_i)^{\mathrm{T}} \quad \Longrightarrow \quad R_{eDe} = (\Gamma_5 P_{eDe})^{\mathrm{T}}, \tag{3.39}$$

since D is Γ_5-hermitian, cf. Eq. (2.68). In case of the squared operator $D^\dagger D$, hermitian by construction, left- and right-eigenvectors are connected by

$$\hat{v}_i^{\mathrm{T}} = v_i^\dagger \quad \Longrightarrow \quad R_{eDe} = P_{eDe}^\dagger. \tag{3.40}$$

The coarse-grid correction in Eq. (3.32) is then equivalent to the exact deflation of the linear operator D in Eq. (3.28).

Even though we can express exact deflation in terms of multigrid operators, i.e. could go under the category of AMG approaches since P and R are constructed from D, it is not a multigrid approach for the following reasons:

- it brings no benefit to additionally coarsen the diagonal matrix D_{eDe} and thus we cannot define a multilevel approach;

- one application of the coarse-grid correction is enough to deflate the low-mode and thus it brings nothing to apply it iteratively, i.e. the solution would not change anymore;

- the computation of R_{eDe} and P_{eDe} is extremely expensive and grows with V^2, which is not in line with the objectives of multigrid approaches.

3.3 Inexact deflation

The first work paving the path for AMG in lattice QCD is by M. Lüscher [96]. The approach presented is named *inexact deflation*, as opposed to *exact* deflation. Inexact deflation is very close to an AMG approach but small differences are significant to change the method and moreover there is no reference to multigrid methods in the presentation of the approach in Ref. [96]. The remarkable innovation brought by inexact deflation regards mostly the reduction of the setup costs as compared to exact deflation. The main idea is indeed to compute the eigenvectors "inexactly", i.e. at low precision, which reduces the computational costs.

3.3.1 Local Coherence

In Ref. [96] it has been observed that low-modes of the Dirac operator are approximately similar over block-aggregates of the lattice. Loosely speaking the long-wavelength fluctuations of the Dirac operator have waves similarly repeated all over the lattice. More formally we can state that a block-segment of a low-mode eigenvector is highly correlated with a large amount of block-segments of other low-mode eigenvectors. This property is known as *local coherence* and for a qualitative analysis we refer to Ref. [96].

Thanks to the property of local coherence we can motivate the construction of a prolongation operator as follow. Let $v_1, ..., v_{N_v}$ be a set of N_v low-mode right-eigenvectors of D, then the prolongation

operator P constructed by decomposing the eigenvetors into N_b block-aggregates is given by

$$(v_1 \mid \ldots \mid v_{N_v}) = \left(\right) \longrightarrow P = \left(\right) \begin{matrix} \mathscr{A}_1 \\ \\ \mathscr{A}_2 \\ \\ \vdots \\ \\ \mathscr{A}_{N_b} \end{matrix}, \tag{3.41}$$

will approximate $\mathscr{O}(N_b \times N_v)$ low-mode eigenvectors. The block-aggregation acts only on position space aggregating to one site a full set of components belonging to

$$\mathscr{A}_j = \mathscr{V}_j \times \mathscr{S} \times \mathscr{C}, \tag{3.42}$$

as can be seen from Eq.(2.67). Local coherence has several important consequences:

- As mentioned, $\mathscr{O}(N_b \times N_v)$ low-modes of D are approximately projected in the coarse Dirac operator $D_c = RDP$.

- The block size, V_b, of the aggregates depends on the wavelengths of the low-modes, which change accordingly to the lattice spacing a and the quark mass m_q of the operator, and not with the lattice volume V. It is found then that a block-aggregate of volume 4^4 is a good choice at the physical point at any volume.

- Thus, since the block size is fixed at given a and m_q, the number of modes restricted in the coarse grid is proportional to the volume having $N_b = V/V_b$.

- The spectral density of the low-modes in the Dirac operator grows as well with the volume. It follows, then, that the dependence on the volume is taken into account by N_b. Therefore, the number of eigenvectors N_v can be constant. It is found that an order of twenty vectors are enough for sufficiently deflating the low-modes at the physical point at any volume.

- The constructed coarse Dirac operator is still a sparse matrix as depicted in Fig. 3.7 involving only first-neighbor interactions. If we consider a block volume $V_b = 4^4$ and number of vectors $N_v = 24$, then the coarse Dirac operator is 128 times smaller than the Dirac operator.

- The setup cost for computing the eigenvectors is then proportional to the volume V and not to V^2 as in exact deflation.

- Moreover, since P approximates the low-modes of D, also the eigenvectors entering in P can be computed approximately.

3.3.2 The algorithm – A missed AMG approach

In Ref. [96], a different approach to the multigrid error propagation in Eq. (3.31) is used. Two projections are introduced,

$$\bar{\pi} = I - DPD_c^{-1}R \quad \text{and} \quad \pi = I - PD_c^{-1}RD. \tag{3.43}$$

Figure 3.7: Schematic representation of fine- and coarse-grid Dirac operator. The same caption of Fig. 3.1 applies here. Differently from Fig. 3.1, the Dirac operator D as been reordered accordingly to the block decomposition of the lattice. The red and blue arrows show, respectively, the different approach between standard MG and AMG methods. In the first, the coarse operator D_c is defined from the coarse-grid, in the latter instead the coarse operator is defined from the fine grid operator D as in Eq. (3.29).

where π is clearly the coarse-grid error propagation of Eq. (3.31). The following properties[1] result:

(i) $\quad \bar{\pi}\bar{\pi} = \bar{\pi}$ and $\quad \pi\pi = \pi$ $\hfill$ (3.44)

(ii) $\quad \bar{\pi}D = D\pi$ $\hfill$ (3.45)

(iii) $\quad R\bar{\pi} = \pi P = 0$ $\hfill$ (3.46)

(iv) $\quad \bar{\pi}DP = RD\pi = 0$ $\hfill$ (3.47)

Moreover, if R and P satisfy $RP = I_c$, as in multigrid methods, then

(v) $\quad \bar{\pi}(I - PR) = (I - PR)\pi = (I - PR)$ $\hfill$ (3.48)

---

[1]Proofs:

(i) $\quad \bar{\pi}\bar{\pi} = I - 2DPD_c^{-1}R + DPD_c^{-1}\underbrace{RDP}_{D_c}D_c^{-1}R = \bar{\pi}$ and $\quad \pi\pi = I - 2PD_c^{-1}RD + PD_c^{-1}\underbrace{RDP}_{D_c}D_c^{-1}RD = \pi$

(ii) $\quad \bar{\pi}D = D - DPD_c^{-1}RD = D\pi$

(iii) $\quad R\bar{\pi} = R - \underbrace{RDP}_{D_c}D_c^{-1}R = 0$ and $\quad \pi P = P - PD_c^{-1}\underbrace{RDP}_{D_c} = 0$

(iv) combining (ii) and (iii): $\quad \bar{\pi}DP = D\pi P = 0$ and $\quad RD\pi = R\bar{\pi}D = 0$

(v) $\quad \bar{\pi}(I - PR) = I - DPD_c^{-1}R - PR + DPD_c^{-1}\underbrace{RP}_{I_c}R = (I - PR)$ and

$\quad (I - PR)\pi = I - PD_c^{-1}RD - PR + P\underbrace{RP}_{I_c}D_c^{-1}RD = (I - PR).$

These projections are used to decompose the linear system $Dx = b$ as follows

$$(I - \bar{\pi})Dx = (I - \bar{\pi})b \quad \Longrightarrow \quad D\underbrace{(I - \pi)x}_{\chi} = (I - \bar{\pi})b \quad \Longrightarrow \quad \chi = PD_c^{-1}Rb, \qquad (3.49)$$

$$\underbrace{\bar{\pi}\bar{\pi}}_{\bar{\pi}}Dx = \bar{\pi}b \quad \Longrightarrow \quad D\pi\underbrace{\pi x}_{\chi'} = \bar{\pi}b \quad \Longrightarrow \quad \chi' = (D\pi)^{-1}\bar{\pi}b. \qquad (3.50)$$

and $x = \chi + \chi'$. χ, involves the inversion of the coarse Dirac operator; χ' involves the inversion of $D\pi$ with rhs $\bar{\pi}b = (b - D\chi)$ which is the residual of the partial solution χ. The inexact deflation algorithm is listed in Alg. 5.

Algorithm 5: $z = \mathbf{InDe}(b)$

   Input: b
   Output: z
   Parameters: ρ, z_0

 1 $r \leftarrow b - Dz_0$
 2 **solve** χ_c in $D_c\chi_c = Rr$ with residual ρ $\Big\}$ Coarse-grid correction
 3 $r \leftarrow r - DP\chi_c$
 4 **solve** χ' in $D\pi\chi' = r,$ with residual ρ
 5 $z \leftarrow z_0 + P\chi_c + \chi'$
 6 **return** z

By comparing Alg. 5 and the coarse-grid corrections for a two-level multigrid approach in Eqs. (3.32) and (3.33) is clear that the first step of the inexact deflation approach is equivalent to a coarse grid correction. The main difference between the two approaches is then that

- for exact deflation the coarse grid correction is applied only once and the remaining part of the solution is computed directly with a deflated Dirac operator $D\pi$;

- for a multigrid approach, instead, the coarse-grid correction with pre- and post-smoothing is iterated until convergence.

This has been demonstrated in Ref. [80] where inexact deflation has been compared to an AMG approach. After the publication of the latter work the inexact deflation algorithm implemented by M. Lüscher in openQCD[2] has been improved to be effectively an AMG approach. Except for the algorithm itself, the work of M. Lüscher in Ref. [96] opened the path to a successful application of AMG approaches in lattice QCD. Moreover many other important characteristics of AMG methods have been rediscovered in its work. For instance, it has been proposed to use for the inversion of $D\pi$ a preconditioner, the Schwartz alternating procedure (SAP), which reduces the short-wavelength fluctuations and it is a suitable smoother in case of multigrid methods. The construction of P as in Eq. (3.41) is also typical of AMG methods with block aggregation. And the restriction operator is chosen as $R = P^\dagger$ that in AMG methods is known as Galerkin approach.

3.4   MG-GCR – The first AMG approach in LQCD

Going from inexact deflation to a real AMG approach, is straightforward. The USQCD collaboration implemented the first AMG approach for Wilson fermions in LQCD. The method has been called

---

[2]http://luscher.web.cern.ch/luscher/openQCD/

MG-GCR and for details we refer to Refs. [97–102].

The algorithm is a GCR solver preconditioned by a multilevel AMG approach. As explained in Sec. 3.4.1, a Petrov-Galerkin approach is used for the definition of the restriction operator, while the prolongation operator P, is constructed from a set of null-vectors built in the setup procedure, as explained in Sec. 3.4.2. The smoother consists of a few iterations of the MR solver.

3.4.1 Petrov-Galerkin approach for Wilson fermions

For an efficient AMG approach, the restriction operator R should be constructed from the left-eigenvectors of D and P from the right-eigenvectors. Defining R as $R = P^\dagger$ is known as Galerkin approach in the MG literature, while defining R from the right eigenvalues is known as Petrov-Galerkin approach. If left- and right-eigenvectors are not connected by any relation, this requires a separated setup procedure for R which doubles the setup costs. In case of Wilson fermions, we know that left- and right-eigenvectors are connected by the relation in Eq. (3.39). Therefore, if P is built from vectors $v_1, \ldots, v_N$, which approximate right eigenvectors with small eigenvalues of D, then

$$R = P^\dagger \Gamma_5 \tag{3.51}$$

is the most suitable restriction operator for Wilson fermions accordingly to the Petrov-Galerkin approach. The choice done in exact deflation, $R = P^\dagger$, follows instead the Galerkin approach.

As pointed out in Ref. [98], it is possible to combine Petrov-Galerkin and Galerkin approaches, defining a prolongation operator, which takes into account the spin structure of the Dirac operator. It can be done by decomposing the aggregation in the spin space $\mathscr{S}$ with an aggregate

$$\mathscr{A}_{j,+} = \mathscr{V}_j \times \mathscr{S}_{0,1} \times \mathscr{C}, \tag{3.52}$$

which collects the two upper spin components, 0 and 1, and an aggregate

$$\mathscr{A}_{j,-} = \mathscr{V}_j \times \mathscr{S}_{2,3} \times \mathscr{C}, \tag{3.53}$$

which collects the two lower spin components, 2 and 3. In the chiral basis, Γ_5 acts with $+1$ on the components of the aggregate $\mathscr{A}_{j,+}$ while with -1 on the other aggregates. Now, the coarse grid $\Gamma_{5,c}$ can be defined by $\Gamma_{5,c} = I_{\mathscr{V}_c} \otimes \tau_3 \otimes I_{N_v}$, where τ_3 acts on the different spin aggregates. This type of aggregation was proposed in Ref. [100] for the MG-GCR method. A prolongation operator P, which acts on the aggregates $\mathscr{A}_{j,+}$ and $\mathscr{A}_{j,-}$ is said Γ_5-compatible and satisfies

$$\Gamma_5 P = P \Gamma_{5,c}. \tag{3.54}$$

Schematically, we represent the latter equation as

$$
\begin{pmatrix} I_{6 \cdot V_b} & & \\ & -I_{6 \cdot V_b} & \\ & & \ddots \end{pmatrix} \begin{pmatrix} \rlap{$\blacksquare$} & & \\ & \rlap{$\blacksquare$} & \\ & & \ddots \end{pmatrix} = \begin{pmatrix} \rlap{$\blacksquare$} & & \\ & \rlap{$\blacksquare$} & \\ & & \ddots \end{pmatrix} \begin{pmatrix} I_{N_v} & & \\ & -I_{N_v} & \\ & & \ddots \end{pmatrix} . \qquad (3.55)
$$

A Γ_5-compatible prolongation operator preserves the Γ_5-hermiticity of the Wilson operator on the coarse grid, i.e.

$$
\Gamma_{5,c} D_c = \Gamma_{5,c} P^\dagger D P = P^\dagger D^\dagger P \Gamma_{5,c} = D_c^\dagger \Gamma_{5,c}. \qquad (3.56)
$$

The hermitian operator $H_c = \Gamma_{5,c} D_c$ is the same as the operator $H_c^{\mathrm{PG}} = P^\dagger \Gamma_5 D P$, where in the latter a non Γ_5-compatible prolongation operator has been used. Thus, the Γ_5-compatibility has the important consequence of preserving the Γ_5-hermiticity of the Wilson operator and allows to employ in the coarse grid the matrix D_c instead of H_c^{PG}. Indeed, the inversion of the Hermitian matrix H_c^{PG}, which has positive and negative eigenvalues is harder since is close to be maximally indefinite [103], i.e. the number of negative eigenvalues is about the same as the positive ones. Instead the eigenvalues of D_c lie on the right-half of the complex plane as in the case of D. Solving D_c as the same complexity of solving D and this is optimal for defining a stable multilevel approach.

3.4.2 Null-vector setup

The setup procedure employed in Refs. [97, 98, 102] is based on the idea of adaptive setup proposed in Ref. [104]. The idea is to use the smoother itself for finding the error components, which will not be effectively reduced by it. Considering the homogeneous equation

$$
Du = 0, \qquad (3.57)
$$

which has trivial solution, $u = 0$, and solving it with the smoother starting from a random initial guess, one finds a solution vector u, which is the numerical error of the solver with respect to the real solution 0. The numerical error in u is populated by the modes that the smoother does not handle well. In case of an ill-conditioned operator these modes are the low modes of the operator. The solution u is called null-vector and the setup procedure for N_v null-vectors is listed in Alg. 6. It is repeated for every level of the multigrid method.

3.5 Aggregation-based Domain Decomposition Multigrid (DD-αAMG)

In this section we introduce in detail the multigrid approach used in this work DD-αAMG based on Refs. [80, 105, 106]. The main purpose of DD-αAMG is to find the best set of elements to construct an optimal AMG approach for Wilson fermions. Most of the components are inspired by inexact deflation, MG-GCR and other AMG applications outside LQCD.

The resulting method is based on a flexible iterative Krylov solver, FGMRES, which is preconditioned at every iteration step by a multigrid approach. The multigrid preconditioner exploits domain

decomposition strategies: the smoother is given by the Schwarz Alternating Procedure (SAP), see Sec. 3.5.1 and the coarse grid operator $D_c = P^\dagger D P$ is constructed from a Γ_5-compatible prolongation operator P, see Sec. 3.4.1. The method is designed to deal efficiently with both, infrared (IR) and ultra-violet (UV) modes of the operator D, as depicted in Fig. 3.8. The smoother reduces the error components belonging to the UV-modes [80], while the coarse grid correction is designed to deal with the IR-modes. This is achieved by constructing the prolongation operator P from an iterative adaptive setup procedure, see Sec. 3.5.3. Thanks to the property of local coherence explained in Sec. 3.3.1, the subspace can be approximated by aggregating over a small set of $N_v \simeq \mathscr{O}(20)$ test vectors v_i. On the coarse grids, single precision and, on the coarsest grid, even-odd reduction are used to accelerate the time to solution. Note that it could be beneficial to employ even-odd reduction also for higher levels, see Ref. [102] and results in Sec. 4.4 for MG-GCR.
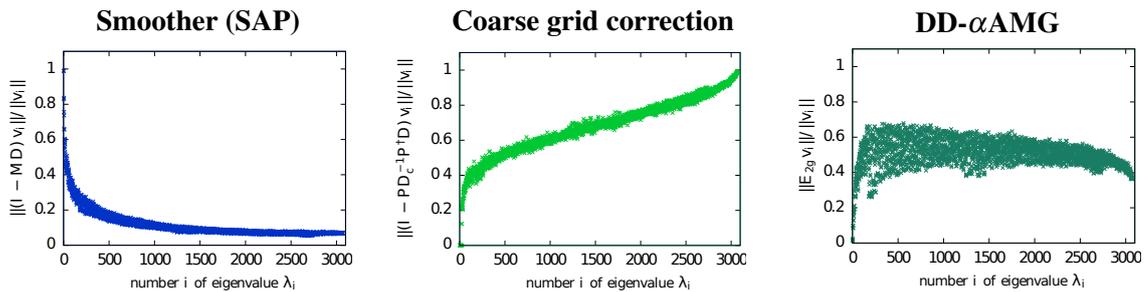
Figure 3.8: Combination of smoother and coarse grid correction in the DD-αAMG approach applied on the eigenvectors of the Dirac operator for a 4^4 trial lattice. *Image credit: Mathias Rottmann*

3.5.1 Smoother

In the DD-αAMG approach, a red-black Schwarz Alternating Procedure (SAP) is used as a smoother. This domain decomposition method was introduced to LQCD in Ref. [107], where it was used as a preconditioner. The lattice is partitioned into alternated "red(r)" and "black(b)" lattice blocks in a checkerboard manner and the subdomains are obtained as the full color-spin space over the respective lattice block, cf. Eq. (3.42). Note that the operator D has only next neighbor interactions and thus blocks of a specific color do not couple to the same color. Re-ordering the Dirac twisted mass operator

---

Algorithm 6: $P_\ell = \textbf{Setup}(\ell = 1)$

Input: $\ell < N_\ell$: level.
Output: P_ℓ
Parameters: N_v, N_b, ρ_{smooth}

1 **for** $j = 1$ **to** N_v **do**
2   v_ℓ^j = random vector
3    $v_\ell^j \leftarrow \textbf{solve}(D_\ell v_\ell^j = 0, \; \rho_{smooth})$ with the smoother
4 **for** $j = 1$ **to** N_b **do**
5    $\{v_\ell \in \mathscr{A}_j\} \leftarrow$ Gram-Schmidt $\left(\{v_\ell \in \mathscr{A}_j\} \right)$ } $\{v_\ell\} \leftarrow$ block-orthonomalization$(N_b, \{v_\ell\})$
6    $\{v_\ell \in \mathscr{A}_j\} \leftarrow \{v_\ell \in \mathscr{A}_j / \|v_\ell \in \mathscr{A}_j\|\}$
7 $P_\ell \leftarrow$ block-aggregation$(N_b, \{v_\ell\})$
8 $D_{\ell+1} \leftarrow P_\ell^\dagger D_\ell P_\ell$
9 **if** $\ell < N_\ell - 1$ **then**
10    $\textbf{Setup}(\ell + 1)$

---

such that the red blocks come first, we obtain

$$D = \begin{pmatrix} D_{rr} & D_{rb} \\ D_{br} & D_{bb} \end{pmatrix},$$ (3.58)

where D_{rr} and D_{bb} are block diagonal matrices filled with the respective subdomains, while D_{rb} and D_{br} connect the neighboring blocks, as schematically depicted in Fig. 3.9.
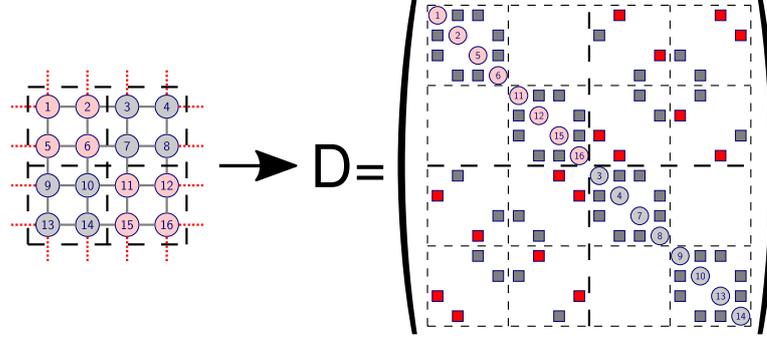
Figure 3.9: Schematic representation of the red-black reduced Dirac operator. The same caption of Fig. 3.1 applies here. Differently from Fig. 3.1, the Dirac operator D as been reordered accordingly to a red-black block decomposition of the lattice.

The eigenmodes of the blocks have a higher cut-off than the full operator, given by $p_v = \pi/L_b$ in the free case due to the Dirichlet boundary conditions. If now the role of the smoother is to reduce the UV-modes, a natural choice for the operator M in the error propagation of Eq. (3.31) is given by the inverse of the block operators $B_{rr}(\mu)$ and $B_{bb}(\mu)$ defined as

$$B_{rr} = \begin{pmatrix} D_{rr}^{-1} & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad B_{bb} = \begin{pmatrix} 0 & 0 \\ 0 & D_{bb}^{-1} \end{pmatrix}.$$ (3.59)

They can be combined in two ways, known respectively as additive and multiplicative SAP,

$$E_{\text{add}} = (I - M_{\text{add}}D) = (I - (B_{rr} + B_{bb})D),$$ (3.60)

$$E_{\text{mul}} = (I - M_{\text{mul}}D) = (I - B_{rr}D)(I - B_{bb}D),$$ (3.61)

where

$$M_{\text{mul}} = B_{rr} + B_{bb} - B_{rr}DB_{bb} = M_{\text{add}} - B_{rr}DB_{bb}.$$ (3.62)

Thus M_{mul} is more computational demanding than M_{add} but numerical tests show a better time to solution when the first is employed [106]. In practice, the inversion of the blocks on the diagonal of B_{rr} and B_{bb} is performed approximately with small computational cost by a few steps of an iterative method like the Minimal Residual method [80, 86]. Note that we fix the block size to coincide with the aggregates on each level of the multigrid hierachy.

3.5.2 Krylov subspace methods

A Krylov subspace method preconditioned by the chosen multigrid approach has to be a flexible algorithm, since the smoother as well as the solver on the coarse grid system are non-stationary processes. Flexible solvers, which have been employed in multigrid preconditioning for lattice QCD are the flexible BiConjugate Gradient Stabilized method (BiCGStab, [108]), Generalized Conjugate

Residual (GCR, [109]) and the Flexible Generalized Minimal RESidual (FGMRES, [110]) solver. In the DD-αAMG approach a FGMRES solver is used for the inversion of the fine grid operator and in the K-cycle for the inversion of the coarser operators except for the coarsest. The latter is inverted by even-odd preconditioned GMRES, i.e. GMRES is run for the even-odd reduced system $\hat{D}\phi_o = \eta_o - D_{oe}D_{ee}^{-1}\eta_e$ with $\hat{D}$ from Eq. (2.110) and then ϕ_e in the solution $[{\phi_e \atop \phi_o}]$ of $D[{\phi_e \atop \phi_o}] = [{\eta_e \atop \eta_o}]$ is retrieved as $\phi_e = D_{ee}^{-1}(\eta_e - D_{eo}\phi_e)$. Even-odd preconditioning is also used in the smoother when inverting the blocks. In both cases a speed-up of about 50% compared to the full operator is achieved. This results in a speed-up of about 20% in the full multigrid procedure.

In general, the accuracy of the coarse grid inversions may be very much relaxed as compared to the target accuracy of the fine grid inversion. Indeed, for DD-αAMG with the K-cycle strategy, optimal results are obtained when requiring the (approximate) inversions of the coarser operators to reduce the residual by just one order of magnitude.

3.5.3 Iterative adaptive setup procedure

In the setup phase we have to compute a sufficient number of approximate low modes v_i, which, when chopped into aggregates, will approximate the IR-modes thanks to the local coherence property. The setup procedure is iterative and adaptive with several differences compared to the null-vectors introduced in Sec. 3.4.2. The iterative setup is obtained considering a set of vectors $\{v_i^{(k)}\}$ where $i = 1, \ldots, N_v$ runs over the number of vectors and $k = 1, \ldots, n_{\text{setup}}$ on the number of setup iterations. The vectors are connected by the relation

$$v_i^{(k)} = D^{-1}v_i^{(k-1)} = D^{-k}v_i^{(0)} \quad \text{with } v_i^{(0)} \text{ chosen randomly.} \tag{3.63}$$

We refer to these vectors as *test-vectors*. The test-vectors $v_i^{(k)}$ converge to the eigenvectors with eigenvalues of smallest modulus. In practice, in order to maintain numerical stability, after each iteration k the test-vectors $v_i^{(k)}$ spanning the space of approximate IR-modes have to be orthonormalized, and D^{-1} is replaced by a multigrid iteration with prolongation operator constructed from the current set of approximate low modes $v_i^{(k)}$. For $k = 1$, where a multigrid hierarchy is not yet available, we just apply some steps of the SAP smoother. This approach results in a iterative self-adapting procedure where the multigrid hierarchy is improved while using it to expose the small eigenmodes. Typically, a small number of setup iterations n_{setup} is sufficient. The setup procedure is listed in Alg. 7.

3.5.4 Software implementation

The DDalphaAMG solver [111], has been released under GNU General Public License. This software package includes an implementation of the DD-αAMG for clover Wilson fermions as described in Ref. [80]. The implementation is of production code quality, it includes a hybrid MPI/openMP parallelization, state-of-the-art mixed precision and odd-even preconditioning approaches and also SSE3 optimizations. Implementation details can be found in Ref. [106]. Based on the DDalphaAMG code we have developed a version, which supports twisted mass fermions [5]. Additionally to the support to twisted mass fermions, as described in the next chapter, many features have been added to the solver making it an independent library to be used in our production codes.

Algorithm 7: $P_\ell = \textbf{Setup}(\ell = 1, i = 0)$

Input: $\ell < N_\ell$ level and i setup iteration. First call with $\ell = 1$ and $i = 0$.
Output: P_ℓ
Parameters: N_{setup}, N_{smooth}

1   **for** $j = 1$ **to** N_v **do**
2   **if** $i = 0$ **then**
3       $v_\ell^j = $ random vector
4       $x = 0$
5   **for** $k = 1$ **to** N_{smooth} **do**   Smoother used in $v_\ell^j \leftarrow D_\ell^{-1} v_\ell^j$
6         $x \leftarrow x + M_\ell(v_\ell^j - D_\ell x)$
7       $v_\ell^j \leftarrow x$
8   **else**
9       $v_\ell^j \leftarrow \textbf{K-Cycle}(\ell, v_\ell^j);$ DD-αAMG used in $v_\ell^j \leftarrow D_\ell^{-1} v_\ell^j$
10   $\{v_\ell\} \leftarrow$ block-orthonomalization$(N_b, \{v_\ell\})$
11   $P_\ell \leftarrow$ block-aggregation$(N_b, \{v_\ell\})$
12   $D_{\ell+1} \leftarrow P_\ell^\dagger D_\ell P_\ell$
13   **if** $\ell < N_\ell - 1$ **then**
14     $\textbf{Setup}(\ell + 1, i)$ Iterative adaptive setup
15   **else if** $i < N_{setup}$ **then**
16     $\textbf{Setup}(1, i + 1)$

4. DD-αAMG for twisted mass fermions

Algebraic multigrid approaches in LQCD are known for a variety of fermion discretizations, such as for the Wilson fermion discretization [14, 80, 98], as a preconditioner for Neuberger overlap fermions [112], for Domain Wall fermions [113] or for staggered fermions [114]. In this work [1, 3], we have extended the DD-αAMG [80] approach to twisted mass Wilson fermions. In this chapter we discuss the construction of the coarse-grid operators for twisted mass fermions and present numerical results on the performance of DD-αAMG for the twisted mass operators.

4.1 Coarse-grid operators for Twisted Mass fermions

The degenerate and non-degenerate TM operator, $D_{\mathrm{TM}}(\mu)$ and $D_{\mathrm{ND}}(\bar{\mu}, \bar{\varepsilon})$ have been introduced in Sec. 2.6. In this section, we construct appropriate coarse-grid operators and preserve properties of the twisted mass operators that successfully deflate the low modes of the target linear system. We also motivate our choices with theoretical arguments.

4.1.1 Degenerate twisted mass operator

The coarse-grid degenerate twisted mass operator in a multigrid approach is defined as, cf. Eq. (3.29),

$$D_{\mathrm{TM},c} = R\left(D + i\mu\Gamma_5\right)P = RDP + i\mu R\Gamma_5 P. \tag{4.1}$$

where D is the Wilson operator with mass tuned to its critical value. Since R and P act only on the Wilson operator, as pointed out in Sec. 3.4.1, the natural choice for the restriction operator is $R = (\Gamma_5 P)^\dagger$ since in this way the restriction operator approximates the subspace of the *right*-eigenvectors of the Wilson operator when the prolongation operator P approximates the *left*-eigenvectors. Using this formulation, though, the resulting coarse grid operator $RD_{\mathrm{TM}}P$ is equivalent to $H_{\mathrm{TM},c} = P^\dagger(\Gamma_5 D + i\mu\Gamma_5)P$, which is a complex shifted maximally indefinite operator, cf. Eq. (2.105). Thus such restriction operator should be avoided for twisted mass fermions.

Requiring a commuting relation between P and Γ_5, as done in Eq. (3.54), the coarse grid corrections obtained by $R = (\Gamma_5 P)^\dagger$ and $R = P^\dagger$ are identical, as noted in Sec. 3.4.1. We then use for twisted mass fermions a Γ_5-compatible operator P, i.e. $\Gamma_5 P = P\Gamma_{5,c}$, and $R = P^\dagger$. The coarse grid twisted mass operator used in DD-αAMG is then defined by

$$D_{\mathrm{TM},c} = P^\dagger D P + i\mu\Gamma_{5,c}, \tag{4.2}$$

where $D_{\mathrm{TM},c}$ preserves several important properties of the fine grid operator D_{TM}. It preserves its sparse structure in that only neighboring aggregates are coupled. The square coarse grid operator,

$$D_{\mathrm{TM},c}^\dagger D_{\mathrm{TM},c} = D_c^\dagger D_c + \mu^2, \tag{4.3}$$

has eigenvalues bounded from below by μ^2. And there is a $\Gamma_{5,c}$-symmetry which reproduces the

connection between the u- and d-quark operators, cf. Eq. 2.102, as

$$D^\dagger_{\text{TM},c}(\mu) = P^\dagger (D + i\mu\Gamma_5)^\dagger P = P^\dagger(\Gamma_5 D\Gamma_5 - i\mu\Gamma_5)P = \Gamma_{5,c} D_{\text{TM},c}(-\mu)\Gamma_{5,c}. \tag{4.4}$$

We construct the operator P using the DD-αAMG setup procedure, see Sec. 3.5.3. In practice we find that the same prolongation operator P constructed for $D_{\text{TM}}(\mu)$ can be effectively used for coarsening all the operators $D_{\text{TM}}(\pm\rho)$ where $\rho \geq \mu$. As shown in Sec. 4.3.3 the time to solution is affected only slightly by this choice. This can be explained by the connection between right-handed eigenvectors v_R of the degenerate TM operator $D_{\text{TM}}(\mu)$ and the left-handed eigenvectors $v_L = v_R^\dagger\Gamma_5$ of $D_{\text{TM}}(-\mu)$, which reads

$$D_{\text{TM}}(\mu)v_R = \lambda v_R \quad \Longleftrightarrow \quad v_R^\dagger\Gamma_5 D_{\text{TM}}(-\mu) = v_R^\dagger\Gamma_5\lambda^\dagger. \tag{4.5}$$

Thanks to the Γ_5-compatibility of the aggregation, this is also true for the coarse operator. Thus, the eigenspaces of $D_{\text{TM}}(\pm\mu)$ are connected and a prolongation operator P constructed for $D_{\text{TM}}(\mu)$ captures the low modes of $D_{\text{TM}}(\mu)$ when acting on the right while acting from the left captures the low modes of $D_{\text{TM}}(-\mu)$.

Although the dimension of the coarse grid operator is reduced, it can develop a large number of small eigenvalues close to μ. This can critically slow down the convergence of a standard Krylov solver to be used on the coarsest grid such that the time spent in the coarsest operator inversions dominates by far the overall inversion time even though only poor accuracy is required. We therefore keep the option to decrease the density of small eigenvalues of the coarsest grid operator by increasing the twisted mass parameter by a factor of δ given by

$$D_c(\mu,\delta) = D_c + i\delta \cdot \mu\Gamma_{5,c} \tag{4.6}$$

with $\delta \geq 1$. We will analyze the effect of δ in detail in Sec. 4.3.2.

4.1.2 Non-degenerated twisted mass operator

The idea behind adapting the DD-αAMG approach to the non-degenerate twisted mass operator $D_{\text{ND}}(\bar{\mu}, \bar{\varepsilon})$ is based on preserving the $(\Gamma_5 \otimes \tau_1)$-symmetry, cf. Eq. (2.114), on the coarse grid. We define the non-degenerate coarse grid operator as

$$D_{\text{ND},c}(\bar{\mu}, \bar{\varepsilon}) = P^\dagger_{\text{ND}} D_{\text{ND}}(\bar{\mu}, \bar{\varepsilon}) P_{\text{ND}}. \tag{4.7}$$

with P_{ND} being a suitable prolongation operator. If P_{ND} is $(\Gamma_5 \otimes \tau_1)$-compatible, i.e.

$$(\Gamma_5 \otimes \tau_1)P_{\text{ND}} = P_{\text{ND}}(\Gamma_{5,c} \otimes \tau_1), \tag{4.8}$$

it follows that the $(\Gamma_5 \otimes \tau_1)$-hermiticity of D_{ND} in Eq. (2.114) is also preserved on the coarse grid and the coarse grid operator fulfills

$$(\Gamma_{5,c} \otimes \tau_1) D_{\text{ND},c} = P^\dagger_{\text{ND}}(\Gamma_5 \otimes \tau_1) D_{\text{ND}} P_{\text{ND}} = P^\dagger_{\text{ND}} D^\dagger_{\text{ND}}(\Gamma_5 \otimes \tau_1)P_{\text{ND}} = D^\dagger_{\text{ND},c}(\Gamma_{5,c} \otimes \tau_1). \tag{4.9}$$

The property in Eq. (4.8) is satisfied by a prolongation operator P_{ND}, which is Γ_5-compatible and diagonal in flavor space. We choose identical components in flavor space defining $P_{\text{ND}} = P \otimes I_2$.

Thus, we obtain

$$D_{\mathrm{ND},c}(\bar{\mu},\bar{\varepsilon}) = (D_{,c}\otimes I_2) + i\bar{\mu}\,(\Gamma_{5,c}\otimes\tau_3) - \bar{\varepsilon}\,(I_c\otimes\tau_1) = \begin{bmatrix} D_{\mathrm{TM},c}(\bar{\mu}) & -\bar{\varepsilon}\,I_c \\ -\bar{\varepsilon}\,I_c & D_{\mathrm{TM},c}(-\bar{\mu}), \end{bmatrix} \qquad (4.10)$$

which follows from the property $P^\dagger P = I_c$. We note that the flavor and spin components of the coarse operator preserve a similar sparse structure and properties of the fine grid operator $D_{\mathrm{ND}}(\bar{\mu},\bar{\varepsilon})$ in Eq. (2.112).

It follows from Eq. (4.10), that the prolongation operator P has to project onto a subspace, which captures the IR-modes. While P_{ND} is degenerate in flavor space, the low modes of the non-degenerate twisted mass operator are defined in the full space. Thus our solution to Eq. (4.8) i.e. $P_{\mathrm{ND}} = P\otimes I_2$ could spoil the efficiency of the coarse grid correction since the same prolongation operator P has to act on both flavor spaces. A possible solution is to use the prolongation operator P constructed for the degenerate twisted mass operator. This has also the advantage that if the multigrid solver is used during the HMC the same setup, built up for one flavor of the TM Wilson operator, can be reused in any step of the HMC for both light degenerate and heavy non-degenerate sector. Saving additional setup expenses makes the usage of the multigrid in HMC more effective. We motivate this choice as follow:

- In the construction of $D_{\mathrm{ND},c}(\bar{\mu},\bar{\varepsilon})$, see Eq. (4.10), the prolongation operator P is used for constructing both $D_{\mathrm{TM},c}(\pm\bar{\mu})$. As mentioned in the previous subsection, a P constructed for $D(\mu)$ can be successfully employed for coarsening any $D_{\mathrm{TM}}(\pm\rho)$ where $\rho \geq \mu$. Thus when $\bar{\mu} \geq \mu$ we are well motivated to use the same P. This is the case in $N_f = 2+1+1$ simulations where the non-degenerate operator is used in the strange and charm sector, while P is constructed for the degenerate operator used for the light, up and down, sector.

- Considering the case $D_{\mathrm{ND}}(0,\bar{\varepsilon})$ where $\bar{\mu} = 0$ and $\bar{\varepsilon} \neq 0$, it follows that the eigenvalues have a linear dependency in $\bar{\varepsilon}$ and the eigenvectors are degenerate in flavor space. Indeed the relation

$$D_{\mathrm{ND}}(0,\bar{\varepsilon})w_\pm = (\lambda\pm\bar{\varepsilon})w_\pm \quad\text{with}\quad w_\pm = (v,\pm v) \quad\text{and}\quad Dv = \lambda v \qquad (4.11)$$

holds in flavor space. Thus, for $\bar{\mu} = 0$ the eigenspace is invariant under changing $\bar{\varepsilon}$ and it motivates the choice via Eq. (4.8) for the coarse grid projector.

- From the two above points follows that the choice for the projector of Eq. (4.8) is well-motivated for the special cases $\bar{\mu} \neq 0 \wedge \bar{\varepsilon} = 0$ and $\bar{\mu} = 0 \wedge \bar{\varepsilon} \neq 0$. However, in our simulations we have $\bar{\mu} \sim 0.1 \wedge \bar{\varepsilon} \sim 0.1$. In order to verify if the properties mentioned in the previous two items are a good approximation, we study numerically the eigenvalue spectrum when changing the parameters $\bar{\mu}$ and $\bar{\varepsilon}$ on a small lattice of a size of 4^4. The eigenvalue spectrum with positive imaginary part of the ND twisted mass operator is depicted in Fig. 4.1. Note that due to the $(\Gamma_5\otimes\tau_1)$-hermiticity the real-axis is a symmetry axis, thus eigenvalues come in complex conjugated pairs or are real. For the one flavor operator $D_{\mathrm{TM}}(\mu)$ this symmetry is broken for $\mu \neq 0$. This is shown in the left lower panel, where we focus on the spectrum of the low modes. By taking a closer look to the middle panel, the parameter $\bar{\varepsilon}$ for the non-degenerate operator $D_{\mathrm{ND}}(\bar{\mu},\bar{\varepsilon})$ acts on the eigenvalues via a linear shift $\pm\bar{\varepsilon}$ similar to the ideal case given by $D_{\mathrm{ND}}(0,\bar{\varepsilon})$. Thus, the projector given by Eq. (4.8) should project on the small eigenvalues of $D_{\mathrm{ND}}(\bar{\mu},\bar{\varepsilon})$. For our case study, this is indeed the case as depicted in the right panel of Fig. 4.1.
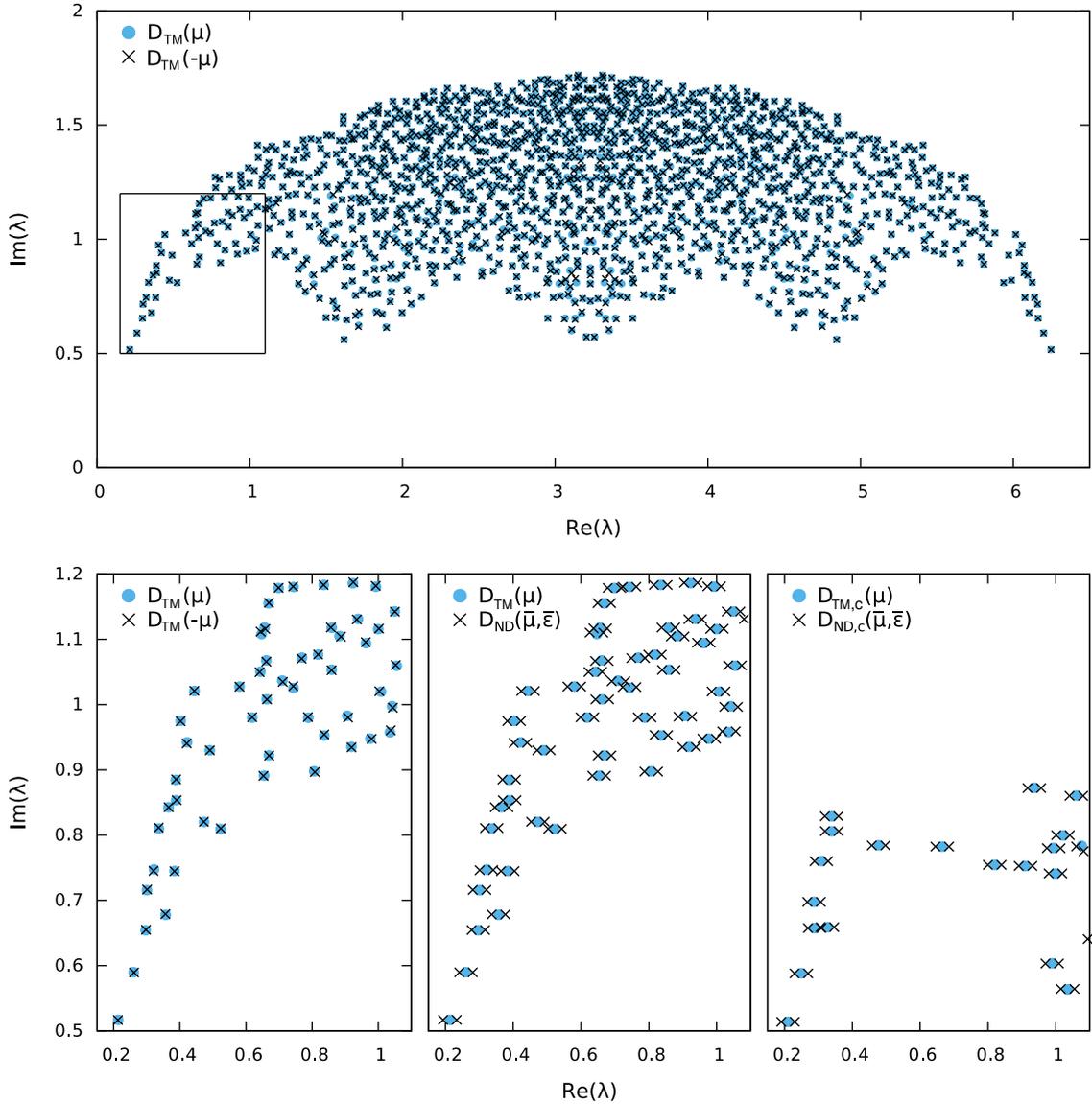
Figure 4.1: Complete spectrum of twisted mass operators on a gauge configuration for a lattice of size 4^4. The twisted mass parameters are $\mu, \bar{\mu}, \bar{\varepsilon} = 0.01$. In the top panel, we depict the spectrum of the degenerate $D_{\text{TM}}(\mu)$ operator folded with respect to the imaginary axis. In the bottom left panel we show part of the spectrum close to the origin for $D_{\text{TM}}(\mu)$ for different sign of $\pm\mu$. In the central lower panel we compare the degenerate and non-degenerate TM operator. In the right lower panel, we show the spectrum for the coarse version of both operators.

The spectrum of the coarse grid operator for the degenerate and non-degenerate twisted mass operator display similar features showing that preserving the operator structure in the coarse grid allows to preserve properties of the fine operator.

4.2   Multigrid performance with twisted mass fermions

The DD-αAMG approach uses a wide range of parameters and its time to solution will depend on a good choice of the parameters. The purpose of this section is therefore threefold. We first provide a set of default parameter choices, which in our extensive numerical testing turned out to yield good overall performance. Secondly, we show that for physically relevant configurations, appropriately chosen parameters yield speed-ups of about two orders of magnitude compared to standard methods for twisted mass fermions. Finally, we compare the critical slowing down of the DD-αAMG solver against CG solver. All the presented numerical results have been obtained on SuperMUC phase 2 at the Leibniz Supercomputing Centre, an Intel Haswell Xeon computer on which we used up to 4096 cores. We have also performed runs on JURECA at the Julich Supercomputing Centre and on Cy-Tera at the Computation-based Science and Technology Research Center Cyprus obtaining compatible results. The stopping criterion of the overall iteration was fixed such that the residual is reduced by a factor of 10^9.

4.2.1   Default parameters

Table 4.1 summarizes our default parameter set used for DD-αAMG. The parameter tuning was done for the physical ensemble $cA2.48$, see Tab. 3.1, already used in Sec. 3.1.2.

As compared to the optimal parameters for the Wilson Dirac operator, given in ref. [105], when inverting the TM operator at the physical point turns out that, additionally to the employment of the shift δ, cf. Eq. (4.6), the overall execution time is minimized if one relaxes even further the accuracy of the coarsest-grid correction using here 10^{-1} as stopping criterium. Also the aggregation parameters differ slightly to the ones suggested in ref. [105], namely $N_{v,1/2} = 28$ instead of 24. This is a consequence of the extensive analysis done in Sec. 4.3.1, where the parameter space for the block-aggregation is studied. Here a dependence on the number of right-hand-side (rhs) to be solved is found and in the following we always consider $\mathcal{O}(100)$ and more rhs, if not otherwise stated.

As shown in Tab. 4.2, the setup costs needs to be optimized accordingly to the application. The value $n_{\text{setup}} = 5$ in Tab. 4.1 yields a good time to solution when several inversions ($\mathcal{O}(100)$ and more) with the same operator are desired. The relatively large setup time can be neglected in this case. We did not find that more setup iterations than $n_{\text{setup}} = 5$ yield substantial further improvement in the solve time. On the other hand, when solving for a few rhs, a good balance of setup and solve time has to be found. As explained in Section 3.5.3, in each iteration of the adaptive setup routine the currently available multigrid hierarchy is used to perform one iteration with the multigrid preconditioner on each test vector. As the setup iteration proceeds the test vectors become more rich in low mode components. This also leads to a more ill-conditioned coarse operator D_c and higher iteration counts on the coarse grid (cf. Ref. [80]), i.e. one can observe a higher cost per setup iteration as the setup proceeds. This can be seen in Tab. 4.2 where the first two setup iterations together costed 5.5 core-hrs, the third costed 5.7 core-hrs and from the fourth on they have an average cost of 6.5 core-hrs each. Therefore a smaller number of setup iterations like $n_{\text{setup}} = 3$ might be more suitable when few rhs are solved. Note that we use the parameters from Table 4.1 also in the multigrid preconditioner during

	parameter		optimal
Multigrid	number of levels	n_ℓ	3
	number of setup iterations	n_{setup}	5
	number of test vectors on level 1	$N_{v,1}$	28
	number of test vectors on level 2	$N_{v,2}$	28
	size of lattice-blocks for aggregates on level 1	$V_{b,1}$	4^4
	size of lattice-blocks for aggregates on level ℓ, $\ell > 1$	$V_{b,2}$	2^4
Solver	mixed precision FGMRES		
	relative residual tolerance (restarting criterion)		10^{-6}
Smoother (SAP)	red-black multiplicative SAP		
	size of lattice-blocks on level 1		4^4
	size of lattice-blocks on level ℓ, $\ell > 1$		2^4
	number of post-smoothing steps		4
	MINRES iterations to invert the blocks		3
K-cycle	with single precision FGMRES		
	restart length		5
	number of maximal restarts		2
	relative residual tolerance (stopping criterion)		10^{-1}
Coarsest grid	solved by even-odd preconditioned GMRES		
	twisted mass parameter	$\mu_{\text{coarse}} = \delta \cdot \mu$	$5 \cdot \mu$
	restart length		100
	number of maximal restarts		5
	relative residual tolerance (stopping criterion)		10^{-1}

Table 4.1: The parameter set used in DD-αAMG, obtained by parameter tuning for the TM fermion ensemble cA2.48 [76].

the setup phase when running the method itself on each test vector (cf. the call of K-cycle in Alg. 7 at line 9).

Setup iters n_{setup}	Setup time [core-hrs]	Iteration cost [core-hrs]	Inversion time [core-hrs]	Inversion iterations
2	5.5	–	4.6	48
3	11.2	5.7	1.9	20
4	17.8	6.6	1.7	18
5	24.4	6.6	1.6	17
6	30.8	6.4	1.7	18
7	37.2	6.4	1.7	18

Table 4.2: Analysis on the number of setup iterations using a configuration from the ensemble cA2.48. The same parameters of Tab. 4.1 are used changing n_{setup} as reported here in the table.

As mentioned in Sec. 4.1.1, we use the same subspace for the u-quark with $+\mu$ and the d-quark with $-\mu$, i.e. we run the setup phase only once for both quarks. Although the eigenspace changes, numerically we do not find a large difference. This also saves computing time for many applications with twisted mass fermions. When solving at larger pion mass than the physical one, a smaller number of smoothing step can improve the time to solution. For the non-degenerate operator at the strange and charm quark mass we find that 2 post-smoothing step are enough for the optimal time to solution. Also the shift δ of the twisted mass term on the coarsest-grid depends on the pion mass and for large

pion masses a value between $1 \leq \delta < 5$ should be used, where 5 is the optimal value at the physical point as suggested in Table 4.1. For the non-degenerate operator at the strange and charm quark mass we find that $\delta = 1$ gives the optimal time to solution. Moreover, for larger pion masses one can, in principle, improve the generation of the subspace of small eigenmodes in the setup by using a smaller twisted mass parameter μ, which helps to probe the small eigenmodes more rapidly within the setup phase. However, at least at the physical point, we do not find a significant improvement by using different mass parameters.

4.2.2 Comparison with CG and CG-eDe for the degenerate twisted mass operator

To put the improvements for twisted mass simulations in perspective, we start with an experiment for Wilson Dirac fermions, thus complementing the results from Ref. [105] for our target machine. We use the default parameter set from Ref. [105] and a configuration from ensemble *VII* [115] with a pion mass of $am_\pi = 0.05786$, a lattice spacing of $a = 0.0071$ fm and a lattice volume of $V = 64 \times 48^3$. Table 4.3 presents a comparison of the inversion times of the Conjugate Gradient (CG) solver, the CG solver with exact deflation (CG-eDe) and the DD-αAMG solver. A similar comparison between CG and CG-eDe has been shown in Sec. 3.1.3 for twisted mass fermions at the physical point.

The conjugate gradient solver, cf. Sec. 3.1.1, requires a positive definite hermitian matrix, which is obtained by solving the linear system with the squared even-odd reduced operator given by $\hat{D}^\dagger \hat{D} x = b'$ with $b' = \hat{D}^\dagger b$. This squares the condition number of the involved matrix.

The CG-eDe and the DD-αAMG solver involve a setup phase, which has to be done for each new configuration once before the linear system is solved. In the case of the CG-eDe solver, $\mathcal{O}(100)$ lowest eigenmodes of the squared even-odd reduced Dirac operator are calculated (here we use 800 for the case of Wilson case and 1600 for the twisted mass case). The number of eigenvectors in CG-eDe is obtained by optimizing the time to solution (setup + inversion time) for computing $\mathcal{O}(1000)$ rhs. Indeed the setup phase is extremely expensive which makes CG-eDe inefficient for a small number of rhs. The low mode computation is done by using the publicly available package ARPACK with tmLQCD [90].

Solver	Setup time [core-hrs]	Inversion time [core-hrs]	Total iteration count of the fine grid solve	Total iteration count of coarse grid solvers
CG	˘	174.8	26 937	˘
CG-eDe	1 527.4	5.4	649	˘
DD-αAMG	13.3	0.9	16	2 988

Table 4.3: Results for clover Wilson fermions[†]. Comparison of CG (tmLQCD), CG-eDe (tmLQCD+ARPACK) with 800 eigenvectors and DD-αAMG with parameters from Ref. [105]. The results are for a $48^3 \times 64$ lattice from ensemble *VII* of Ref. [115] with $m_\pi = 0.1597(15)$ GeV.

Table 4.3 shows that in the case of Wilson Dirac fermions, the DD-αAMG solver speeds up the time to solution by roughly a factor of 200 compared to CG and roughly by a factor five compared to the CG-eDe solver. When including the setup time, DD-αAMG is roughly a factor of 12 faster than CG for one rhs, while CG-eDe is not competitive due to its computationally demanding setup.

[†]In Tables 4.3 and 4.4 the timings for CG and eigCG have been normalized to 1.0 Gflop/s per core (average of standard performance ~ 0.7 Gflop/s and optimal performance ~ 1.3 Gflop/s); the rhs of the equation $D\psi = b$ has been randomly generated and in all the cases the propagator ψ has been computed to a relative precision of 10^{-9}.

Solver	Setup time [core-hrs]	Inversion time [core-hrs]	Total iteration count of the fine grid solve	Total iteration count of coarse grid solvers
CG	˘	338.6	34 790	˘
CG-eDe	6 941.1	9.8	695	˘
DD-αAMG for $n_{\text{setup}} = 3$	7.7	2.5	28	16 619
DD-αAMG for $n_{\text{setup}} = 5$	38.3	1.5	15	11 574

Table 4.4: Results for TM fermions[†]. Comparison of CG (tmLQCD), CG-eDe (tmLQCD+ARPACK) with 1600 eigenvectors and DD-αAMG. The results are computed for the *cA*2.48 ensemble, see Tab. 3.1.

(a) clover fermions

(b) twisted mass fermions
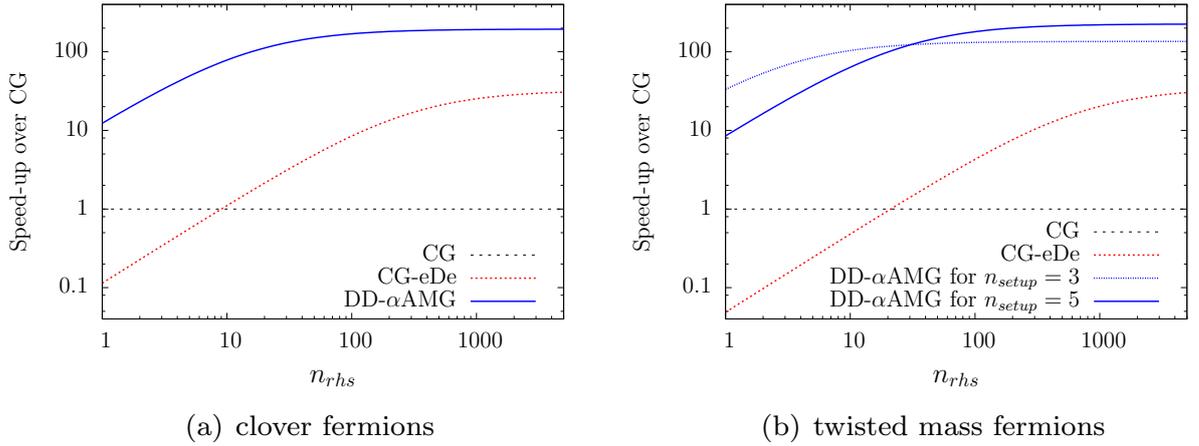
Figure 4.2: Speed-up over CG using the results of Table 4.3 and Table 4.4 depicting $f(n_{rhs}) = (t_{setup} + n_{rhs} \cdot t_{setup})/(n_{rhs}t_{CG})$.

In the twisted mass fermion case, we are able to achieve the same speed-ups as for the Wilson fermion case. This it not straightforward as it requires the coarsest grid twisted mass μ_c to be chosen different from the fine grid twisted mass μ. We choose $\mu_c = 5\mu$ (cf. Table 4.1) for reasons that will be explained in Section 4.3.2. The results for twisted mass fermions are shown in Table 4.4. We are able to achieve a speed-up in inversion time of roughly a factor of 220 compared to CG and roughly by a factor six compared to the CG-eDe solver, with the time for the setup being almost a factor of 100 less as compared to CG-eDe. For $n_{\text{setup}} = 3$ we used the parameters from Table 4.1 except for the number of test vectors on the fine level $N_{v,1}$ being 20 instead of 28. For $n_{\text{setup}} = 5$ we used a further optimized set which is given in the last row of Table 4.5. This yields another factor of 1.5 speed-up in the inversion time at the expense of increasing the setup time.

In Figure 4.2 we show the speed-up in total time (setup+solve) compared to one CG solve for DD-αAMG and CG-eDe as a function of the number of rhs n_{rhs}. The difference between the two blue curves in the figure on the right hand side is due to results from a different number of setup iterations (3 and 5) and a different number of test vectors (20 and 28), where 3 setup iterations are optimal for few rhs (blue dotted line) and 5 setup iterations for many rhs (blue line). In summary, the results for twisted mass fermions show that for one rhs DD-αAMG is roughly 30 times faster than CG for one rhs, 120 times faster for twenty rhs and 220 times faster for a thousand rhs.

4.2.3   Solution to the critical slowing down of the degenerate TM operator

We have discussed the critical slowing down of standard Krylov solvers in Sec. 3.1.2. It arises due to the singularity of the twisted mass operator at $\mu = 0$ when the Wilson mass is tuned to its critical value, i.e. for simulations at maximal twist. As depicted in Fig. 4.3, DD-αAMG definitely resolve the critical slowing down problem.
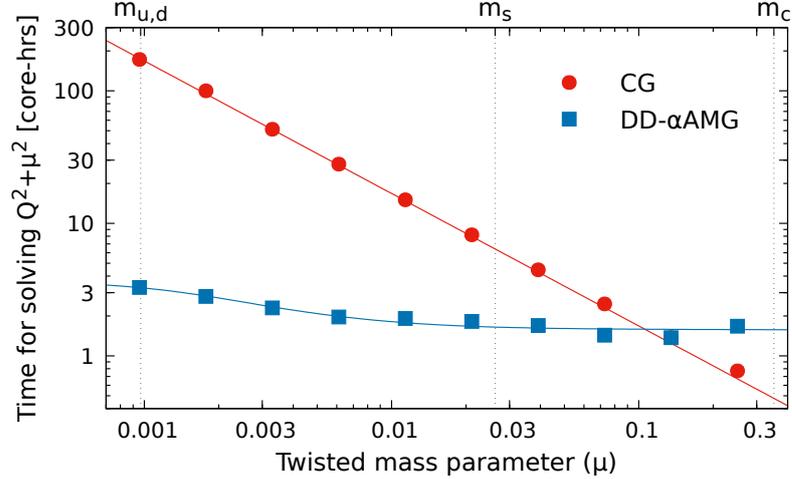
Figure 4.3: Comparison between time to solution for computing the twisted mass fermion propagator at different quark mass using the odd-even (oe) CG solver and the DD-αAMG approach. The value of the light, strange and charm quark masses are shown by the vertical lines with the labels $m_{u,d}$, m_s and m_c, respectively.

The time to solution of CG have been fitted in Sec. 3.1.2. The fit approach used for the time to solution of the multigrid solver is motivated by the convergence of the general minimal residual solver (GMRES) [116]. We use the functional form

$$t_{\text{GMRES}}(K) = t_{\text{init}} + t_{\text{iter}} \frac{\ln(2/\rho)}{\ln(K) - \ln(K-1)}, \tag{4.12}$$

where t_{init} and t_{iter} are the initialization and iteration costs respectively. In case of MG we employ a modified condition number

$$K_{\text{MG}} = \frac{\Lambda_{\max}^2 + \mu^2}{\Lambda_{\min}^2 + \mu^2} \tag{4.13}$$

where $\Lambda_{\max}/\Lambda_{\min}$ defines the condition number of the equivalent multigrid deflated Wilson operator. Note that in case of MG we invert twice the non-squared operator D_{TM}, since it performs better when applied to a linear equation with a single Wilson operator[97]. The reason is simply that an implementation for the squared operator is far more complicated, involving next to next neighbor interactions, which increase the complexity of the coarse operator. Thus we fit the MG timings with

$$t_{\text{MG}}(K_{\text{MG}}) = 2 \cdot t_{\text{GMRES}}(\sqrt{K_{\text{MG}}}), \tag{4.14}$$

For the cA2.48, the time per iteration is $t_{\text{iter}} \simeq 7$ core-minutes and we keep as free parameters t_{init}, $\Lambda_{\max}$ and $\Lambda_{\min}$. We find an initialization cost $t_{\text{init}} \simeq 0.6$ core-hours and a condition number for the equivalent multigrid deflated Wilson operator $\Lambda_{\max}/\Lambda_{\min} \simeq 18$ and the values satisfy $\lambda_{\min}(D^\dagger D) > \Lambda_{\max}^2 > \Lambda_{\min}^2 > \lambda_{\min}(D^\dagger D)$. A finite condition number for the deflated Wilson operator proves that the multigrid operator solves the critical slowing down of the degenerate maximally twisted TM operator. The time to solution still slows down approaching the physical point but only slightly and
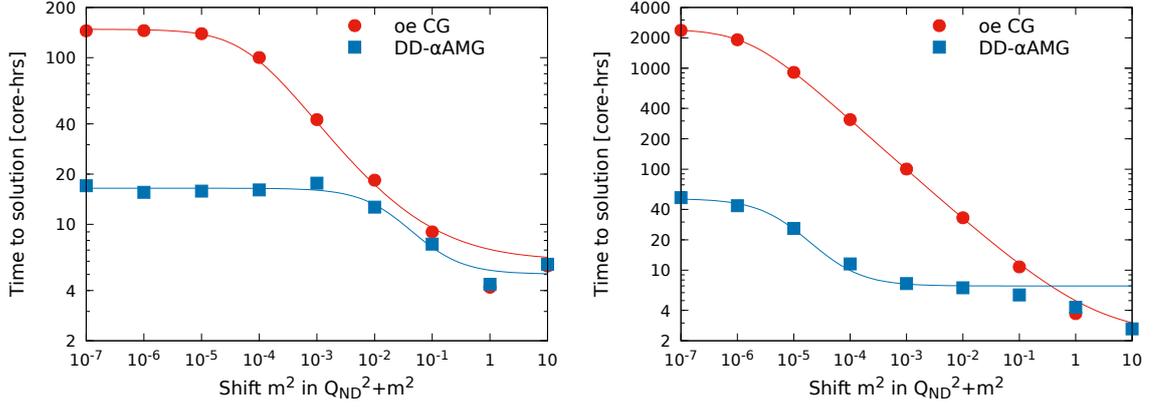
Figure 4.4: Comparison between time to solution for computing the inverse of the squared even-odd reduced ND twisted mass operator at different shift m^2 using the odd-even (oe) CG solver and the DD-αAMG approach. We used the physical test ensemble employing physical strange and charm quark masses (left panel) and physical up and down quark masses (right panel).

not critically. For clarity sake, note that we used the same multigrid parameters for all the mass shifts. The timings at higher mass can be significantly improved by adopting the observation described in Sec. 4.2.1.

4.2.4 Comparison with CG for the non-degenerate operator at increasing mass shift

Here we test the effectiveness of our choice $P_{\mathrm{ND}} = P \otimes I_2$ using a physical test ensemble simulated with N_f=2+1+1 TM fermions, ensemble cB211.64 in Tab. 6.1 at page 94. We compare the performance of the multigrid solver against CG when solving

$$(D_{\mathrm{ND}}^{\dagger} D_{\mathrm{ND}} + m^2)x = b. \tag{4.15}$$

for several shifts in m^2. We obtain the results shown in Fig. 4.4. On the left we have tuned $\bar{\mu}, \bar{\varepsilon}$ for giving the strange and charm quark mass accordingly to Eq. (2.113). On the right we have tuned them to give the mass shift of the physical up and down quark masses, i.e. useful to simulate a $N_f = 1 + 1 + 1 + 1$ ensemble. As in the previous section we have fit the time to solution using Eqs. (3.26) and (4.14).

At the physical strange and charm quark masses – i.e. $m^2 \to 0$ see left panel of Fig. 4.4 – we find an order of magnitude speed-up of the multigrid approach compared to the CG solver. Moreover, the ND-multigrid solver is even more effective than the multigrid solver for the degenerate twisted mass operator at physical strange quark mass. This can be seen by comparing the relative speed-up for strange quark mass of $m_q \sim 95$ MeV, shown in Fig. 4.3 for the degenerate TM operator. The relative speed-up for two application of the ND-multigrid solver is comparable with the speed-up of a single application in case of the TM Wilson operator. This shows that the choice of the coarse grid projector, built up from the TM Wilson operator, yields a very effective multigrid approach for the non-degenerate twisted mass operator. This is also confirmed at physical non-degenerated light quark masses, as shown in the right panel of Fig. 4.4. Here we found a speed-up of around two orders of magnitude similar to the case of the TM Wilson operator.

Due to the large parameter set of multigrid approaches, optimization for a specific lattice can become a major task. In the next section we outline our strategy and we gave a set of parameters in Tab. 4.1. For the non-degenerate operator we use this set of optimized parameters with a few

adjustments in case of heavy quark masses. Namely, the shift δ of the TM parameter in the coarse grid is set to unity and the number of smoothing iterations is reduced from 4 to 2.

4.3 Analysis of parameter settings and tuning strategy

As demonstrated in the previous section, well-tuned parameters are important for a good performance and they tend to be stable at least for configurations from a given ensemble. It is therefore advisable to invest some effort for obtaining good parameters. Since the parameter space is large, it cannot be searched exhaustively and there is thus a need for a strategy how to tune the method in practice. Our suggestion for twisted mass simulations is to keep the default parameters given in Table 4.1, but tune the aggregation parameters and the twisted mass parameter on the coarsest level, μ_c. This is justified by the analysis that we explain in the subsequent sections as follows. In Sec. 4.3.1, we present an analysis of the aggregation parameters without tuning μ_c, i.e. we fix $\delta = 1$. In Sec. 4.3.2, we show the benefits obtained by increasing μ_c and we repeat the previous analysis. In Sec. 4.4.1, we demonstrate that also other multigrid approaches can benefit by an increasing μ_c. All the tests are performed on one configuration averaging the time to solution for the u- and d-propagator. This choice is motivated by the stability of the solver presented in the previous section.

4.3.1 Aggregation parameters

Aggregation parameters are the number of the test vectors $N_{v,\ell}$ and the size of the lattice-blocks $V_{b,\ell}$ on each level ℓ. They should be tuned simultaneously since they define the size of the coarser Dirac operator and consequently the size of the projected subspace. In the present analysis, the solver is restricted to a 3-level implementation. We do not find an improvement in the time to solution by using a 4th level, which is also the result found in Ref. [105] for the Wilson operator with similar lattice sizes. On the other hand, the inversion time increases when just a 2-level multigrid method is used.

To optimize the aggregation parameters on the first level, we fix the values of the parameters on the second level for which we find a block size of $V_{b,2} = 2^4$ and a number of test vectors of at least 28. More precisely the latter is set as $N_{v,2} = \max(28, N_{v,1})$ due to software limitations which require the number of test vector to increase. We analyze how the setup time and time per inversion change dependently on the block size $V_{b,1}$ and the number of test vectors $N_{v,1}$. The results are depicted in Figure 4.5, for the cA2.48 ensemble and a scaling parameter $\delta = 1$ on all levels has been used.

We find that every block size has an optimal number of test vectors as shown in Figures 4.5(a) and 4.5(b) that minimizes the time to solution. By fitting the data to a polynomial of order 2 in $N_{v,1}$ we estimate the minimums and we find that the optimal number of test vectors grows approximately linearly for block sizes $V_{b,1} < 8 \cdot 4^3$ with the block volume

$$N_{v,1} = \alpha + \beta V_{b,1},$$ (4.16)

which is shown in Figure 4.5(c). This indicates a non-trivial connection between the fine-grid size and the dimension of the coarse-grid operator given by

$$\dim(D_{c,1}) = 2 \left(\frac{\alpha}{V_{b,1}} + \beta \right) V.$$ (4.17)

This implies that the optimal size of the coarse-grid operator increases linearly with the volume V at
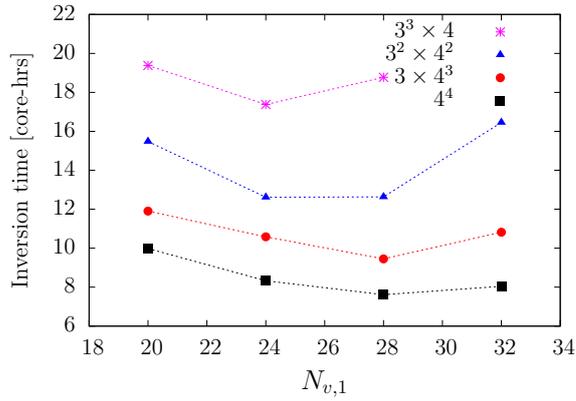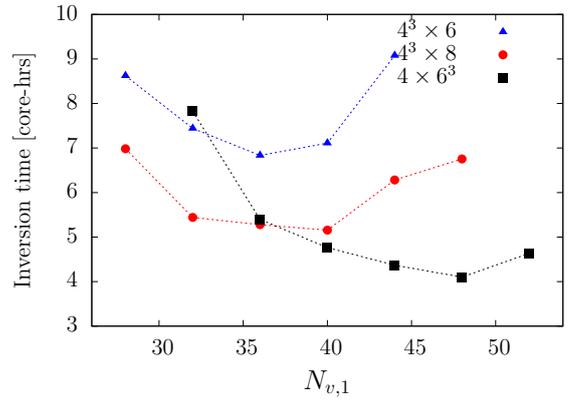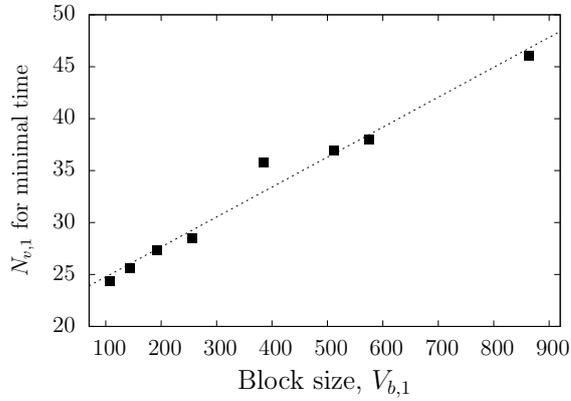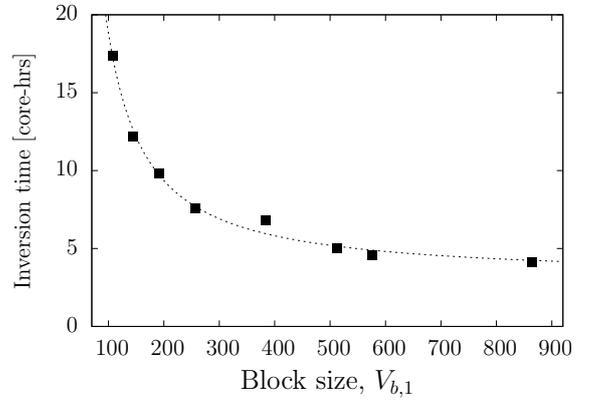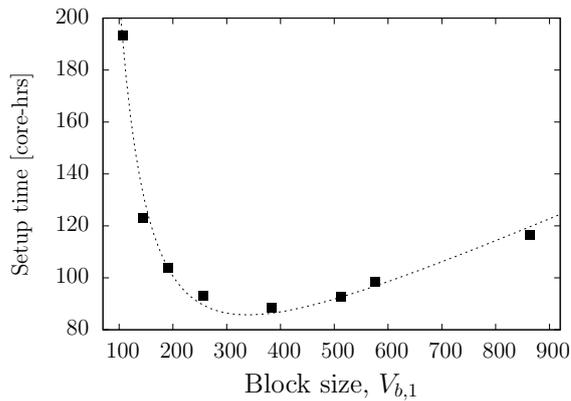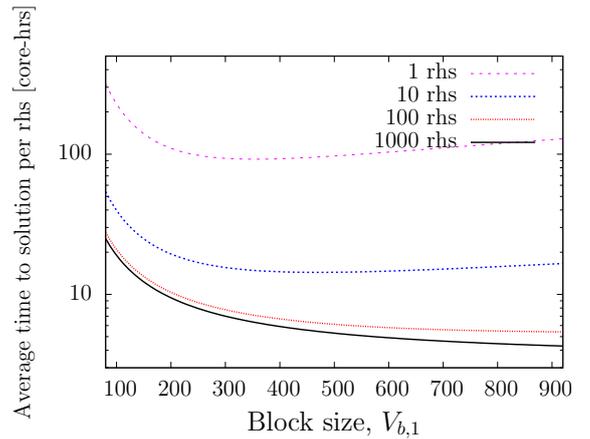
Figure 4.5: Analysis of the aggregation parameters for the *cA*2.48 ensemble. For each set of raw data presented in (a) and (b) the position and value of the minimum has been extracted with a parabolic interpolation and displayed in (c) and (d), respectively. In (d), (e) and (f), $N_{v,1}$ and $V_{b,1}$ are connected according to the minimums found in (c), i.e. (d) shows the inversion time and (e) the setup time, both for the minimums from (c). The estimated average total time per right hand side $(t_{\text{setup}} + n_{rhs}t_{solve})/n_{rhs}$ is shown in (f). The fitting functions are explained in the text.
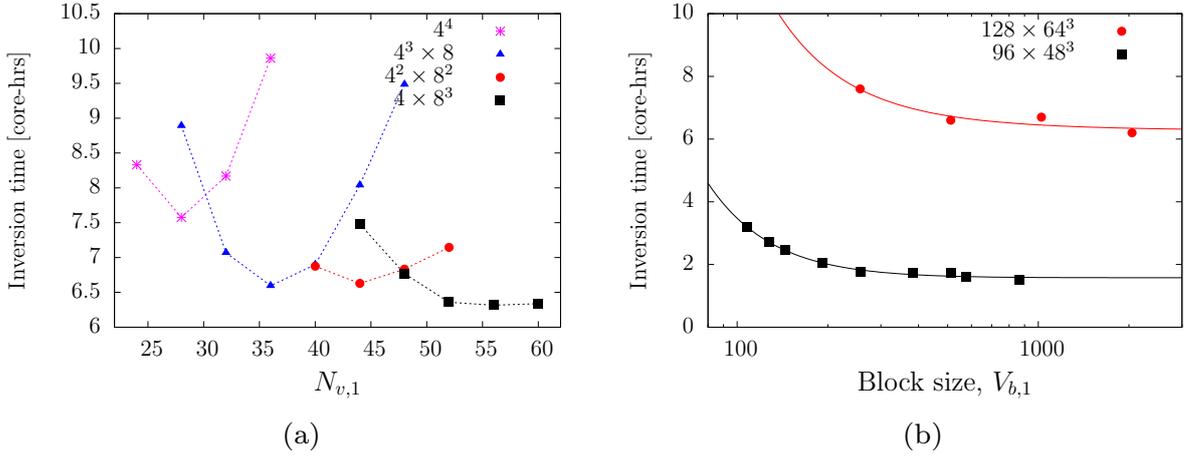
Figure 4.6: Analysis of aggregation parameters for $cA2.64$, see Tab. 6.1, and $\delta = 5$ on the coarsest grid. The sets of raw data are presented in (a); from each set, the value of the minimum has been extracted with a parabolic interpolation and plotted with red points in (b). The black points are for the ensemble $cA2.48$ and $\delta = 5$ on the coarsest grid.

fixed aggregation block size. For larger block sizes the behavior deviates from this linear dependence, however. For this case the minimum of the time to solution is already reached for a smaller number of test vectors. We can interpret this complex behavior as an insight into the non-trivial link between coarse grid size and local coherence.

The time to solution of the multigrid method is dominated by the solves with the coarse grid operator. We use a K-cycle and a 3-level approach where, due to fixing the blocks of the coarsest grid, it follows that $\dim(D_{c,1}) \propto \dim(D_{c,2})$. The numerical effort needed for a matrix-vector multiplication involving $D_{c,1}$ depends linearly on $\dim(D_{c,1})$. Allowing for an additional second order term we model the time for an inversion as

$$t_{solve} = \gamma' + \delta' \dim(D_{c,1}) + \varepsilon' \dim(D_{c,1})^2. \tag{4.18}$$

By using the dependence on $V_{b,1}$ of Eq. (4.17) we can rewrite t_{solve} with

$$t_{solve} = \gamma + \delta V_{b,1}^{-1} + \varepsilon V_{b,1}^{-2}. \tag{4.19}$$

Within this model, connecting $N_{v,1}$ and $V_{b,1}$ according to Eq. (4.16) and Fig. 4.5(c), the time to solution can be fitted very well up to $V_{b,1} \approx 8^2 \times 4^2$ as shown in Fig. 4.5(d). In Fig. 4.6, we display the data obtained for the ensemble $cA2.64$ with a lattice size of 128×64^3. The other lattice parameters are the same as those for $cA2.48$. In contrast to the discussed analysis of $cA2.48$, the coarse grid scale factor is set to $\delta = 5$ on the coarsest grid. The full analysis on the dependence of the algorithm on δ is reserved for the next section. For both lattice sizes the inversion times reach the minimum for the same block volumes. The asymptotes γ from Eq. (4.19) are given by 1.64(13) for the $cA2.48$ ensemble and by 6.3(5) for the $cA2.64$ ensemble, which is an increase of $V^{5/4}$. As mentioned above, for larger block sizes, the linear connection of $N_{v,1}$ and $V_{b,1}$ does not hold and the solution time is increased. This is also observed for the time spent in the setup. We also observe that with increasing lattice volume the region where the linear dependence holds is shifted.

During the setup procedure the solver is applied on $N_{v,1}$ test vectors for several iterations. Here, the test vectors are orthonormalized at the end of each iteration and used for building the multigrid

hierarchy. We model the setup time by allowing for a linear and quadratic term as

$$t_{\text{setup}} = \zeta + \eta \, N_{v,1} \, t_{solve} + \theta \dim(D_{c,1}),$$ (4.20)

where the amount of computation which does not involve the solver scales with $\dim(D_{c,1})$ at leading order. Fig. 4.5(e) shows that the measured timings are indeed very well described by the Ansatz given in Eq. (4.20). We remark that for block sizes $V_b \sim 4^4$ up to $6^2 \times 4^2$ the time to solution shows a relatively large plateau, which makes the timings for the multigrid solver relatively stable. Further fine tuning in this region would lead only to small improvements.

The optimal choice for $N_{v,1}$ and $V_{b,1}$ depends on the number of rhs n_{rhs}. The total time is given by

$$t_{total} = t_{\text{setup}} + n_{rhs} t_{solve},$$ (4.21)

and we find a non-trivial dependence of $V_{b,1}$ on n_{rhs} for an optimal time to solution, as depicted in Figure 4.5(f). This motivated our suggestion to consider two different values for n_{setup}, depending on the number of rhs.

4.3.2 Tuning the coarse grid scale factor δ

Block size $V_{b,1}$	Test vectors $N_{v,1}$	δ for the coarsest μ $\mu_c = \delta\mu$	Setup time [core-hrs]	Inversion time [core-hrs]
$3^3 \times 4$	24	7.8	31.1	3.19(3)
2×4^3	24	6.6	27.5	2.71(5)
$3^2 \times 4^2$	24	8.6	23.5	2.47(2)
3×4^3	28	5.4	28.0	2.04(5)
4^4	28	5.2	22.2	1.75(4)
$4^3 \times 6$	36	4.5	37.4	1.74(6)
$4^3 \times 8$	40	4.0	40.7	1.73(5)
$4^2 \times 6^2$	40	4.1	37.7	1.59(6)
4×6^3	44	4.0	38.3	1.52(4)

Table 4.5: Summary of δ parameters yielding the optimal solve time for various block sizes $V_{b,1}$. The shown numbers were computed for the $cA2.48$ ensemble and a relative residual tolerance of 10^{-9}. The number of test vectors was chosen as in Figure 4.5(c) according to the block size $V_{b,1}$. All other parameters were fixed to the values in Table 4.1.

At physical quark masses, the density of the low-lying eigenvalues for the twisted mass operator increases compared to the Wilson Dirac operator, see Figs. 2.4 and 2.5. Densely populated low eigenvalues slow down the iteration of the Krylov subspace solvers on the coarsest grid and thus of the whole multigrid method. This is much more pronounced for the twisted mass operator than it is for the Wilson Dirac operator. For the latter, Tab. 4.3 reports a total of around $3\,000$ coarse grid iterations for the considered ensemble, whereas data depicted in Fig. 4.7(a) show that for $\delta = 1$ we have roughly 40 times more coarse grid iterations in the twisted mass case, an unexpectedly large increase.

When using a larger twisted mass value on the coarsest gird operator, given by $\mu_c = \delta\mu$ with $\delta \geq 1$, we make the small eigenvalues less dense in the low-lying part of the spectrum. This speeds
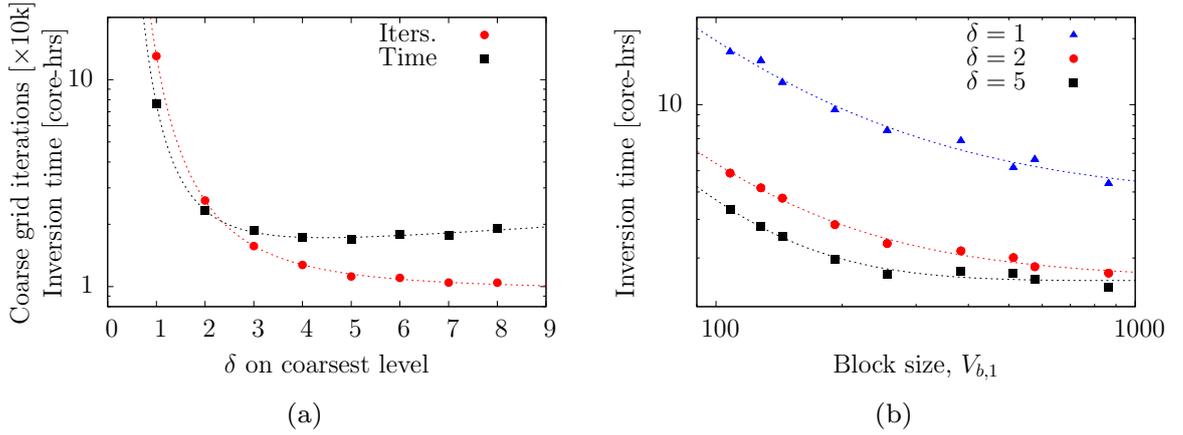
Figure 4.7: The number of coarse grid iterations as a function of the δ parameter and the inversion time for different δ-values depending on the block size. The results depicted in (a) are for a block size of $V_{b,1} = 4^4$ and compare the number of iterations on the coarse grid with the fine grid inversion time. The number of iterations has been scaled down by a factor 10000. The behavior for the coarse grid iterations is proportional to $1/\delta^2$, while the inversion time has an additional term linear in δ due to the increase in the fine grid iteration count. The results depicted in (b) show the inversion time for different values of δ and different block sizes. The behavior is similar to the one observed in Figure 4.5(d).

up the inversion time on the coarsest grid. We also analyze the effect of the scaling factor δ as a function of the block size. The results are depicted in Fig. 4.7 and we find that the optimal aggregation parameters determined in the previous section do not depend on the chosen coarse grid scale factor δ. In Tab. 4.5 we present the final results for the $cA2.48$ ensemble, summarizing the δ parameters which minimize the inversion time for various block sizes and the corresponding optimal numbers of test vectors. By using a scaling factor $\delta = 2$ the iteration count on the coarsest grid is already reduced by a factor of five, which results into a speed-up of the fine grid inversion time by roughly a factor of three. Optimal performance is achieved for a relatively large plateau around $\delta \sim 5$. Obviously, a large scaling factor δ causes a distinct violation of the Galerkin condition, i.e. $P^\dagger D(\mu)P \neq D_c(\delta\mu)$. However, we only find a minor increase in the iteration count of the fine grid solver from 14 to 16 iterations.

4.3.3   Stability of optimal parameters

Empirically we find that the solver performance is stable for the tuned parameter set. Within the ensemble $cA2.48$ we do not find any configurations where the iteration counts or the time to solution differ by more than 5%, as can be seen in Fig 4.8. This behavior is also corroborated by the performance of the multigrid solver during the force computation in the HMC algorithm, where it shows very stable iteration counts for simulations at the physical point, see Fig. 4.9.

4.4   Comparison with inexact deflation and MG-GCR

In this section we consider an implementation of the MG-GCR algorithm provided by the QUDA library [117] and the implementation of inexact deflation provided in openQCD [1] [118]. The latter is restricted to a two-level approach. Differences between the three multigrid approaches are summarized in in Tab. 4.6.

---

[1] We use for the tests a modified openQCD-version, which is optimized for twisted mass fermions, i.e. where the even–odd reduced twisted mass Dirac operator is implemented.
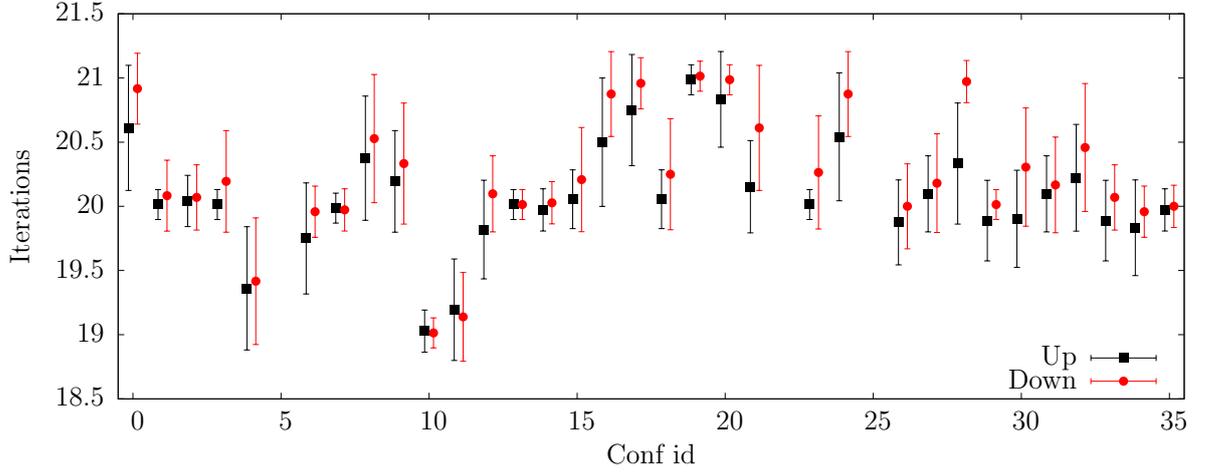
Figure 4.8: We depict the average iteration counts on computing quark propagators for several configuration of the ensemble $cA2.48$. For the black square points the setup is generated with the same μ while for the red circle points the setup is generated with a TM parameter μ with opposite sign.
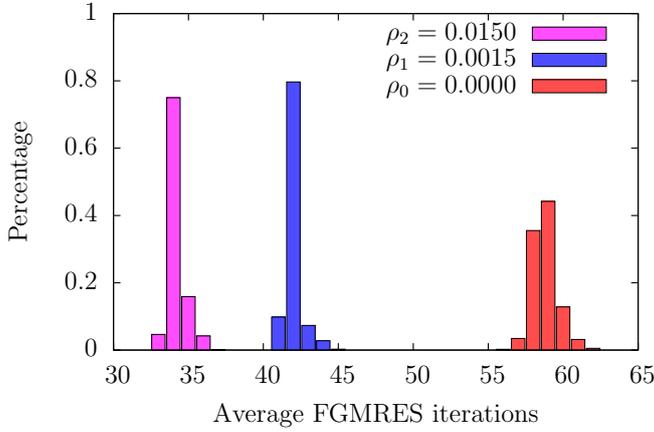
Figure 4.9: We show the FGMRES iteration counts averaged over the trajectory when the solver is used for computing the force terms in an HMC simulation. The data is for the ensemble $cA2.64$, see Tab. 6.1, with statistics of 2000 trajectories and the squared operator $\hat{D}^\dagger\hat{D} + \rho_i^2$ is inverted, i.e. two inversions are performed with DD-αAMG inverting $\hat{D} + i\rho_i\hat{\Gamma}_5$ and $\hat{D}^\dagger - i\rho_i\hat{\Gamma}_5$ where $\hat{\Gamma}_5$ is the restriction of Γ_5 to the odd lattice sites.

Component	Inexact deflation	MG-GCR	DD-αAMG
Multilevel AMG	No	Yes	Yes
Γ_5-compatibility	No	Yes	Yes
Setup	Test-vectors	Null-vector	Test-vectors
Smoother	SAP	MR	SAP
Odd-even reduction	Everywhere	Everywhere	Smoother and coarsest level
Flexible solver	GCR	GCR	FGMRES
Coarse-grid solver	GCR	GCR	GMRES

Table 4.6: Main differences between DD-αAMG and MG-GCR approaches.

4.4.1 The coarse-grid scale factor δ

We tested the influence of a delta parameter in both cases. For both approaches we find that the iteration counts on the coarsest grid are reduced by using a scale factor $\delta > 1$, also resulting in reduced inversion and setup times.

The MG-GCR implementation in QUDA allows to chose whether we want to use the even-odd reduction only on the coarsest-level or in all levels. The first is the same as in DD-αAMG, while the latter is the same as suggested for MG-GCR since one achieves the best performance. In Tab. 4.7 we show results at several shifts δ in three different situations. The first without even-odd reduction, the second employing even-odd reduction only on the coarsest level and the third employing the reduction

directly from the fine-level inverting the even-odd reduced operator. The first and second case have a similar dependence on the shift δ, finding the best performance at $\delta \sim 5$ as in DD-αAMG, cf. Tab. 4.5. Going from the first to the second case gives a speed-up of around 20%. In the third case, instead, one has to employ a factor δ at least four times larger than before. Going from the second to the third case gives an additional speed-up of around 20%.

No even-odd			Only on the coarsest level			On all the levels		
δ_μ	Inv. iters	Inv. time	δ_μ	Inv. iters	Inv. time	δ_μ	Inv. iters	Inv. time
2	26	47.1 s	2	26	31.7 s	5	29	66.8 s
3	28	12.8 s	3	29	9.9 s	10	30	9.7 s
4	30	10.8 s	4	30	8.5 s	15	32	7.0 s
5	31	10.7 s	5	31	8.2 s	20	33	6.7 s
6	31	10.5 s	6	31	8.1 s	25	36	7.1 s
7	32	11.2 s	7	32	8.4 s	30	37	7.4 s
8	33	11.4 s	8	33	8.6 s	35	37	6.7 s

Table 4.7: Analysis of the shift δ dependently on the levels where the even-odd reduction is employed. The results have been obtained with the QUDA package for the ensemble $cA2.48$ running on 54 Nvidia Kepler GPUs. In the table on the left no even-odd reduction has been use. In the middle the even-odd reduction has been employed only on the coarsest level. On the right the even-odd reduction has been employed on all the three levels of the multigrid inverting directly the even-odd reduce operator.

For inexact deflation in openQCD we observe similar result as the latter case, since also there the even-odd reduced operator is inverted. Furthermore, when solving for $D(-\mu)$ with a setup built for $D(\mu)$, we find that the optimal δ increases by a factor of four within the openQCD solver whereas it remains constant for DD-αAMG. This is a consequence of the Γ_5-compatibility which is not preserved by the prolongation operator in inexact deflation. For the Wilson fermion case it was reported in Ref. [105] that the two-level openQCD solver shows about the same performance as the three-level DD-αAMG approach. For the twisted mass ensemble $cA2.48$, the openQCD solver is roughly a factor of four slower than DD-αAMG.

4.4.2 Null-vector setup

We analyze the performance of the null-vector setup in MG-GCR, cf. Sec. 3.4.2. For its successful application we had to implement the possibility of changing solver in the setup procedure of QUDA. The default solver in QUDA is BiCGStab which has a very bad convergence for twisted mass fermions as shown in the top-left of Fig. 4.10. GCR shows a better convergence which although stops at a relative residual of 10^{-4}. As expected the optimal convergence is given by CG where although the squared operator is used in the definition of null-vectors, cf. Eq. 3.57, i.e. $D^\dagger D u = 0$. Therefore the CG solver has to be used in the setup procedure. We also show the convergence of the solver at several values of the twisted mass term μ in the top-right of Fig. 4.10. Interestingly the convergence is the same at different μ down to a certain relative residual, then the solver starts to converge faster. Also the convergence in the coarse grid is stable as shown in the bottom left of Fig. 4.10. Finally employing setups at different tolerance in the inversion we find out that the best time to solution is achieved with a setup tolerance of about 10^{-7} relative residual.
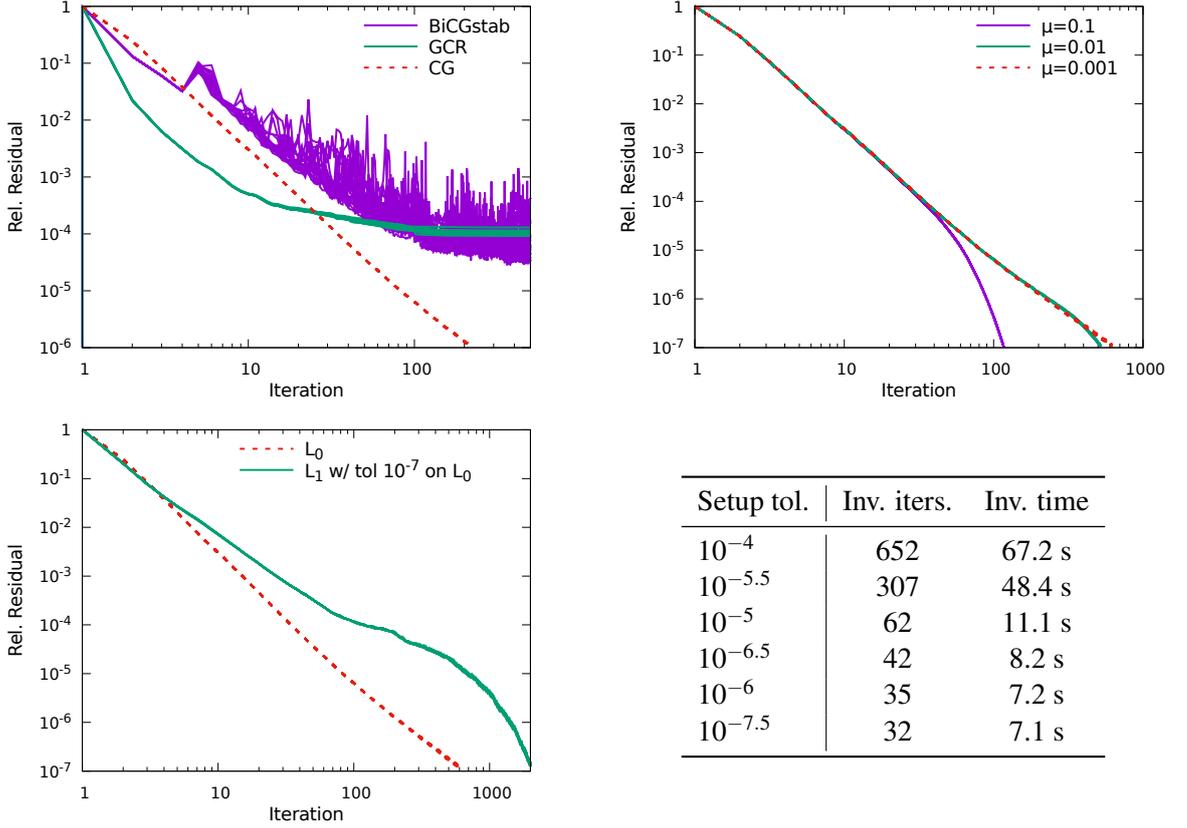
Figure 4.10: Null-vector setup procedure in MG-GCR for twisted mass fermions using the ensemble *cA*2.48. Top-left, comparison between the convergence of different solvers in the calculation of the null vectors. Top-right, convergence of CG at different twisted mass parameter μ (target $\mu = 0.0009$). Bottom-left, convergence of the CG solver on different multigrid levels with L_0 being the fine-level and L_1 the first coarse level. Bottom-right, table with the iterations and time to solutions for inversions with MG-GCR using setup procedure at different relative residual.

4.5 Error and residual of multigrid solutions

In this section we compare the numerical error and residual of solutions which are equivalent in exact arithmetic but not numerically. For clarity sake, our notation is the following: a linear system

$$Ax = b, \tag{4.22}$$

with *A* being an invertible matrix, *b* a known right hand side (rhs) and *x* the numerical solution, has residual and error, respectively,

$$r = b - Ax \quad \text{and} \quad e = A^{-1}b - x = A^{-1}r. \tag{4.23}$$

The residual is commonly used as stopping criterium for the solvers, since it does not require the knowledge of the *exact solution* $A^{-1}b$. However the error vector *e* is the statistical deviation which one introduces using the solution *x*. The error is connected to the residual by the relation

$$e = A^{-1}r \quad \text{which states} \quad \|e\| = \|A^{-1}r\| \leq \|A^{-1}\|\,\|r\| = \frac{\|r\|}{\sqrt{\lambda_{\min}(A^{\dagger}A)}}, \tag{4.24}$$

where the latter equality holds in the Euclidean norm and $\lambda_{min}(A^\dagger A)$ is the smallest eigenvalue of $A^\dagger A$. Iterative solvers are usually stopped when $\|r\| < \rho \|b\|$ with ρ being a fixed tolerance. This fix the relative norm of the residual to be below a given threshold.

Hereafter we study and show numerical results for equivalent solutions in several cases. We use the twisted mass operator on the physical test ensemble. In the numerical results we also show the error with maximal and minimal deviation of the local components. In the following examples we start from a random vector taken as the solution x. Then we apply the involved operator A to obtain the right hand side $b = Ax$. Then the error is given by $e = x' - x$ with x' the iterated solution. The minimal and maximal local deviation are given min/max of the norm of e restricted to the lattice site. The maximal and minimal deviation are potentially interesting for future simulations where very large lattices will be used [119]. Indeed since $\|b\|$ grows with $V^{1/2}$, the stopping criterium $\|r\| \leq \rho \|b\|$ does not guarantee the residual or error to be uniformly small and at larger V larger local deviation of the vectors are allowed.

4.5.1   Inverting the twisted mass operator
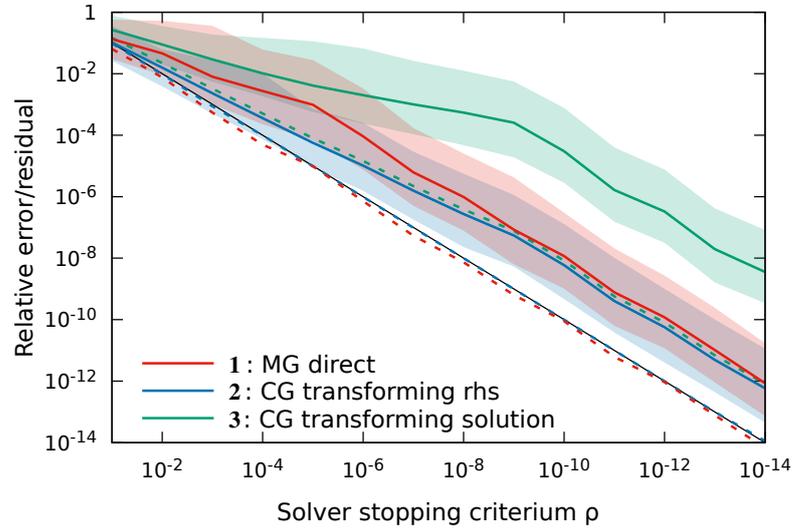
Figure 4.11: The dependence of the relative error and residual on the solver stopping criterium, ρ for the three different approaches for solving the linear equation involving one single TM Wilson operator. Method **1** is given by the direct solution using the multigrid solver (red), method **2** is given by the CGNR method (blue) and method **3** is given by the CGNE method (green). The relative residual is shown as dashed line while the error is illustrated by the solid line with a shaded band determined by the largest and smallest local deviations as defined in Sec. 4.5. The black line shows where y-axis is equal to the x-axis (i.e. relative residual equals to ρ).

In general retrieving the solution of the linear system in Eq. (4.22) can be done by different methods, for instance by using a linear iterative solver like the conjugate gradient (CG) solver. In case of CG, which requires A to be hermitian instead of Eq. (4.22) a modified linear equation involving a hermitian operator has to be solved. Two methods are available, known as CGNR and CGNE given by solving the equations

$$(A^\dagger A)x = A^\dagger b \quad \text{or} \quad (AA^\dagger)y = b \rightarrow x = A^\dagger y \tag{4.25}$$

for obtaining x respectively. Although both solutions are equivalent in exact arithmetic, in practice the numerical error and residual, as defined in Eq. (4.23), are in general different. In order to approach

smaller quark masses and increasing the dimension of the lattice require to study different ways in order to control the resulting error.

Here we consider the following three methods for solving the linear equation $(Q \pm i\mu I)x_{\pm} = b$:

1: The linear system is solved directly by obtaining

$$x_{\pm,1} + e_{\pm,1} = (Q \pm i\mu I)^{-1}b. \tag{4.26}$$

The solver stopping criterium is based on the relative residual of the solution, thus

$$\|r_{\pm,1}\| = \|b - (Q \pm i\mu I)x_{\pm,1}\| \leq \rho \|b\|. \tag{4.27}$$

Following Eq. (4.24), the norm of the error satisfies

$$\|e_{\pm,1}\| = \left\| (Q \pm i\mu I)^{-1}r_{\pm,1} \right\| \leq \frac{\|r_{\pm,1}\|}{\sqrt{\lambda_{\min}^2 + \mu^2}} \leq \frac{\rho \|b\|}{\sqrt{\lambda_{\min}^2 + \mu^2}} \tag{4.28}$$

where $\lambda_{\min}^2$ is the smallest eigenvalue of Q^2.

2: A normal equation is used by applying a transformation on the rhs (equivalent to CGNR)

$$x_{\pm,2} + e_{\pm,2} = (Q^2 + \mu^2 I)^{-1}(Q \mp i\mu I)b. \tag{4.29}$$

Since the rhs is $(Q \mp i\mu I)b$, the solver stopping criterium is

$$\left\| r'_{\pm,2} \right\| = \left\| (Q \mp i\mu I)b - (Q^2 + \mu^2 I)x_{\pm,2} \right\| \leq \rho \left\| (Q \mp i\mu I)b \right\| \leq \rho \|b\| \sqrt{\lambda_{\max}^2 + \mu^2}. \tag{4.30}$$

The norm of the residual is then

$$\|r_{\pm,2}\| = \left\| (Q \mp i\mu I)^{-1}r'_{\pm,2} \right\| \leq \frac{\rho \left\| (Q \mp i\mu I)b \right\|}{\sqrt{\lambda_{\min}^2 + \mu^2}} \leq \rho \|b\| \sqrt{\frac{\lambda_{\max}^2 + \mu^2}{\lambda_{\min}^2 + \mu^2}} \tag{4.31}$$

and the norm of the error satisfies

$$\|e_{\pm,2}\| \leq \frac{\|r_{\pm,2}\|}{\sqrt{\lambda_{\min}^2 + \mu^2}} \leq \frac{\rho \|b\| \sqrt{\lambda_{\max}^2 + \mu^2}}{\lambda_{\min}^2 + \mu^2}. \tag{4.32}$$

3: A normal equation is used by applying a transformation on the solution (equivalent to CGNE)

$$y_3 + e_3 = (Q^2 + \mu^2 I)^{-1}b \quad \longrightarrow \quad x_{\pm,3} + e_{\pm,3} = (Q \mp i\mu I)y_3 \tag{4.33}$$

where $e_{\pm,3} = (Q \mp i\mu I)e_3$. The solver stopping criterium

$$\|r_3\| = \left\| b - (Q^2 + \mu^2 I)y_3 \right\| = \|b - (Q \pm \mu I)x_{\pm,3}\| = \|r_{\pm,3}\| \leq \rho \|b\| \tag{4.34}$$

is equivalent to computing the residual of the solution $r_{\pm,3}$. The norm of the error then satisfies

$$\|e_{\pm,3}\| \leq \frac{\|r_{\pm,3}\|}{\sqrt{\lambda_{\min}^2 + \mu^2}} \leq \frac{\rho \|b\|}{\sqrt{\lambda_{\min}^2 + \mu^2}}. \tag{4.35}$$

From this analysis, we conclude that method **1** and **3** generates solutions which have compatible residuals and errors. On the other hand, method **2** has upper limits increased by the condition number $K = \sqrt{(\lambda_{\max}^2 + \mu^2)/(\lambda_{\min}^2 + \mu^2)}$ compared to **1** or **3**. The numerical results depicted in Fig. 4.11 for the degenerate twisted mass operator at the physical light quark mass ($\lambda_{\min}^2 = 0$ and $\mu = 0.00072$) confirm these conclusions. From the numerical results we also notice that the error of the method **2** is close to the upper limits obtained in Eq. (4.32). This shows that the residual of CG solver is dominated by the lowest eigenmodes. Indeed if $r = v_{\min}(A^\dagger A)$ holds, where $v_{\min}$ is the eigenvector of the smallest eigenvalue, then it follows

$$\|e\| = \|A^{-1}r\| = \|A^{-1}\| \, \|r\| = \frac{\|r\|}{\sqrt{\lambda_{\min}(A^\dagger A)}} \, . \tag{4.36}$$

4.5.2 Solving the squared linear equation

Here we take a closer look to the linear system

$$(Q^2 + \mu^2 I)x = b \tag{4.37}$$

which involves a squared hermitian operator $Q^2 = D^\dagger D$. We consider here the following three methods:

A: the linear system is directly solved by

$$x_A + e_A = (Q^2 + \mu^2 I)^{-1}b. \tag{4.38}$$

The solver stopping criterium is based on the relative residual of the solution, thus

$$\|r_A\| = \|b - (Q^2 + \mu^2 I)x_A\| \le \rho \, \|b\| \tag{4.39}$$

and norm of the error satisfies

$$\|e_A\| = \|(Q^2 + \mu^2 I)^{-1}r_A\| \le \frac{\|r_A\|}{\lambda_{\min}^2 + \mu^2} \le \frac{\rho \, \|b\|}{\lambda_{\min}^2 + \mu^2} \, . \tag{4.40}$$

B: the system is solved in two consecutive steps, by computing

$$x_\pm + e_\pm = (Q \pm i\mu I)^{-1}b \quad \text{and then} \quad x_B + e_B' = (Q \mp i\mu I)^{-1}x_\pm \tag{4.41}$$

using either x_+ or x_-. The solution

$$x_B + e_B = (Q^2 + \mu^2 I)^{-1}b \quad \text{has error} \quad e_B = e_B' + (Q \mp i\mu I)^{-1}e_\pm. \tag{4.42}$$

The solver stopping criteria are in the two steps respectively

$$\|r_\pm\| = \|b - (Q \pm i\mu I)x_\pm\| \le \rho \, \|b\| \tag{4.43}$$

$$\|r_B'\| = \|x_\pm - (Q \mp i\mu I)x_B\| \le \rho \, \|x_\pm\| \le \frac{\rho \, \|b\|}{\sqrt{\lambda_{\min}^2 + \mu^2}}. \tag{4.44}$$

The residual of the solution is then

$$r_B = b - (Q^2 + \mu^2 I)x_B = b - (Q \pm i\mu I)(x_\pm + r'_B) = r_\pm + (Q \pm i\mu I)r'_B$$

$$\implies \|r_B\| < \|r_\pm\| + \|Q \pm i\mu I\| \, \|r'_B\| \leq \left(1 + \sqrt{\frac{\lambda^2_{\max} + \mu^2}{\lambda^2_{\min} + \mu^2}}\right)\rho \, \|b\|. \qquad (4.45)$$

For obtaining a smaller upper limit of the error in Eq. (4.42), we consider

$$e_\pm = (Q \pm i\mu I)^{-1} r_{\pm,B} \implies \|e_\pm\| \leq \frac{\rho \, \|b\|}{\sqrt{\lambda^2_{\min} + \mu^2}} \qquad (4.46)$$

$$e'_B = (Q \mp i\mu I)^{-1} r'_B \implies \|e'_B\| \leq \frac{\rho \, \|x_\pm\|}{\sqrt{\lambda^2_{\min} + \mu^2}} \leq \frac{\rho \, \|b\|}{\lambda^2_{\min} + \mu^2} \qquad (4.47)$$

from which we obtain

$$\|e_B\| = \left\|e'_B + (Q \mp i\mu I)^{-1} e_\pm\right\| \leq \|e'_B\| + \frac{\|e_\pm\|}{\sqrt{\lambda^2_{\min} + \mu^2}} \leq \frac{2\rho \, \|b\|}{\lambda^2_{\min} + \mu^2}. \qquad (4.48)$$

C: the solution is given by a difference of two solutions of a linear combination of $x_\pm + e_\pm = (Q \pm i\mu)^{-1}b$,

$$x_C = b - (Q^2 + \mu^2)x_C = \frac{i}{2\mu}(x_+ - x_-) = (Q^2 + \mu^2 I)^{-1}b - \frac{i}{2\mu}(e_+ - e_-) \qquad \text{for } \mu \neq 0. \quad (4.49)$$

The solver stopping criteria are in Eq. (4.43) and the errors $e_\pm$ in Eq. (4.46). The residual of the solution is then

$$r_C = \frac{i}{2\mu}\left((Q + i\mu I)r_+ - (Q - i\mu I)r_-\right)$$

$$\implies \|r_C\| \leq \frac{1}{2\,|\mu|}\left(\|Q + i\mu I\| \, \|r_-\| + \|Q - i\mu I\| \, \|r_+\|\right) \leq \frac{\rho\sqrt{\lambda^2_{\max} + \mu^2} \, \|b\|}{|\mu|} \qquad (4.50)$$

while the error of the solution is

$$e_C = \frac{i}{2\mu}(e_+ - e_-) \implies \|e_C\| \leq \frac{1}{2\,|\mu|}\left(\|e_+\| + \|e_-\|\right) \leq \frac{\rho \, \|b\|}{|\mu|\sqrt{\lambda^2_{\min} + \mu^2}}. \qquad (4.51)$$

In exact arithmetic all ways yield to the same solution. This is not so obvious anymore if iterative solvers are used since they involve the stopping criterium ρ and the different approaches have different upper limits as just derived. In case of the CG solver, only method **A** can be employed. On the other hand, in case of MG solver only methods **B** and **C** are available accordingly to our software implementation, i.e. only the un-squared twisted mass operator can be solved. Analytically comparing the bounds of the residual of the iterated solution of Eq. (4.41) and Eq. (4.49) one would be in favor of case **B** for $\mu < 1$. Indeed following the above discussion for the residuals we obtain

$$\frac{\|r_B\|}{\|b\|} \leq \left(1 + \sqrt{\frac{\lambda^2_{\max} + \mu^2}{\lambda^2_{\min} + \mu^2}}\right)\rho \qquad \text{and} \qquad \frac{\|r_C\|}{\|b\|} \leq \frac{\sqrt{\lambda^2_{\max} + \mu^2}}{|\mu|}\rho \,, \qquad (4.52)$$

indicating that $\|r_C\|$ can be much larger than $\|r_B\|$ at $\mu \ll 1$. However, comparing the error

$$\frac{\|e_B\|}{\|b\|} \leq \frac{\rho}{\lambda_{\min}^2 + \mu^2} \qquad \text{and} \qquad \frac{\|e_C\|}{\|b\|} \leq \frac{\rho}{|\mu| \sqrt{\lambda_{\min}^2 + \mu^2}}. \qquad (4.53)$$

it follows the error e_B of case **B** is numerical equivalent to the error e_C of case **C** if $\lambda_{\min} \ll \mu$.

In the case of twisted mass fermions at maximal twist this relation is fulfilled since $\lambda_{\min} \sim 0$. This is the case in our example, shown in Fig. 4.12 for the degenerate operator at light quark mass. The dependence of the relative error and residual on the solver stopping criterium ρ for solutions obtained from the CG solver and the two aforementioned multigrid approaches is analyzed on our physical test ensemble. For this ensemble, $\lambda_{min}^2 + \mu^2 \simeq 0.00072^2$, thus yielding a difference around six orders of magnitude between the stopping criterium ρ and error, as shown in Fig. 4.12 for all the three methods after $\rho > 10^{-9}$.
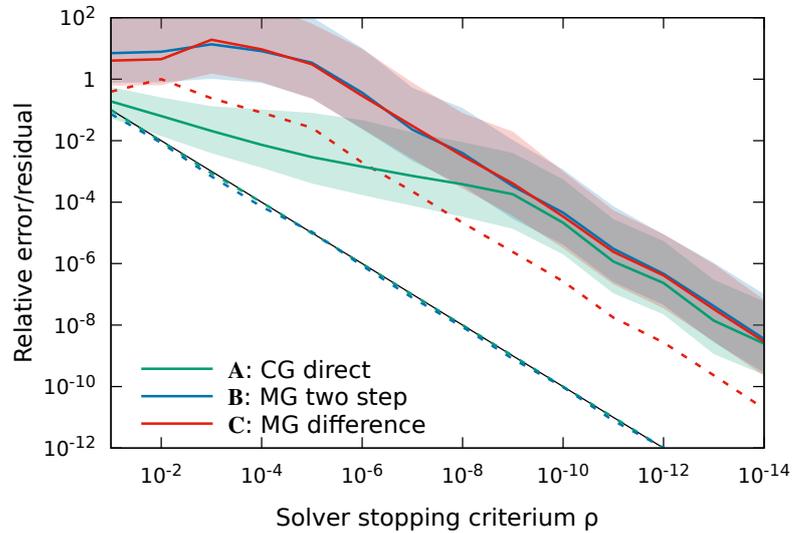
Figure 4.12: The dependence of the relative error and residual on the solver stopping criterium, ρ, for the three different ways of solving Eq. (4.37). Method **A** uses the conjugated gradient solver (green), method **B** uses the consecutively ordered single inversions of Eq. (4.41) (blue) and method **C** uses the single inversions involving differences (red) given by Eq. (4.49). The relative residual is shown as dashed line while the relative error is illustrated by the solid line with a shaded band determined by the largest and smallest local deviations as defined in 4.5. The black line shows where y-axis is equal to the x-axis (i.e. relative residual equals to ρ).

Despite the fact that as long as $\mu \gg \lambda_{min}$ all approaches yield a similar error, the approach **C** has advantages. The first advantage is about software optimization where one can solve both shifts together, as outlined in Ref. [120, 121]. The second advantage is that for the shifted linear equations we can give a common initial guess where available since we are solving twice the same linear system (except for the sign of μ) on the same rhs, cf. Eq. (4.49).

The comparison between the numerical error and the analytical upper bound can give some information on the modes the solver tackles effectively. Namely if the numerical error coincidences with the upper bound, the error is dominated by the mode of the smallest eigenvalue. In contrast, if the error is much smaller than the upper bound, the error is dominated by much larger modes and thus the solver would treat the small eigenmodes very effectively. In our numerical test we found that in all cases the smallest eigenvalues are dominating the error. Interestingly, this is also the case if a multigrid solver is used, which is in contrast to the observation in the case of Wilson fermions [102]. It is surprising because multigrid methods threat the low mode subspace separately and the conver-

gence is expected to be similar for all the modes, cf. Fig. 3.8. For instance an exactly deflated solver would have the residual dominated by the first non-deflated eigenvalue. Therefore the analytical upper bound would be reduced accordingly. This has been observed in Ref. [102] for Wilson fermion but here we do not observe it for twisted mass fermions. Our explanation is that the difference is related to a finite twisted mass term which has a higher density of small eigenvalues at the cut-off scale μ, as shown in Fig. 2.4. The higher density increases the condition number of the coarse operator which makes its inversion harder. For accelerating the time to solution of DD-αAMG we have employed a shift in the coarsest grid, cf. Sec. 4.3.2. Most probably the latter spoils the effective deflation of the twisted mass operator and some low modes survive in the error. To corroborate this observation we also notice that the norm of the error has a slope parallel to the target relative residual at a distance $\lambda_{\min}^{-1}$, only when the residual vector is dominated by the eigenvector with the smallest eigenvalue. Indeed $\|e\| = \left\|(Q^2 + \mu^2 I)^{-1} r\right\| = \lambda_{\min}^{-1} \|r\|$ only if r is such eigenvector. This is indeed the separation we observe in Fig. 4.12 between error and relative target residual.
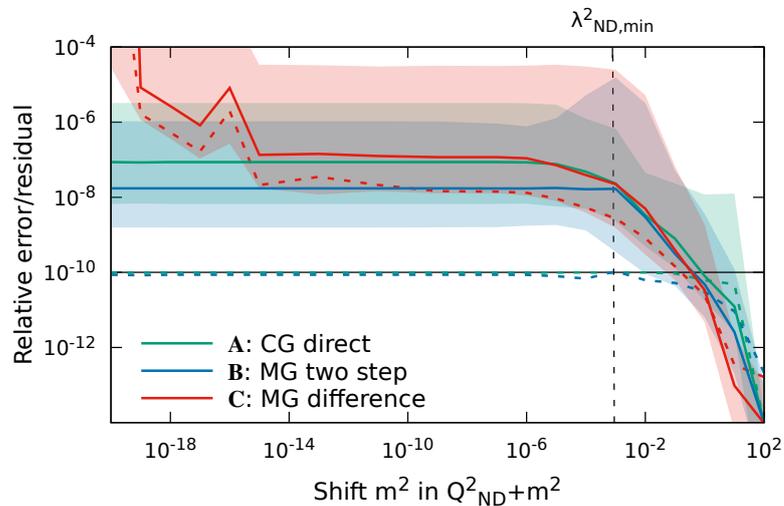
Figure 4.13: The dependence of the relative error and residual on the shift m^2 for the three different approaches for solving the linear equation involving the squared even-odd reduced ND twisted mass operator. Method **A** is given by the direct solution using the conjugate gradient solver (green), method **B** is given by the consecutively ordered single solves (blue) and method **C** is given by the difference of the two single solves (red). The relative residual is shown as dashed line while the relative error is illustrated by the solid line with a shaded band determined by the largest and smallest local deviations as defined in 4.5. The black solid line shows where the relative residual is equal to the solver stopping criterium $\rho = 10^{-10}$. The black dashed line shows where m^2 is equal to the smallest eigenvalue of Q_{ND}^2, $\lambda_{\mathrm{ND},min}^2 \sim 0.0008$

In Fig. 4.13 we analyze the connection of the error with the residual for the squared non-degenerate twisted mass operator at different shifts m^2 where the operator has been tuned at physical strange and charm quark mass. The solved equation is the following,

$$\left(D_{\mathrm{ND}}^\dagger D_{\mathrm{ND}} + m^2\right) x = b. \tag{4.54}$$

We fix the stopping criteria for each inversions to $\|r\| / \|b\| < \rho = 10^{-10}$. As discussed above, the relative error and residual depend on the smallest eigenvalue of the operator $Q_{\mathrm{ND}}^2 + m^2$, which is given by $\lambda_{\mathrm{ND},min}^2 \sim 0.0008$. As follows from Eqs. 4.40 and 4.48, the inverse of the smallest eigenvalue bounds the relative error for the approach **A** and **B** to be by around 3 orders of magnitudes larger compared to the stopping criterium ρ if $m^2 < \lambda_{\mathrm{ND},min}^2$. This is shown in Fig. 4.13. For the case **C**

this is not true since from Eq. 4.51 follows that the relative error is bounded by $\rho/|m\lambda_{ND,\min}|$ which should diverge at $m^2 \to 0$. However, numerically we find that the real relative error is similar to case **A** and **B** when $m^2 > 10^{-14}$. This shows that method **C** can be used for our application since in the rational approximation of the matrix square root all shifts are larger than $m^2 > 10^{-10}$. At $m^2 < 10^{-14}$ the error of method **C** diverges. This is probably due to a significance loss of the floating point precision. In the test we have used the well-known 64 bytes double precision which has significance loss at a relative value smaller than 10^{-14}. To corroborate this statement it would be interesting to perform the same test as in Fig. 4.13, for single precision which we expect to diverge at $m^2 < 10^{-7}$.

5. Monte Carlo algorithms for Lattice QCD

Monte Carlo techniques are used to evaluate the Feynman path integrals needed for the extraction of observables. In this chapter, we introduce Monte Carlo techniques starting from pure gauge field theory and then discuss how fermionic degrees of freedom are taken into account. We explain the Hybrid Monte Carlo (HMC) algorithm, which is a combination of Monte Carlo and Molecular Dynamics (MD) approaches. In the simulation of the non-degenerate twisted mass operator one has to compute the square-root of the inverse of the operator. This is done by employing a rational approximation of the square-root. We also present our approach for accelerating the inversions involved in the rational approximation and show the speed-up we achieve by employing multigrid methods in the simulations.

5.1   Monte Carlo simulations

The goal is to have an efficient method to compute expectation values of gauge invariant operators that can be related to physical observables. Such an observable, as follows from Sec. 2.1, can be computed via the integral

$$\langle O(\bar{\psi}, \psi, U) \rangle = \frac{1}{Z} \int \left[\prod_{\mu, x_i}^{4V} dU_\mu(x_i) \right] O\left(D^{-1}(U, m_f), U \right) \prod_f \det\left(D(U + m_f) \right) e^{-S(U)}, \qquad (5.1)$$

which can be evaluated using Monte Carlo methods. The idea stems from similar approaches used in statistical mechanics where a representative set of states of the systems is generated by an ergodic process. Utilizing these representative set of states, also referred to as configurations, one can estimate the integral over the fundamental degrees of freedom of the system.

Markov chain Monte Carlo (MC) approaches were first studied under a secret project, coined Monte Carlo, at the Los Alamos National Laboratory during second world war and that gave this name to these methods. The Metropolis algorithm [122] revolutionized the field, since it provided a rigorous algorithm to generate representative configurations and evaluate integrals of the form of Eq. (5.1) for a general action S. The basic steps of the Metropolis algorithm is as follows:

1. Start from a random gauge field configuration $C^{(0)}$, i.e. a set of values for all the gauge variables $U_\mu(x)$.

2. Change one link, $U_\mu(x) \to U'_\mu(x)$, using a procedure that has a symmetric probability, i.e.

$$P(U \to U') = P(U' \to U). \qquad (5.2)$$

3. Accept $U'_\mu(x)$ if the action is decreased or accept with probability $e^{-\Delta S}$ if the action is increased where $\Delta S = S(U') - S(U)$ is the change in the action. Otherwise reject, i.e $U_\mu(x)$ remains unchanged.

4. Repeat steps 2-3 for all $U_\mu(x)$. When all $U_\mu(x)$ are updated then a new configuration $C^{(1)}$ is generated.

5. Repeat step 4 for N times.

Steps 2 and 3 of the Metropolis algorithm generate configurations C with a probability distribution proportional to $e^{-S(C)}$. For example, let $U \to U'$ be a change that increases the action, then the new element U' is accepted with a transition probability

$$T(U \to U') = P(U \to U')e^{-\Delta S} = P(U' \to U)e^{-\Delta S}.$$ (5.3)

On the other hand, the change $U' \to U$ would decrease the action and thus it is always accepted having

$$T(U' \to U) = P(U' \to U).$$ (5.4)

Therefore, from the property of the detailed balance,

$$P(U)T(U \to U') = P(U')T(U' \to U),$$ (5.5)

which is satisfied by the Markov chain Monte Carlo methods, we obtain

$$\frac{P(U')}{P(U)} = \frac{T(U \to U')}{T(U' \to U)} = \frac{P(U' \to U)e^{-\Delta S}}{P(U' \to U)} = \frac{e^{-S(U')}}{e^{-S(U)}}.$$ (5.6)

The same can be proved if the change $U \to U'$ decreases the action. Eq. (5.6) is verified if each link U is distributed accordingly to a probability $P(U) = e^{-S(U)}$.

5.1.1 Monte Carlo expectation values

The expectation value in Eq. (5.1) is evaluated by using an *ensemble average* over the N sampled configurations. Since the Metropolis algorithm starts from a random set of elements, $C^{(0)}$, the number of iterations required to approach the correct probability distribution is known as thermalization period. Assuming that after n_0 steps the probability distribution has approached the correct limit we can use the subsequent configurations to compute the expectation value of our operator given as

$$\langle O \rangle = \frac{1}{N - n_0} \sum_{k=n_0}^{N} O(C^{(k)})$$ (5.7)

where $O(C)$ stands for the value of O computed on the configuration $C^{(k)}$. This average is computed on a finite ensemble of configurations and it is necessary to estimate the error. In general, the error of an observable O is given by

$$\Delta O = \sigma(O)\sqrt{\frac{2\tau_{int}}{N_{cnf}}} \propto \frac{1}{\sqrt{N_{cnf}}} \quad \text{where} \quad \sigma(O) = \sqrt{\langle O^2 \rangle - \langle O \rangle^2},$$ (5.8)

where $N_{cnf} = N - n_0$ is the number of thermalized configurations and $\tau_{int} \geq 0.5$ is the integrated autocorrelation time, which takes into account that the configurations are correlated with each other. In LQCD, the autocorrelation time increases with decreasing lattice spacing a and its value depends on the specific Monte Carlo algorithm. This can spoil simulations for very fine lattices [123].

5.1.2 Pure-gauge simulations

In case of pure-gauge theories or quenched LQCD, i.e. where only the gauge action is considered in the MC sampling, we have

$$P(U) \propto e^{-S_g(U)} \propto e^{\beta \sum_{x,\mu < \nu} \text{Re Tr} \, U_{\mu\nu}(x)}. \tag{5.9}$$

An important element is that the gauge action is local, which means that each time a link is updated only next neighboring terms are affected making the computation of the difference in the action very simple. If $U_\mu(x_i) \to U'_\mu(x_i)$, ΔS is

$$\Delta S = \beta \sum_{\nu \neq \mu} \text{Re Tr} \left(U'_{\mu\nu}(x_i) - U_{\mu\nu}(x_i) \right), \tag{5.10}$$

which does not have a sum over the volume and consists of a very small number of operations. This is crucial when one has to repeat such a calculation an enormous number of times. Moreover the variation ΔS is kept small and one avoids drastic changes of configuration which, implying a very large positive ΔS, would almost invariably lead to a rejection of the move.

An improved approach for simulations of $SU(2)$ pure-gauge theories is the heat bath algorithm developed by M. Creutz in Ref [31]. Here the weight of the distribution can be solved exactly generating directly link distributed as $e^{-S_g(U)}$. This approach has been extended to any $SU(N)$ theory in Ref. [124] where the heat bath algorithm is applied to $SU(2)$ subgroups of $SU(N)$.

5.2 Pseudo-fermions integral

When we include the fermionic interactions of QCD in the action we obtain

$$\int \mathscr{U} \mathscr{D}\bar{\psi} \mathscr{D}\psi \, e^{-S_f(U,\bar{\psi},\psi)} = \int \mathscr{U} \int \mathscr{D}\bar{\psi} \mathscr{D}\psi \, e^{-\sum_{i=1}^{n_f} \bar{\psi} D(U,m_i)\psi} = \int \mathscr{U} \prod_{i=1}^{n_f} \det D(U,m_i), \tag{5.11}$$

where m_i is the bare quark mass. The fermionic determinant makes the Boltzmann weight non-local. This explains why it took more than 20 years to develop algorithms and have large enough machines to perform unquenched computations.

We now show how to rewrite the determinant in a different form that helps us in the simulations. For a generic matrix $M \in \mathbb{C}^{n \times n}$ and considering a complex vector $\eta = \{x_i + iy_i\} \in \mathbb{C}^n$, we have the identity

$$\frac{\pi^n}{\det M} = \int d^n x_i \, d^n y_i \, e^{-\eta^\dagger M \eta} \quad \text{if} \quad \text{Re}(\eta^\dagger M \eta) > 0 \quad \forall \eta \in \mathbb{C}^n \backslash \{0\}. \tag{5.12}$$

In the case of Grassmann variables we have instead

$$\det D = \int d^n \bar{\psi}_i \, d^n \psi_i \, e^{-\bar{\psi} D \psi} = \frac{1}{\pi^n} \int d^n x_i \, d^n y_i \, e^{-\eta^\dagger D^{-1} \eta}, \tag{5.13}$$

where ψ_i are Grassmann variables, while η_i are complex numbers that can be represented on a computer. The η variables are referred to as pseudo-fermions [125, 126], since they are used to represent the $detD$ but they are commuting variables. The condition $\text{Re}(\eta^\dagger M \eta) > 0 \quad \forall \eta \in \mathbb{C}^n \backslash \{0\}$ is equivalent to requiring

$$\text{Re}(\eta^\dagger M \eta) > 0 \quad \forall \eta \quad \Longrightarrow \quad \lambda_i(M + M^\dagger) > 0 \quad \forall i, \tag{5.14}$$

where $\lambda_i(M + M^\dagger)$ are the eigenvalues of the matrix $M + M^\dagger$. But this is not the most stringent

condition on the eigenvalues. If we want to evaluate the integral in Eq. (5.13) with MC sampling, i.e. using $e^{-\eta^\dagger D^{-1}\eta}$ as a propability, we need $\eta^\dagger D^{-1}\eta$ to be real and positive $\forall \eta$. Therefore, we need all the $\lambda_i(D^{-1})$ to be real and positive, which is consistent with the condition given in Eg. (5.14). This leads to the following formulations for dynamical fermion simulations

Wilson fermions, N_f=1: The Wilson operator D given in Eq. (2.66), has a real determinant thanks to the Γ_5-hermiticity, defined in Eq. (2.68), but the eigenvalues of $H = \Gamma_5 D$ are not positive-definite. In order to avoid the sign problem in the MC evolution one can rewrite as follows:

$$\det D = \det \sqrt{D^\dagger D} = \frac{1}{\pi^n} \int d^n x_i \, d^n y_i \, e^{-\eta^\dagger \left(D^\dagger D\right)^{-1/2}\eta} \qquad (5.15)$$

and

$$S_{N_f=1}^{W}(U,\eta) = \eta^\dagger \left[D^\dagger(U,m_f)D(U,m_f)\right]^{-1/2}\eta. \qquad (5.16)$$

Wilson fermions, N_f=2: In the case of a degenerate doublet, such as in the case of taking the same mass for the u- and d- quarks then one takes advantage of the Γ_5-hermiticity of the Wilson Dirac operator to write

$$\det D(m_{ud}) \det D(m_{ud}) = \det \left(\Gamma_5 D(m_{ud})\Gamma_5 D(m_{ud})\right) = \det \left(D^\dagger(m_{ud})D(m_{ud})\right), \qquad (5.17)$$

where $D^\dagger(m_{ud})D(m_{ud})$ is a positive-definite matrix by construction and invertible for $m_{ud} > \bar{m}$ where $\bar{m}$ is the critical mass of the Wilson operator. Therefore, the action for two degenerate quarks is

$$S_{N_f=2}^{W}(U,\eta) = \eta^\dagger \left[D^\dagger(U,m_f)D(U,m_f)\right]^{-1}\eta. \qquad (5.18)$$

Twisted mass fermions, N_f=2: In the case of the degenerate twisted mass operator given in Eq. (2.90) we have

$$\det \begin{bmatrix} D_{\text{TM}}(\mu) & 0 \\ 0 & D_{\text{TM}}(-\mu) \end{bmatrix} = \det \left(D_{\text{TM}}(-\mu)D_{\text{TM}}(\mu)\right) = \det \left(\Gamma_5 D_{\text{TM}}(-\mu)\Gamma_5 D_{\text{TM}}(\mu)\right)$$
$$= \det \left(D_{\text{TM}}^\dagger(\mu)D_{\text{TM}}(\mu)\right) = \det \left(D^\dagger D + \mu^2\right). \qquad (5.19)$$

The squared operator, cf. Eq. (2.107), is again positive definite by construction for $\mu^2 > 0$. The degenerate twisted mass action for fermions at maximal twist, i.e. $D \equiv D(U,\bar{m})$ with $\bar{m}$ being the critical mass of the Wilson operator, is

$$S_{N_f=2}^{\text{TM}}(U,\eta) = \eta^\dagger \left[D^\dagger D + \mu^2\right]^{-1}\eta. \qquad (5.20)$$

Twisted mass fermions, N_f=1+1: In the case of the non-degenerate twisted mass operator, given in Eq. (2.112) we have

$$\det D_{\text{ND}}(\bar{\mu},\bar{\varepsilon}) = \det \begin{bmatrix} D + i\bar{\mu}\Gamma_5 & -\bar{\varepsilon}I \\ -\bar{\varepsilon}I & D - i\bar{\mu}\Gamma_5 \end{bmatrix}, \qquad (5.21)$$

for which the determinant is real thanks to the $(\Gamma_5 \otimes \tau_1)$-hermiticity, cf. Eqs. (2.114) and (2.115). However, the positiveness of $H_{\text{ND}}(\bar{\mu},\bar{\varepsilon}) = (\Gamma_5 \otimes \tau_1) D_{\text{ND}}(\bar{\mu},\bar{\varepsilon})$ is not ensured and in order to avoid the sign problem we proceed with the same construction as for Eq. (5.15). The non-

degenerate twisted mass action is then given by

$$S_{N_f=1+1}^{\text{TM}}(U,\eta) = \eta^\dagger \left[D_{\text{ND}}^\dagger(\bar{\mu},\bar{\varepsilon}) D_{\text{ND}}(\bar{\mu},\bar{\varepsilon}) \right]^{-1/2} \eta. \tag{5.22}$$

In all the actions listed above we need to compute either the inverse of the Dirac operator or the square root of the inverse. These calculations are non-local and computationally very demanding with a cost proportional to the lattice size. The algorithms outlined in Sec. 5.1 for MC sampling are too expensive when *dynamical fermions* are used and improved algorithms are needed. Indeed, a Monte Carlo algorithm has either a very high rejection rate if significant changes are proposed or highly auto-correlated configurations if very small changes are applied. Standard MC approaches are effective only when the locality of the action can be exploited as in the case with pure-gauge simulations. Therefore, with dynamical fermions one needs an algorithm that computes the change in the action as less times as possible. A suitable algorithm is introduced in the next section.

5.3 Hybrid Monte Carlo (HMC) algorithm

The Hybrid Monte Carlo (HMC) algorithm [127, 128] combines MC and molecular dynamics (MD) approaches. In molecular dynamics a physical system is evolved in real time using the equations of motion. Therefore, compared to the Metropolis algorithm in Sec 5.1, the update of the configuration in HMC is not proposed randomly but from the evolution of a MD *trajectory*. For combining MD with MC, one needs

- to satisfy the detailed balance condition of Eq. (5.5),

- to ensure that the system can explore all the configuration space (ergodicity).

The first condition is achieved by using a symplectic time reversible integrator, i.e. one that can return to the initial point staring from the final point of the trajectory by reversing the momenta (time reversible) and preserve the Hamilton equation (symplectic). Such an integrator then satisfies

$$P(U \to U') = P(U' \to U), \tag{5.23}$$

where $U' \equiv U(t_f)$ is evolved from $U \equiv U(t_i)$ using the MD integration along a time $\tau = t_f - t_i$. The second condition is ensured by integrating over all the possible values of the momentum of the Hamiltonian. Adding a momentum to the action is allowed since the observable expectation value is conserved, i.e.

$$\langle O \rangle = \frac{1}{Z} \int \mathscr{D}U \, O(U) e^{-S(U)} = \frac{1}{Z \int \mathscr{D}\pi e^{-\pi^\dagger \pi/2}} \int \mathscr{D}U \, \mathscr{D}\pi \, O(U) e^{-H(U,\pi)} \quad \text{with} \quad H(U,\pi) = \frac{1}{2}\pi^\dagger \pi + S(U). \tag{5.24}$$

This is equivalent to the introduction of a new field, $\pi_\mu^a(x)$, in the action that we identify as the conjugate momentum of the Hamiltonian. The conjugate momentum has a weight $e^{-\pi^\dagger \pi/2}$ which is integrated by generating fields $\pi_\mu^a(x)$ from a Gaussian distribution.

For a gauge-theory the equations of motion are

$$\dot{U}_\mu(x) = \pi_\mu(x) U_\mu(x) \quad \text{and} \quad \dot{\pi}_\mu(x) = -F_\mu(x) = \frac{\partial S(e^\omega U_\mu)}{\partial \omega(x)} \bigg|_{\omega(x)=0} \tag{5.25}$$

where $\omega(x)$ is the generator of an infinitesimal gauge-field transformation.

The HMC algorithm proceeds as follow:

1. **Generation of conjugated momenta:** The momenta are generated via a normalized Gaussian distribution

$$P(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \tag{5.26}$$

2. **Integration of Hamilton's equations:** The Hamilton's equations given in Eq. (5.25) are integrated via a numerical integration scheme. Detailed balance requires that the integrator is time reversible and area preserving. In general the integration scheme can be constructed by the operations (see e.g. Ref. [129])

$$\mathscr{U}_\pi(\varepsilon): \ \pi \to \pi - \varepsilon F \quad \text{and} \quad \mathscr{U}_U(\varepsilon): \ U \to e^{\varepsilon \pi} U, \tag{5.27}$$

where $\mathscr{U}_\pi(\varepsilon)$ describes the update of the field π and $\mathscr{U}_U(\varepsilon)$ the update of the gauge-field U. Now, the integration from $t = 0$ to $t = \tau$ can be given by the leap frog scheme with

$$\mathscr{U}(\varepsilon, N) = \left(\mathscr{U}_\pi \left(\frac{\varepsilon}{2} \right) \mathscr{U}_U(\varepsilon) \mathscr{U}_\pi \left(\frac{\varepsilon}{2} \right) \right)^N \tag{5.28}$$

and $\varepsilon = \tau/N$. The integrator is time reversible since starting from the final point with opposite momenta, $\pi(\tau) \to -\pi(\tau)$, one returns to the starting point. The discretization error increases as $\delta H \propto N\varepsilon^3 = \varepsilon^2 \tau$ for this integration scheme. Note that the discretization error contributes randomly to the change of the energy resembling the update in the Metropolis algorithm.

3. **Metropolis accept-reject step:** The usage of a numerical integration scheme introduces discretization errors and the energy conservation is violated. For a proper sampling the weight of the new configuration has to be corrected. The correction is done by a Metropolis accept–reject step with the acceptance probability

$$P_{acc}(U(0) \to U(\tau)) = \min\left(1, e^{-\delta H} \right) \tag{5.29}$$

where δH is the change of the Hamiltonian and it is given by

$$\delta H = H(U(\tau), \pi(\tau)) - H(U(0), \pi(0)). \tag{5.30}$$

The steps 1-3 are iterated in a HMC simulation. Several remarks are in order here.

- The trajectory length τ and the number of integration steps N need to be tuned to get the best performance. A good compromise between high acceptance rate and small auto-correlation time needs to be found in order to reduce the computational cost.

- One can replace the leap-frog algorithm with higher order integration schemes to reduce the discretization error and therefore improve the acceptance. For more details see Ref. [130].

- The MD integration can be split into nested monomials [131], where different pieces of the action are integrated on different levels at increasing number of steps N. This allows to compute more precisely the parts of the Hamiltonian, which produce bigger forces.

5.3.1 HMC simulations for degenerate quark masses

When a degenerate fermionic doublet is introduced in the action, as done for Wilson fermions and for degenerate twisted mass fermions, the partition function of the theory takes the form

$$Z = \int \mathscr{D}U e^{-S_g(U)} \det \left[D^\dagger(U)D(U) \right] = \int \mathscr{D}U \, \mathscr{D}\eta \, e^{-S_g(U)-S_F(U,\eta)}, \tag{5.31}$$

where $S_F(U,\eta) = \eta^\dagger \left(D^\dagger(U)D(U) + \mu^2 \right)^{-1} \eta$ with $\mu^2 = 0$ for Wilson fermions and $\mu^2 > 0$ for twisted mass fermions. The pseudo-fermionic field η is integrated in HMC by considering the Hamiltonian

$$H(U,\pi,\eta) = \frac{1}{2}\pi^\dagger \pi + S_g(U) + \eta^\dagger \left(D^\dagger(U)D(U) + \mu^2 \right)^{-1} \eta. \tag{5.32}$$

The field η is generated randomly at the beginning of the trajectory and not evolved in the MD integration. The acceptance step in Eq. (5.29) involves then the difference of the Hamiltonian

$$\delta H = H(U(\tau), \pi(\tau), \eta(0)) - H(U(0), \pi(0), \eta(0)). \tag{5.33}$$

It is not necessary to evolve the pseudo-fermionic field η since it is only an auxiliary field introduced in the computation of the determinant of the Dirac operator whose value depends on U and not η. Therefore, generating η randomly at the beginning of every HMC step is enough for the MC procedure.

In the HMC algorithm we need to compute the inverse of $D^\dagger D + \mu^2$ at the beginning and at the end of the trajectory, in order to evaluate δH and twice for every calculation of the force for updating the momenta, as given in Eqs. (5.27) and (5.25). The fermionic contribution to the force is

$$-F_{\mu,f}(x) = \partial S_f = \eta^\dagger D_{\mathrm{TM}}^{-1\dagger}(\mu) \left. \frac{\partial D^\dagger(e^\omega U_\mu)}{\partial \omega(x)} \right|_{\omega=0} \left(D^\dagger D + \mu^2 \right)^{-1} \left. \frac{\partial D(e^\omega U_\mu)}{\partial \omega(x)} \right|_{\omega=0} D_{\mathrm{TM}}^{-1}(\mu)\eta, \tag{5.34}$$

which is computed in two steps performing

$$\phi = \left. \frac{\partial D(e^\omega U_\mu)}{\partial \omega(x)} \right|_{\omega=0} D_{\mathrm{TM}}^{-1}(\mu)\eta \quad \text{and} \quad -F_{\mu,f}(x) = \phi^\dagger \left(D^\dagger D + \mu^2 \right)^{-1} \phi. \tag{5.35}$$

The calculation of the inverse of D is then a critical component in the HMC, where $4N+2$ inversions of $D^\dagger D + \mu^2$ are required with N being the number of MD steps. It is worth noting that the force can be computed at low precision since this is corrected by the acceptance step.

A common improvement to the HMC integration with dynamical fermions is the so-called Hasenbusch mass preconditioning [132]. The fermionic determinant is transformed into

$$\det \left[D^\dagger D + \mu^2 \right] = \prod_{i=1}^{n-1} \det \left[\frac{D^\dagger D + \mu_i^2}{D^\dagger D + \mu_{i+1}^2} \right] \det \left[D^\dagger D + \mu_n^2 \right] \tag{5.36}$$

with $\mu_1 = \mu$ and $\mu_{i+1} > \mu_i$. This results in the integration of the action

$$S_{N_f=2}(U,\eta) = \sum_{i=1}^{n-1} \eta_i^\dagger \frac{D^\dagger(U)D(U) + \mu_{i+1}^2}{D^\dagger(U)D(U) + \mu_i^2} \eta_i + \eta_n^\dagger (D^\dagger(U)D(U) + \mu_n^2)^{-1}\eta_n, \tag{5.37}$$

where a number of n pseudo-fermionic fields are used. Although the Hasenbusch mass preconditioning seems to increase significantly the number of inversions, it speeds-up the algorithm. It is

important to note that the terms that produce a larger contribution to the force are the ones with a larger shift μ. Thus

$$F_n > F_{n-1} > \cdots > F_1, \tag{5.38}$$

where F_i is the force generated by the terms $\eta_i^\dagger \frac{D^\dagger D + \mu_{i+1}^2}{D^\dagger D + \mu_i^2} \eta_i$ for $i < n$ and F_n is the force generated by $\eta_n^\dagger (D^\dagger D + \mu_n^2)^{-1} \eta_n$. This allow to split the integration of the fermionic action in multiple monomials, computing more frequently the contribution to the force due to a large shift μ_i and few times for the target mass μ. Therefore, the number of inversion are effectively increased but the inversions at large μ_i are several order of magnitude faster than an inversion done for a μ tuned to the physical quark mass when CG is employed. Before the employment of multigrid methods, this reduced significantly the computational costs. Moreover, the terms with a small forces, like F_1, can be computed at low precision without increasing the integration errors. Then the physical point inversions are also accelerated requiring a less precise solution. This also reduces the integration errors since the total force F is split in several terms F_i which can be better controlled.

5.3.2 Rational HMC simulations for single flavor or non-degenerate quark masses

When we include a single flavor quark term or a non-degenerate twisted mass doublet to the action, as done in Eqs. (5.16) and (5.22), we need to compute the squared root of the inverse of $D^\dagger D + \mu^2$. This is significantly more complicated than computing an inverse since a direct numerical approach is not known and approximations of the square root need to be used. A first approach used in LQCD has been a polynomial approximation of the squared root, resulting in a polynomial HMC (PHMC) approach [133]. A more recent and better solution involves the employment of a rational approximation. The resulting approach is known as Rational Hybrid Monte Carlo (RHMC) [134] and we present it below.

Any continuous function $f(x)$ can be approximated by a rational function of generic order $[m,n]$

$$R(x) = A \frac{\prod_{i=1}^m (x + n_i)}{\prod_{i=1}^n (x + d_i)} \ . \tag{5.39}$$

The maximal deviation in a fixed interval of the rational approximation $R(x)$ from the function $f(x)$ is then bounded from below, as stated in the de-Vallée-Poussin's theorem [135]. A rational approximation is optimal when the maximal deviation is equal to the bound. In the case of the function $1/\sqrt{x}$ an optimal rational approximation is given by Zolotarev's solution [136]. The rational function of order $[n,n]$

$$R_{n,\varepsilon}(x) = a_{n,\varepsilon} \prod_{j=1}^n \frac{x + c_{n,\varepsilon,(2j-1)}}{x + c_{n,\varepsilon,2j}} \tag{5.40}$$

optimally approximates $1/\sqrt{y}$ in the interval $\varepsilon < x < 1$ with a maximal deviation

$$\delta_{n,\varepsilon} = \max_{\varepsilon < x < 1} \left| 1 - \sqrt{x} R_{n,\varepsilon}(x) \right| \ . \tag{5.41}$$

The parameters in Eqs. (5.40,5.41) can be computed analytically [129] and they are given by

$$c_{n,\varepsilon,k} = \mathrm{cs}^2\left(k \cdot v_{n,\varepsilon}, \sqrt{1-\varepsilon}\right) \qquad \text{with } v_{n,\varepsilon} = \frac{K\left(\sqrt{1-\varepsilon}\right)}{2n+1} \tag{5.42}$$

$$a_{n,\varepsilon} = \frac{2}{1+\sqrt{1-d_{n,\varepsilon}^2}}\prod_{j=1}^{n}\frac{s_{n,\varepsilon,(2j-1)}}{s_{n,\varepsilon,2j}} \qquad \text{with } s_{n,\varepsilon,k} = \mathrm{sn}^2\left(k \cdot v_{n,\varepsilon}, \sqrt{1-\varepsilon}\right) \tag{5.43}$$

$$\delta_{n,\varepsilon} = \frac{d_{n,\varepsilon}^2}{1+\sqrt{1-d_{n,\varepsilon}^2}} \qquad \text{with } d_{n,\varepsilon} = (1-\varepsilon)^{\frac{2n+1}{2}}\prod_{j=1}^{n}s_{n,\varepsilon,(2j-1)}^2, \tag{5.44}$$

where $\mathrm{sn}(u,k)$ and $\mathrm{cs}(u,k) = \mathrm{cn}(u,k)/\mathrm{sn}(u,k)$ are Jacobi elliptic functions and $K(k)$ is the complete elliptic integral.

The squared root of the inverse of $Q_{\mathrm{ND}}^2(\bar{\mu},\bar{\varepsilon}) = D_{\mathrm{ND}}^{\dagger}(\bar{\mu},\bar{\varepsilon})D_{\mathrm{ND}}(\bar{\mu},\bar{\varepsilon})$ is then approximated by a rational approximation of the kind

$$\sqrt{Q_{\mathrm{ND}}^{-2}(\bar{\mu},\bar{\varepsilon})} \simeq R_{n,\varepsilon}\left(Q_{\mathrm{ND}}^{-2}(\bar{\mu},\bar{\varepsilon})\right) = a_{n,\varepsilon}\prod_{i=1}^{N}\frac{Q_{\mathrm{ND}}^2 + c_{n,\varepsilon,(2i-1)}}{Q_{\mathrm{ND}}^2 + c_{n,\varepsilon,2i}}$$

$$= a_{n,\varepsilon}\left(1 + \sum_{i=1}^{n}q_i \cdot (Q_{\mathrm{ND}}^2(\bar{\mu},\bar{\varepsilon}) + m_i^2)^{-1}\right), \tag{5.45}$$

where in the latter we have split the product of fractions in a sum of fractions using

$$m_i^2 = c_{n,\varepsilon,2i} \qquad \text{and} \qquad q_i = (c_{n,\varepsilon,(2i-1)} - c_{n,\varepsilon,2i})\prod_{j=1,j\neq i}^{n}\frac{c_{n,\varepsilon,(2j-1)} - c_{n,\varepsilon,2i}}{c_{n,\varepsilon,2j} - c_{n,\varepsilon,2i}}. \tag{5.46}$$

The parameter ε has to be take such that the maximal and minimal eigenvalue of $Q_{\mathrm{ND}}^2(\bar{\mu},\bar{\varepsilon})$, $\lambda_{\max}$ and $\lambda_{\min}$ respectively, belong to the interval $\varepsilon < \lambda < 1$ with $\varepsilon = \lambda_{\min}/\lambda_{\max}$. Unfortunately, theoretical proofs, which determine the limits to the eigenspace of $Q_{\mathrm{ND}}^2(\bar{\mu},\bar{\varepsilon})$ are missing but from empirical arguments we can deduce that

$$|\bar{\mu} - \bar{\varepsilon}|^2 < \lambda_{\min} \lesssim |\bar{\mu}^2 - \bar{\varepsilon}^2| \qquad \text{and} \qquad \lambda_{\max} \lesssim (4 + |4 + m_c| + |\bar{\mu}| + |\bar{\varepsilon}|)^2, \tag{5.47}$$

where the upper limit of $\lambda_{\max}$ is proved in Ref. [1] in the case where no clover term is included in the operator.

Practically, in RHMC simulations values for $\lambda_{\min}$ and $\lambda_{\max}$ are computed for a large amount of configurations in the pre-production phase of the ensemble and then cross-checked frequently during the simulation. In Fig. 5.1 we depict the fluctuations of the smallest and largest eigenvalue for the ensemble produced in Ref. [8]. We measured the distribution of the smallest and largest eigenvalues of the squared even-odd reduced non-degenerate twisted mass operator on an ensemble of gauge-field configurations with a lattice of size 128×64^3 and a lattice spacing of $a \sim 0.08$ fm. It is generated with fermions parameters, such that the light, strange and charm quark masses are tuned close to their physical values. The value of the light quark mass is set by parameter $\mu = 0.00072$ and for the strange by $\bar{\mu} = 0.12469$ and $\bar{\varepsilon} = 0.13151$ sets the mass splitting between the strange and charm quarks. More details on this ensemble are given in Sec. 6.3. Using the values of these parameters the limits in Eq. (5.47) are

$$0.0000465 < \lambda_{\min} \lesssim 0.00175 \qquad \text{and} \qquad \lambda_{\max} \lesssim 68.2 \tag{5.48}$$

and the values depicted in Fig. 5.1 are all within the bounds.
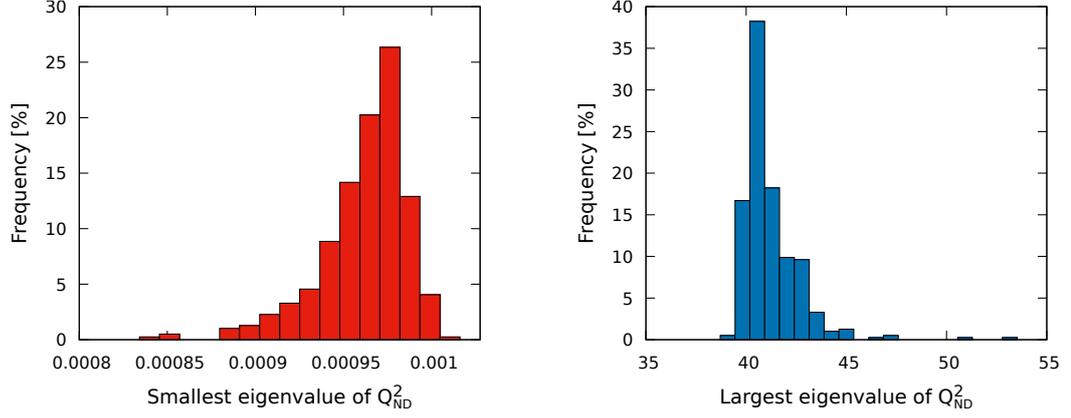
Figure 5.1: Distribution of the smallest (left) and largest (right) eigenvalues of the squared even-odd reduced non-degenerate twisted mass operator using an ensemble of gauge-field configurations tuned to the physical values of the light, strange and charm quark masses.

In the $N_f = 1 + 1$ twisted mass fermionic action, given in Eq. (5.22), we consider the rational approximation

$$R_{1+1} \equiv R_{n,\varepsilon}\left(Q_{1+1}^2\right) = a_{n,\varepsilon}\left(1 + \sum_{i=1}^{n} q_i \cdot (Q_{1+1}^2 + m_i^2)^{-1}\right), \qquad (5.49)$$

where we have defined the non-degenerate operator

$$Q_{1+1}^2 \equiv \frac{Q_{\text{ND}}^2(\bar{\mu}, \bar{\varepsilon})}{\lambda_{\text{max}}} \qquad \text{such that} \quad \varepsilon = \frac{\lambda_{\text{min}}}{\lambda_{\text{max}}} < \lambda_i\left(Q_{1+1}^2\right) < 1 \quad \forall i. \qquad (5.50)$$

The determinant of the non-degenerate twisted mass operator is transformed as

$$\det\left[D_{\text{ND}}(\bar{\mu}, \bar{\varepsilon})\right] = \det\left[\sqrt{Q_{1+1}^2}\right] = \det\left[R_{1+1}^{-1}\right] \det\left[|Q_{1+1}| R_{1+1}\right], \qquad (5.51)$$

where $|Q_{1+1}| R_{1+1}$ is a correction term. This term needs to be introduced since R_{1+1}^{-1} only approximates the square root.

The correction term is treated with an additional approximation given by the Taylor expansion

$$|Q_{1+1}| R_{1+1} = (1 + Z)^{1/2} = \sum_{n=0}^{\infty} \binom{1/2}{n} Z^n \quad \text{with } Z = Q_{1+1}^2 R_{1+1}^2 - 1, \qquad (5.52)$$

which is rapidly converging since $\|Z\|_{\infty} \leq \delta_{n,\varepsilon}^2$ and can be truncated after the Z^2 or Z^3 term [129, 137]. The more accurate the rational approximation, the closer the correction term is to the identity. Thus, the contribution to the force in the MD trajectory is small and for this reason the correction term is excluded from the MD trajectory. It is instead included either in the acceptance step or as a reweighting factor.

As for the Hasenbusch mass preconditioning, given in Eq. (5.36), the several fractions composing the rational approximation can be split in multiple determinants. We consider the following fractions

$$R_{1+1} \propto r_{1+1}^{1,a} \cdot r_{1+1}^{a+1,b} \cdot \ldots \cdot r_{1+1}^{n+1,N} \quad \text{with } r_{1+1}^{i,j} \equiv r_{n,\varepsilon}^{i,j}\left(Q_{1+1}^2\right) = \prod_{k=i}^{j} \frac{Q_{1+1}^2 + c_{n,\varepsilon,2k-1}}{Q_{1+1}^2 + m_k}, \qquad (5.53)$$

where $1 < a < b < \ldots < n < N$ are integers. Nested monomials can be used in the MD integration and

each term $r_{1+1}^{i,j}$ is placed on a different timescale and integrated at different frequency. The frequency is determined by studying the force intensity of the respective monomials in the molecular dynamic (MD) trajectory. As depicted in Fig. 5.2, the force intensity depends on the mass shifts m_k included in $r_{1+1}^{i,j}$ and the terms that have larger shifts produce higher contributions to the force – note m_1 is the largest mass and m_N the smallest.
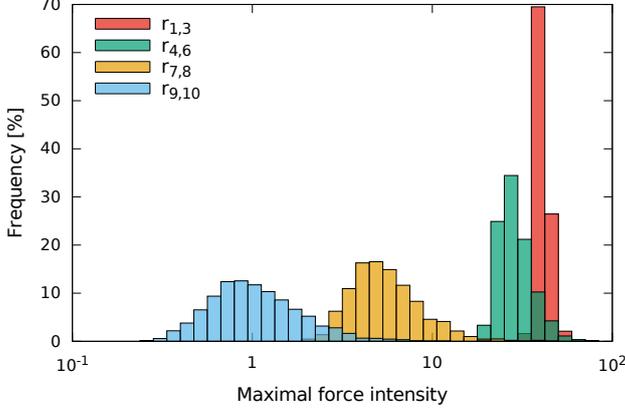
Figure 5.2: Distribution of the force intensity for the terms $r_{i,j} \equiv r_{1+1}^{i,j}$ using the squared even-odd reduced non-degenerate twisted mass operator on an ensemble of gauge-field configurations tuned to the physical values of the light, strange and charm quark masses.

5.4 Multigrid in shifted linear systems

The inverse operators entering the rational approximation of Eq. 5.45, can be treated as shifted linear systems. These are linear equations of the kind

$$(Q + m_i I)\, x_i = b \tag{5.54}$$

defined for a set of different mass shifts m_i acting on the diagonal of the operator Q and for a common right hand side b. Additionally to the case of the rational approximation, shifted linear systems have to be solved also for a polynomial approximation of the square root [133] or when using an integral definition via Stieltjes function [138, 139]. In general, the matrix roots have to be calculated in case of HMC simulations involving a single quark or in case of fermions represented by the Neuberger overlap operator [140].

The approaches routinely used to solve shifted linear equations like Eq. (5.54) are given by Krylov space solvers [138, 139, 141–143]. In what follows we will consider the multi-shift conjugated gradient (MS-CG) solver as our standard solver for shifted linear systems [141, 142]. The advantage of multi-shifts solvers is that the numerical effort scales only with the cost for solving the smallest shift and thus is almost independent of the total number of shifts. The idea behind these approaches is to exploit the identical eigenspaces by using the same Krylov space generated for the most ill-conditioned system to iterate all other shifts. However, if the operator Q^2 develops very small eigenvalues, multigrid solvers can be by far more effective as compared to the conjugate gradient solver. But a multi-shift solver version of mix-precision approaches or flexible preconditioners have not been successfully developed so far and this can further hinder the efficiency of multi-shift solvers.

The central idea of this section is to introduce a hybrid method, using MS-CG solver for systems with a larger mass-shift, while for the systems with smaller shifts employ a multigrid approach using extrapolated initial guesses. Usually, initial guesses are chosen to be zero, since it is the most safe starting point when the inverse is unknown. However, if parts of the inverse is known the iteration count can be reduced by starting with an initial vector close to the solution. Here, we will discuss two

approaches: one using Lagrange interpolation based on the solution of other shifts and the second using the MS-CG solver to construct guesses that will be used for the systems involving the smallest mass-shifts.

5.4.1 Initial guesses via Lagrange interpolation

The idea for the employment of multigrid in multi-mass shifted systems is to use previously computed solutions, e.g. $x_1, \ldots, x_n$, to generate an initial guess for the next inversion of the solution x_{n+1}. This can be done by a Lagrangian interpolation of the previous solutions, where the previous shifts enter as node points. The Lagrangian interpolation of a function $f(x)$ is given by

$$L_k(x) = \sum_{i=1}^{k} f(x_i) l_{i,k}(x), \tag{5.55}$$

where $k > 1$ and $l_{i,k}(x)$ satisfies $l_{i,k}(x_j) = \delta_{ij}$ for all $i, j \in [1, k]$ with δ_{ij} being the Kronecker delta. A polynomial solution to the latter property is

$$l_{i,k}(x) = \prod_{\substack{j=1 \\ j \neq i}}^{k} \frac{x - x_j}{x_i - x_j} \quad \implies \quad l_{i,k}(x_j) = \delta_{ij} \text{ for } i, j \in [1, k], \tag{5.56}$$

where the Lagrangian interpolation defined via $l_{i,k}(x)$ is denoted as the Lagrange's form. In this case, $L_k(x)$ is the unique polynomial of degree $(k-1)$ that exactly interpolates k fixed points of the function $f(x)$, i.e. $L(x_i) = f(x_i)$ for all $i \in [1, k]$. Additionally, we define $l_{1,1}(x) \equiv 1$ that gives

$$\sum_{i=1}^{k} l_{i,k}(x) = 1 \qquad \text{for all } k \in \mathbb{N}^+ . \tag{5.57}$$

Furthermore, it follows that the Lagrangian interpolation of a constant is exact, $L_k(x) = \sum_{i=1}^{k} b\, l_{i,k}(x) = b \sum_{i=1}^{k} l_{i,k}(x) \equiv b$ and the Lagrangian interpolation of $(Q + mI)^{-1}$ with grid points $\{(Q + m_i I)^{-1}\}$ is given by

$$L_k(m) = \sum_{i=1}^{k} l_{i,k}(m)(Q + m_i I)^{-1} \quad \text{with } l_{i,k}(m) = \prod_{\substack{j=1 \\ j \neq i}}^{k} \frac{m - m_j}{m_i - m_j}. \tag{5.58}$$

The interpolated solution reads as

$$\tilde{x}_k(m) = L_k(m)b = \sum_{i=1}^{k} l_{i,k}(m)(Q + m_i I)^{-1} b = \sum_{i=1}^{k} l_{i,k}(m)x_i, \tag{5.59}$$

where x_i are solutions of $(Q + m_i I)x_i = b$ computed with a residual $r_i = b - (Q + m_i I)x_i$ that fulfills the solver stopping criterion $\|r_i\| < \rho \|b\|$. For the residual of the interpolated solution $\tilde{x}_k(m)$ we have

$$\tilde{r}_k(m) = b - (Q + mI)\tilde{x}_k(m) = b - \sum_{i=1}^{k} l_{i,k}(m)(Q + mI)x_i \tag{5.60}$$

$$= \sum_{i=1}^{k} l_{i,k}(m)\left((m_i - m)x_i + b - (Q + m_i I)x_i\right) = \sum_{i=1}^{k} l_{i,k}(m)\left((m_i - m)x_i + r_i\right),$$

which is a Lagrangian interpolation of the residuals, i.e. $\tilde{r}_k(m_i) = r_i$. Studying the norm of the Lagrange form of the residuals we obtain

$$\left\| \sum_{i=1}^{k} l_{i,k}(m) r_i \right\| \leq \sum_{i=1}^{k} |l_{i,k}(m)| \, \|r_i\| < \rho \, \|b\| \sum_{i=1}^{k} |l_{i,k}(m)| = \rho \, \gamma_k(m) \, \|b\|, \tag{5.61}$$

where $\gamma_k(m) = \sum_{i=1}^{k} |l_{i,k}(m)|$ is the Lebesgue function defined from the Lagrange's polynomials $l_{i,k}(m)$. The Lebesgue function in the interval $[m_{\min}, m_{\max}]$ assumes values

$$1 \leq \gamma_k(m) \leq \Gamma_k, \tag{5.62}$$

where $m_{\min}$ and $m_{\max}$ are the smallest and largest shifts of m_i, respectively. $\Gamma_k = \max_{m \in [m_{\min}, m_{\max}]} \gamma_k(m)$ is referred to as Lebesgue constant. Depending on the shifts, the Lebesgue constant could grow exponentially, logarithmically or as an asymptotic function of k [144]. In case of diverging growths one can truncate the degree of the interpolation in order to keep the error under control. The relation

$$\left\| \sum_{i=1}^{k} l_{i,k}(m) r_i \right\| \leq \rho \, \Gamma_k \, \|b\| \quad \text{for} \quad m \in [m_{\min}, m_{\max}] \tag{5.63}$$

fixes the maximal contribution of the residuals to the residual of the interpolated solution.

The additional term to the Lagrange form in Eq. (5.60) can be re-written as

$$\sum_{i=1}^{k} l_{i,k}(m)(m - m_i) x_i \simeq \left(\prod_{j=1}^{k} (m - m_j) \right) \sum_{i=1}^{k} \left(\prod_{\substack{j=1 \\ j \neq i}}^{k} \frac{1}{m_i - m_j} \right) (Q + m_i I)^{-1} b$$

$$= \left(\prod_{j=1}^{k} \frac{m - m_j}{Q + m_j I} \right) b, \tag{5.64}$$

where $x_i \simeq (Q + m_i I)^{-1} b$ is used and the partial fraction decomposition is re-summed in a product of fractions. For the norm of it follows that

$$\left\| \sum_{i=1}^{k} l_{i,k}(m)(m - m_i) x_i \right\| \simeq \left\| \left(\prod_{i=1}^{k} \frac{m - m_i}{Q + m_i I} \right) b \right\| \leq \prod_{i=1}^{k} \frac{|m - m_i|}{|\lambda_{\min,i}|} \|b\|, \tag{5.65}$$

where $\lambda_{\min,i}^2$ is the smallest eigenvalue of $(Q + m_i)^\dagger (Q + m_i)$. If Q is a positive-definite matrix then $\lambda_{\min,i} = \lambda_{\min} + m_i$ with $\lambda_{\min} > 0$ being the smallest eigenvalue of Q.

Considering now the full residual interpolation in Eq. (5.60) we find the following upper limits

$$\|\tilde{r}_k(m)\| = \left\| \sum_{i=1}^{k} l_{i,k}(m) \left((m - m_i) x_i + r_i \right) \right\| \leq \left\| \sum_{i=1}^{k} l_{i,k}(m)(m - m_i) x_i \right\| + \rho \, \gamma_k(m) \, \|b\|$$

$$\lesssim \left(\prod_{i=1}^{k} \frac{|m - m_i|}{|\lambda_{\min,i}|} + \rho \, \Gamma_k \right) \|b\|. \tag{5.66}$$

The first upper bound requires a knowledge of the solutions x_i and residuals r_i, making it dependent on the numerical approach. The second instead depends only on the analytical properties of the shifted systems.

From the second upper bound in Eq. (5.66) we obtain that the residual of the $n+1$ solution is

$$\frac{\|\tilde{r}_{n+1}\|}{\|b\|} \lesssim \gamma_n \rho + \prod_{i=1}^{n} \frac{|m_{n+1} - m_i|}{\lambda_{\min} + m_i} \tag{5.67}$$

if $m_i > 0$ and Q a positive-(in)definite matrix with $\lambda_{\min} \geq 0$. The residual of the initial guess $\tilde{r}_{n+1}$ is bounded by two terms. The first term depends on the nodes m_i while the second term depends additionally on the smallest eigenvalue of the matrix. It follows that $\|\tilde{r}_{n+1}\| / \|b\| > \rho$ since $\gamma_n > 1$. However, γ_n is known *apriori* such that the interpolation strategy can be adapted if γ_n becomes too larger. In our case γ_n is around 2 and for $\rho < 10^{-6}$ the second term dominates the right hand side of Eq. (5.67) due to the dependence on λ_{min}. Moreover the second term is strictly smaller than one for the case $m_{n+1} < m_i$. This gives the optimal ordering to solve the multi-shifted problem via initial guesses with

$$m_1 > m_2 > \ldots > m_n > m_{n+1} > \ldots > m_N > 0, \tag{5.68}$$

whereby it follows that

$$0 \leq \prod_{i=1}^{n} \frac{|m_{n+1} - m_i|}{\lambda_{\min} + m_i} < 1. \tag{5.69}$$

Finally, we remark that if the upper bound of Eq. (5.67) is smaller than 1, starting from $\tilde{x}_{n+1}$ will be always more efficient than from a zero vector.

5.4.2 Initial guesses via MS-CG

Another possibility to guess a starting vector for the last $(N - n)$ shifts is given by using the MS-CG solver. The general idea of the MS-CG solver is to exploit the fact that the eigenspace of the shifted systems are identical. Thus the Krylov space generated for one of the shifts can be simultaneously used to iterate the other shifts.

If one generates the Krylov space for the most ill-conditioned system, here the Nth system, all other iteration vectors will converge to smaller residuals than the residual of the target system. However, if the system is too ill-conditioned, like it is in our case, the iteration count of the MS-CG solver increases drastically and a hybrid-approach is potentially more efficient.

Our proposal is to solve the first n systems via the MS-CG solver by generating the Krylov space for the nth system. This will generate the first n solutions $x_1, x_2, \ldots, x_n$. Furthermore, the MS-CG solver can also predict guesses for the next m systems by iterating those together with the first n systems. Although the iteration vectors of these systems will not reach the target precision, at the iteration count where the nth system is converged, the iteration vectors will contain the full information of the generated Krylov space for the nth system. Based on this fact, for the $(n+1)$th system, the MS-CG solver generates an initial vector, which is in general closer to the target residual than using a Lagrange interpolation based on the n solutions.

5.5 Multigrid in the rational approximation of the square root

In this section, we analyze the behavior of multigrid for the rational approximation of the square root of Eq. (5.45), for the non-degenerate twisted mass operator using the ensemble $cB211.64$, cf. Tab. 6.1, at physical strange and charm quark masses. We choose a rational approximation consisting of 10 terms and restring the eigenvalues of the operator in the interval $(\varepsilon_{\mathrm{ND}} ; 1]$ with $\varepsilon_{\mathrm{ND}} = 0.000065/4.7$,

as follow from Fig. 5.1.

5.5.1 Initial guesses for the rational approximation of the square root

Using the Lagrange interpolation we obtain an upper bound for the initial guesses through Eq. (5.67), bounded by two terms, which depend on γ_n and the previous mass-shifts m_i, respectively. The coefficient γ_n depends as well on the nodes of the mass-shifts and with an ordering $m_1 > m_2 > \ldots > m_N$ for the shifts we find that γ_n is never larger than 2. This bounds the first term of Eq. (5.67) by 2ρ where ρ is the precision of the stopping criterion. Thus, the first term is suppressed compared to the second term of Eq. (5.67) such that we can neglect it in the following.

The second term of Eq. (5.67) is generated by

$$\left\| \sum_{i=1}^{n} c_{i,n}(m_{n+1} - m_i) x_i \right\| \lesssim \|b\| \prod_{i=1}^{n} \frac{|m_{n+1} - m_i|}{\lambda_{\min} + m_i} \tag{5.70}$$

with $\lambda_{\min} \sim 0.0008$ for the squared even-odd reduced non-degenerate twisted mass operator. Using an ordering $m_1 > m_2 > \ldots > m_N$ this yields an upper bound of 0.0003 for the last, the Nth, initial guess, which is close to the residual shown in Fig. 5.3. However, this upper bound significantly overestimates the relative residual of the first shifts being up one order of magnitude larger than the values depicted in the figure. We find that the data can be described effectively via

$$\left\| \sum_{i=1}^{n} c_{i,n}(m_{n+1} - m_i) x_i \right\| \simeq \|b\| \prod_{i=1}^{n} \frac{|m_{n+1} - m_i|}{A_i m_i}, \tag{5.71}$$

where the A_i are similar to each other and smaller then 1.9 for all i. Using $B = [\max(A_i)]^{-1} < 1$ it follows

$$\left\| \sum_{i=1}^{n} c_{i,n}(m_{n+1} - m_i) x_i \right\| \simeq B^n \|b\| \prod_{i=1}^{n} \left(1 - \frac{m_{n+1}}{m_i} \right) \tag{5.72}$$

with $m_{n+1} < m_i$ and $B \cong 0.6$. Thus, the initial guesses using the Lagrange interpolation become more efficient with increasing n, as shown in figure Fig. 5.3.

As pointed out in the previous section, the MS-CG solver, used for solving the first n systems, can be used to predict an initial guess by including the $(N - n)$ systems in the MS-CG iteration. This effectively interpolates the initial vector x_j in the Krylov space of the nth system. We depict the residual for the predicted initial guesses by the MS-CG solver in Fig. 5.4 if the nth system is converged. As shown, the relative residual of the $(n+1)$th system depends only slightly on n and is given by $\sim 10^{-5}$. Thus, for the first step after using the MS-CG solver the residual is smaller then the residual of the initial guess generated by the Lagrange interpolator. However, this changes for $m > 1$. While for small n the guess using the MS-CG solver for the $(n+2)$th system is better, for $n \gtrsim 5$ the guesses of the Lagrange interpolation yield similar results and becomes better for $m > 2$.

Based on these results, the optimal approach for our example is a combination of all three elements, namely the MS-CG solver and the multigrid solver with initial guesses using the MS-CG solver and the Lagrange interpolation. Thus, solving the first nth systems using MS-CG solver involves an additional iteration for the $(n+1)$th and $(n+2)$th systems, which can be used as an initial guess. The $(N - n)$ remaining systems are then solved via a multigrid approach one by one, where for the last $(N - n - 2)$ systems the Lagrange interpolation is used to start the iteration. In the following section, we will discuss the optimal n.
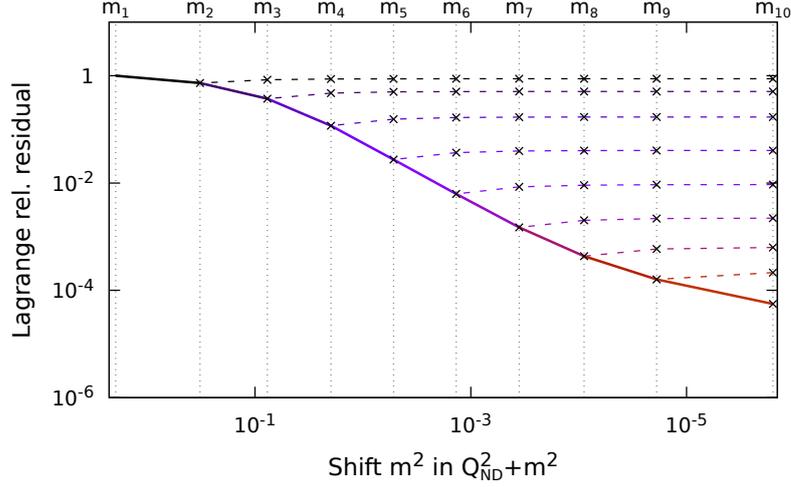
Figure 5.3: The relative residuals of the Lagrange interpolation based on the n solutions versus the shifts m^2 used in the approximation of the square root of the squared even-odd reduced ND twisted mass operator. The solid lines illustrate the norm of the relative residual of the initial guess for decreasing m taking all available nodes into account. The nodes used in the interpolation, i.e. the shifts m^2, are denoted by vertical dotted lines. We depict the dependence of each polynomial $p^{(i)}$ for $m^2 > m_i^2$ by the dashed lines.
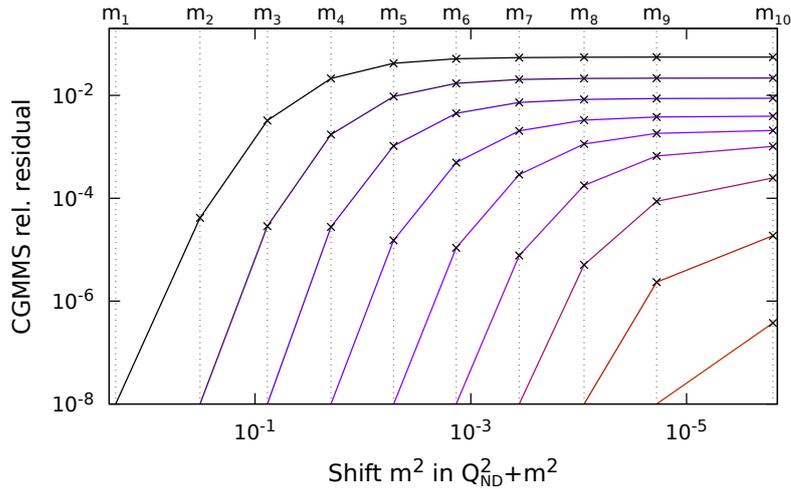
Figure 5.4: The relative residual of the iteration vectors of the MS-CG versus the mass shifts when varying the target Krylov space for the non-degenerate twisted mass case. The solver is stopped when the residual of the target system reaches the stopping criterion given here by $\|r\| / \|b\| < 10^{-8}$. The iteration vectors of systems with smaller shift than the target one have a larger residual. This is shown for all $N-1$ cases depicted by the solid line changing the color from black, for the largest target shift to orange for the $N-1$ shift. Note that we do not show the case where the target system is given by the last, Nth system, because here all other systems have converged. The vertical dotted lines illustrate the shifts m_i^2.

5.5.2 DD-αAMG speed-up in the rational approximation of the square root

Based on the observations in the previous section, we propose to use a hybrid approach to solve a system of linear equations with N shifts. Namely, use the MS-CG solver for the largest n shifts and the DD-αAMG approach for the remaining $(N-n)$ systems solving each one via the different methods discussed in section 4.5.2. The $(N-n)$ systems can be started by initial guesses, proposed for the $(n+1)$th and $(n+2)$th systems using the MS-CG solver and for the rest using Lagrange interpolation and employing the previous solutions.

The optimal n depends on the ratio of the performance of the solvers, which includes environ-

mental parameters, software implementation and computer hardware. Here, we consider the Haswell-nodes partition of SuperMUC and use an MPI parallelization employing 1024 task on 37 nodes. The software used is the tmLQCD package [7, 90], which is linked to the DDalphaAMG library [111] and is publicly available.

The question we would like to answer is what is the optimal n of the hybrid approach introduced in this section, i.e. how many n shifts should be solved with the MS-CG solver in order to solve the total system with N shifts in an efficient way employing for the last $N - n$ shifts the DD-αAMG approach. We discuss it on the $cB211.64$ ensemble using the non-degenerate twisted mass operator at the strange and charm quark masses with 10 mass-shifts. We employ Eq. (4.49) for solving the linear equation with the squared operator when the multigrid solver is used.
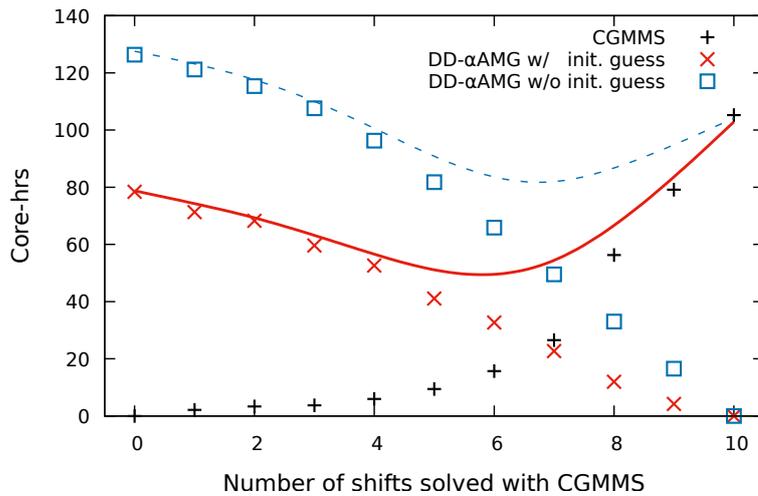
Figure 5.5: The CPU hours necessary to solve the multi-mass shifted equation involving the squared even-odd reduced non-degenerate twisted mass operator at the strange and charm quark masses versus the number of shift n used in the MS-CG solver shown as black crosses. The left $10 - n$ shifts are solved via the multigrid approach with initial guesses as red crosses and without initial guesses as blue squares. The total CPU hours needed to solve the system is shown by the red solid line for the hybrid method using initial guesses and for the hybrid method without initial guesses with the red solid line.

The cost for the solution of the multi-mass shift linear equation via the hybrid method is given by

$$c_{HY}(n) = t_{CG}(m_n^2) + \sum_{i=n+1}^{N} t_{MG}(m_i^2) , \qquad (5.73)$$

where the time to solution of the MS-CG solver is approximated with the time of CG solver $t_{CG}(m)$ at the smallest shift m_n. For the case without initial guesses, the cost $c_{HY}(n)$ can be minimized using the fits of Eqs. (3.26) and (4.14). This yields an optimal $n_{opt} \cong 7$ as shown in Fig. 5.5. However, the total speed-up of the hybrid method without initial guesses only improves slightly the time to solution compared to the application of the MS-CG solver. This changes by using initial guesses. Here, the cost for the multigrid part is significantly reduced as shown in Fig. 5.5 using initial guesses compared to the case without. We find that the initial guesses reduce the total time to solution by about factor of two while the optimal $n_{opt} \sim 6$ is shifted slightly.

The hybrid method becomes even more effective for smaller quark masses. To illustrate this, the n-dependence is examined using the non-degenerate twisted mass operator at physical light quarks using the physical test ensemble. For $\bar{\mu}_\ell = 0.00072$ and $\bar{\varepsilon}_\ell = 0.000348$ the dependence of the hybrid method using a rational approximation with $N = 15$ terms is shown in Fig. 5.6. We find that the hybrid
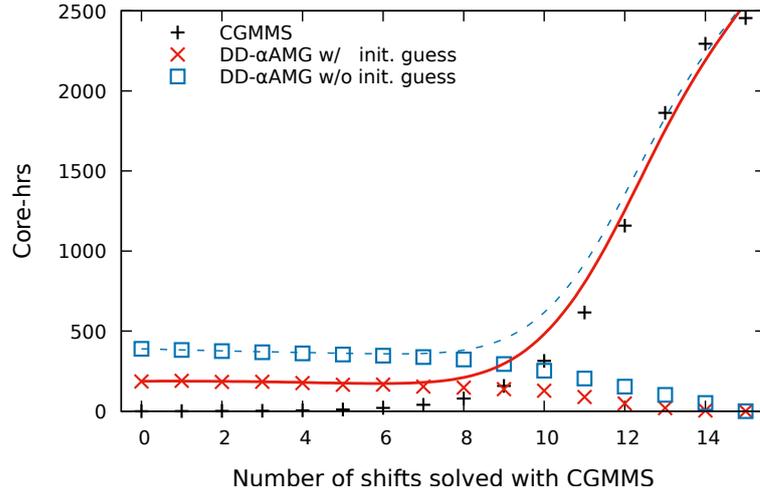
Figure 5.6: The CPU hours necessary to solve the multi-mass shifted equation involving the squared even-odd reduced non-degenerate twisted mass operator for the physical value of the up and down quark mass. The dependence on the number of shift n used in the MS-CG solver is shown as black crosses. The left $15-n$ shifts are solved via the multigrid approach and results with initial guesses are shown with the red crosses while without initial guesses with the blue squares. The total CPU hours needed to solve the system is given by the red solid line for the hybrid method using initial guesses and with the dotted blue line for the hybrid method without initial guesses.

method with initial guesses gives a total speed up at optimal $n_{opt} \cong 7$ by approximately a factor of 15. As in the case of the strange and charm quark masses, the initial guesses, result in a speedup of about a factor of two.

6. Simulating twisted mass fermions at the physical point with multigrid methods

Simulations directly at physical quark masses with large enough volume and small enough lattice spacing have become feasible in the recent year thanks to significant algorithmic improvements, such as multigrid methods, and the availability of substantial computational resources. In fact, state-of-the-art simulations using different discretization schemes are currently being carried out worldwide.

Within the twisted mass formulation, the Extended Twisted Mass Collaboration (ETMC) has carried out the first simulation directly at the physical value of the pion mass [76, 145] with $N_f = 2$ mass-degenerate up and down quarks at a lattice spacing of $a = 0.0913(2)$fm, the ensemble $cA2.48$ in Tab. 3.1. This has been a remarkable result, since explicit isospin breaking effects associated with twisted mass fermions can make physical point simulations at too coarse values of the lattice spacing very difficult. Being able to reach the physical pion mass for the case of $N_f = 2$ flavours was therefore of great importance and many physical quantities have already been computed on the so generated gluon field configurations. Examples are meson properties [9–11, 76, 145, 146], the structure of hadrons [12, 145, 147–150] and the anomalous magnetic moment of the muon [76].

The ensemble $cA2.48$ has been simulated with a huge computational effort since multigrid solvers where not yet available for twisted mass fermions. After our extension of the DD-αAMG to twisted mass fermions, the solver has been employed largely within the collaboration. Many of the observables measured on the $cA2.48$ ensemble have been computed with it. After the success of this first $N_f = 2$ flavor simulation we have performed other three simulations at the physical point and a fourth is currently under production. The properties of these simulations are summarized in Tab. 6.1 and depicted in Fig. 6.1 together with other ensembles outside of the physical point produced by ETMC.

Name[1]	Ref.	N_f	V	a [fm]	m_π [MeV]	$m_\pi L$
$cA2.48$	[76]	2	$48^3 \times 96$	0.0913(2)	131.9(3)	3.004(5)
$cA2.64$	[76][2]	2	$64^3 \times 128$	0.0913(2)	131.9(3)	4.005(5)
$cA2.64.r$	[76][2]	2	$64^3 \times 128$	0.0913(2)	135.0(3)[3]	4.100(5)
$cB211.64$	[8]	2+1+1	$64^3 \times 128$	0.0820(4)	136.1(7)	3.71(1)
$cC211.80$	N.A.[4]	2+1+1	$80^3 \times 160$	0.07 $\sim$	135 $\sim$	3.9 $\sim$

Table 6.1: We give the lattice volume V, the lattice spacing a, the pion mass m_π and the value of $m_\pi L$ for the physical point ensembles produced by the ETM collaboration.

In this chapter we review the production of these ensembles. We start from the improvements in the numerical sector of the simulation where we employ the DD-αAMG solver for the inversions [2, 4, 6]. We start from the $N_f = 2$ simulations showing results for the production of the ensembles $cA2.64$ and $cA2.64.r$ in Sec. 6.1. These ensembles did not require additional effort in setting up the simulation since they use similar parameters to $cA2.48$ whose tuning has been presented in Ref. [76].

---

[1] Our nomenclature is generated by the following rules: c stands for clover, $A/B/C$ stands for the lattice spacing category, 2/211 stands for the number of flavors, 48/64/80 for the lattice size and r for replica.

[2] These ensembles are not part of the publication, but the same parameters and simulation setup of $cA2.48$ have been used.

[3] The ensemble $cA2.64.r$ differ from $cA2.64$ only by the twisted mass term which has been tuned to give a more precise physical pion mass. Namely we used $\mu = 0.000966$ instead of $\mu = 0.0009$.

[4] The ensemble $cC211.80$ is currently under production.
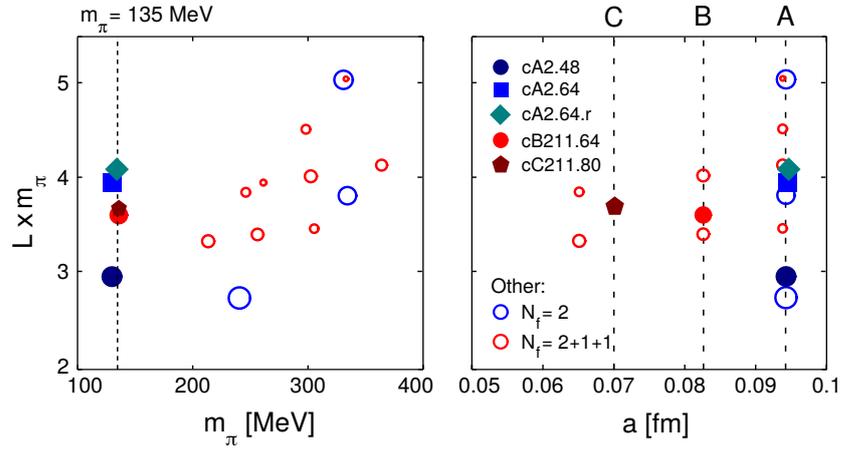
Figure 6.1: Constellation of ensembles produced by the ETM collaboration. The same ensembles are depicted on the left as a function of the pion mass m_π and on the right as a function of the lattice spacing. The radius of the circles scales proportionally to the number of independent configurations in the ensembles. The five ensembles at the physical point (dashed vertical line on the left) are the ensemble listed in Tab. 6.1.

In Sec. 6.2 we review numerical approach in our $N_f = 2 + 1 + 1$ simulations using results from the production of the ensemble $cB211.64$. And we conclude the chapter with a sections dedicated to the ensemble $cB211.64$ where we show the tuning of the parameters and the determination of the lattice spacing in Sec. 6.3.

6.1 N_f=2 twisted mass simulations at the physical point

The $N_f = 2$ twisted mass action is

$$S = S_g + S_{\text{TM},N_f=2} \ , \tag{6.1}$$

where $S_{\text{TM},N_f=2}$ has been given in Eq. (5.20) and for S_g we chose the Iwasaki improved gauge action, cf. Eq. (2.47).

6.1.1 Simulations details

The tuning of the simulation parameters for the ensemble $cA2.48$ was carried out in Ref. [76]. For simulating the ensembles $cA2.64$ and $cA2.64.r$ we used similar parameters where additionally to the volume we changed the hopping parameter κ, in order to be closer to maximal twist, and the twisted mass parameter μ, in $cA2.64.r$ only, in order to have a pion mass closer to the physical one. We list the simulation parameters in Tab. 6.2.

Name	Volume	β	μ	κ	c_{SW}
$cA2.48$	$48^3 \times 96$	2.1	0.0009	0.13729	1.57551
$cA2.64$	$64^3 \times 128$	2.1	0.0009	0.1372938	1.57551
$cA2.64.r$	$64^3 \times 128$	2.1	0.000966	0.1372938	1.57551

Table 6.2: The table shows the simulations parameters for the ensembles $cA2.48$, $cA2.64$ and $cA2.64.r$ which physical properties are listed in Tab. 6.1.

The simulations have been produced using the tmLQCD software package [90] and the DDalphaAMG library for TM fermions [2]. All the simulation codes are released under GNU license. The integrator we used is given by a nested minimal norm scheme of second order [131] with a nested

integration scheme similar to the previously produced simulation [76]. In the $N_f = 2$ sector of the simulations we employ the Hasenbusch mass preconditioning, cf. Eq. (5.36), by introducing additional mass terms and split up the determinants into additional ratios as in the following

$$\det\left[Q^2 + \mu^2\right] = \det\left[\frac{Q^2 + \mu^2}{Q^2 + (\mu + \rho_1)^2}\right] \det\left[\frac{Q^2 + (\mu + \rho_1)^2}{Q^2 + (\mu + \rho_2)^2}\right] \det\left[\frac{Q^2 + (\mu + \rho_2)^2}{Q^2 + (\mu + \rho_3)^2}\right] \det\left[Q^2 + (\mu + \rho_3)^2\right]$$

(6.2)

where $Q = \Gamma_5 D = Q^\dagger$ is the hermitian Wilson Dirac operator.

Each determinant in the right hand side (rhs) of Eq. (6.2) can be placed on a different monomial and integrated on a different time-scale. The time-steps are chosen accordingly to the intensity of the force term. This procedure controls the large fluctuations of the force terms, avoiding instabilities during the HMC. In our simulations we use 5 time-scales, which are integrated respectively $N_{\mathrm{int}} = \{20, 60, 180, 540, 1620\}$ times. The gauge action is placed in the innermost time-scale and in the other time-scales we place one by one the fermionic determinants from the rhs of Eq. (6.2) going from the largest shift to the smallest.

6.1.2 DD-αAMG speed-up in N_f=2 simulations

The simulations are performed for the even-odd reduced operator while the DD-αAMG solver inverts the full operator. The even-odd reduced solution is retrieved by projecting the complete solution provided by DD-αAMG. For the heat-bath inversions and acceptance steps we require as stopping criterion for the solvers the relative residual to be smaller than 10^{-11}. For the force terms in the MD trajectory the criterion is relaxed, using 10^{-7} for the CG solver and 10^{-9} for the DD-αAMG method. This ensures that the reversibility violation of the MD integration is sufficiently reduced. Note that the usage of the multigrid method is efficient if the subspace can be reused at larger integration time. In general, this yields a larger reversibility violation [96]. By choosing a higher accuracy i.e. by using a smaller stopping criterion, for the multigrid solver the reversibility violation can be reduced. We checked that with the values mentioned above the reversibility violations are compatible with the case where a CG solver is used for all the monomials. As depicted in Fig. 6.2, this yields a stable simulation without large spikes in the energy violation δH with an acceptance rate of 84.5%.

For the Hasenbusch mass preconditioning we use the three shifts as reported in Eq. (6.2). The DD-αAMG method is faster than CG solver for all the shifts except the largest, ρ_3. Thus we have used DD-αAMG for the inversions with shifts μ, $\mu + \rho_1$ and $\mu + \rho_2$. The DD-αAMG iterations count averaged per MD trajectory is depicted in Fig. 6.2. No exceptional fluctuations or correlations with larger energy violation δH are seen along the simulation. The stability of the multigrid method is ensured by updating the setup every time the inversion at the physical quark mass, i.e. $(Q^2 + \mu^2)^{-1}$, is performed. The update is based on the previous setup by using one setup iteration, which is possible due to the adaptivity in the DD-αAMG method [105] for the definition of the setup iteration. At the beginning of the trajectory we perform three setup iterations. The final speed-up, including setup costs, is a factor of 8 compared to CG in $N_f = 2$ simulations at the physical pion mass.

6.2 N_f=2+1+1 twisted mass simulations at the physical point

In N_f=2+1+1 simulations we follow the same approach of N_f=2 simulation regarding the gauge field and the N_f=2 sector, cf. Sec. 6.1. Additionally we add to the action a N_f=1+1 term which simulates the strange and charm quark flavors. This term involves the calculation of the square root of the
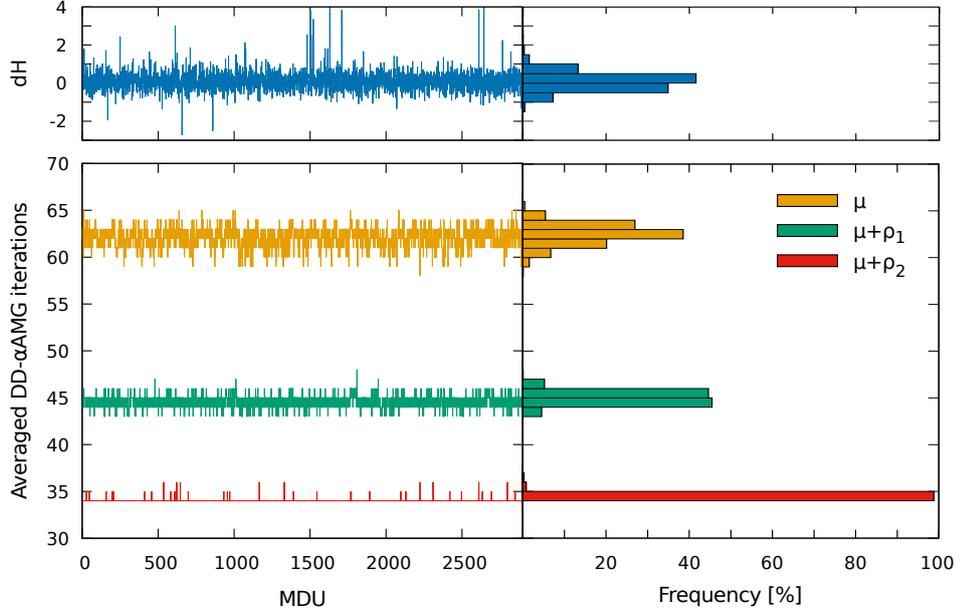
Figure 6.2: Results for the ensemble *cA*2.64.*r* produced on Supermuc phase 2.
Top: The energy violation of the numerical integrator in the MD integration plotted in units of MD trajectories (left) and its frequency (right).
Bottom: DD-αAMG iterations count averaged on MDU considering different monomials of the HMC simulation shown in units of MD trajectories (left) and its frequency (right). The shifts ρ_i are $\{0.0015, 0.015\}$ and $\mu = 0.000966$. The iterations count corresponds to average of the sum of the count of the outer level during the calculation of the force in the MD. A similar frequency plot has been depicted in Fig. 4.9 which was for the ensemble *cA*2.64.

inverse of the non-degenerate TM operator, cf. Eq. (5.22), which cannot be computed directly with numerical methods. We use a rational approximation of the square root which we integrate via a RHMC approach, cf. Sec. 5.3.2. We recall that the determinant of the degenerate operator is written as, cf. Eq.(5.51)

$$\det\left[D_{\mathrm{ND}}(\bar{\mu}, \bar{\varepsilon})\right] = \det\left[\sqrt{Q_{\mathrm{ND}}^2}\right] = \det\left[R_{\mathrm{ND}}^{-1}\right] \det\left[\left|Q_{\mathrm{ND}}\right| R_{\mathrm{ND}}\right], \tag{6.3}$$

where $Q_{\mathrm{ND}} = (\Gamma_5 \otimes \tau_1)D_{\mathrm{ND}} = Q_{\mathrm{ND}}^{\dagger}$ is the hermitian version of D_{ND} and R_{ND} is the optimal rational approximation of $1/\sqrt{Q_{\mathrm{ND}}^2}$. In general, one can relax the approximation of the square root by introducing a correction term $\det\left[\left|Q_{\mathrm{ND}}\right| R_{\mathrm{ND}}\right]$ as in the rhs of Eq. (6.3). It takes into account the deviation from $\left|Q_{\mathrm{ND}}\right|$ being closer to the identity as much as the rational approximation is precise. For this reason we include it only in the acceptance step.

6.2.1 Simulation details

We list the simulation parameters for the *cB*211.64 and *cC*211.80 in Tab. 6.3. In Sec. 6.3 we will review the tuning for generating the *cB*211.64 ensemble [8]. The tuning of *cC*211.80 followed a similar procedure.

For the production of the gauge field configurations we use as a basis the simulation setup for the N_f=2 simulations, see Sec. 6.1.1. In particular, here we employ four determinant ratios for the N_f=2 sector with mass shifts $\rho = \{0.0; 0.0003; 0.0012; 0.01; 0.1\}$ and we use a 6 level nested minimal norm second order integration scheme with integration steps $N_{\mathrm{int}} = \{12, 36, 108, 324, 972, 2916\}$. As for the N_f=2 simulations we place the integration of the gauge field in the innermost level and the N_f=2

Name	Volume	β	μ_ℓ	$\bar{\mu}$	$\bar{\varepsilon}$	κ	c_{SW}
$cB211.64$	$64^3 \times 128$	1.778	0.00072	0.1246864	0.1315052	0.1394265	1.69
$cC211.80$	$80^3 \times 160$	1.836	0.00060	0.1065860	0.1071460	0.1387529	1.6452

Table 6.3: The table shows the simulations parameters for the ensembles $cB211.64$ and $cC211.80$ which physical properties are listed in Tab. 6.1.

monomials one by one in the outer levels. For the N_f=1+1 section we use a rational approximation of order $n = 10$, which has a relative deviation such that $\left\| |Q_{ND}| R_{ND} \right\|_\infty < 1.4 \cdot 10^{-6}$, considering the eigenvalues of Q_{ND}^2 in the interval $\lambda_{min} = 6.5 \cdot 10^{-5}$ and $\lambda_{max} = 4.7$. The product of ratios in the rational approximation is split in four terms R_{1-4}, cf. Eq. (5.53). The first two contain three shifts, the second two contain two shifts. The four monomials R_{1-4} are placed one by one in the four outermost time-scales.

6.2.2 DD-αAMG speed-up in N_f=2+1+1 simulations

The force calculation in the N_f=2 sector is accelerated by a 3-level DD-αAMG approach as described in Sec. 6.1.2. Here we also use the DDalphaAMG method for the non-degenerate twisted mass operator. It is particularly helpful for the lowest terms of the rational approximation, as well as for the correction of rational approximation in the acceptance steps, where it yields a speed up to two over the standard multi-mass shifted conjugate gradient (MMS-CG) solver. As depicted in Fig. 6.3, with this setup we have relative small energy violation during the MD integration and an acceptance rate of 76.8%. We checked the size of reversibility violation yielding a standard deviation < 0.01 for $\delta \Delta H$ and $|1 - \langle \Delta H \rangle| < 0.02$ fulfilling the criteria discussed in Ref. [151]. Here, $\delta \Delta H$ is the difference of the Hamiltonian at integration time $t = 0$ and the Hamiltonian of the reversed integrated field variables after one trajectory is performed.

In Fig. 6.3 we also depict the iterations count for the inversions done with DD-αAMG. The setup update is done on the second time-scale for the shift $\mu + \rho_1$, since it is still close to the light quark mass. In this case, we find slightly larger fluctuations in the iteration count of the outer level of the multigrid solver compared to the $N_f = 2$ simulation shown in Fig. 6.2. Overall the simulation is stable and we find no correlation of the iteration count with larger energy violations δH.

The computational cost of each monomial per MD trajectory is depicted in Fig. 6.4. The previous approach depicted in the left panel involves the employment of multi-mass-shift CG (MMS-CG) for inverting all in once the shifted linear systems. In the right panel we have depicted the costs of the simulation accelerated by using the DD-αAMG solver for the most ill-conditioned linear systems. We achieve a speed-up for the $N_f = 1 + 1$ sector of a factor of 2 compare to a full eo-MMS-CG algorithm. The overall speed-up for $N_f = 2 + 1 + 1$ simulation at the physical light, strange and charm quark mass is a factor of 5.

As last, we report in Fig. 6.5 the time per HMC step for the $cA2.64$ and $cB211.64$ ensembles. The time per MDU is quite stable with fluctuations within the 10%. During the $N_f = 2 + 1 + 1$ simulation we calculate the smallest and largest eigenvalue of the non-degenerated TM operator at least every ten MDUs, which takes additionally 15 mins. This explains the single points that are frequently out of the main distribution. The two longer fluctuations instead are due to machine instabilities since they are limited to a single allocation of the job. Although the multigrid method is limiting significantly the parallelization of the calculation, the speed-up achieved makes feasible to sample enough MD
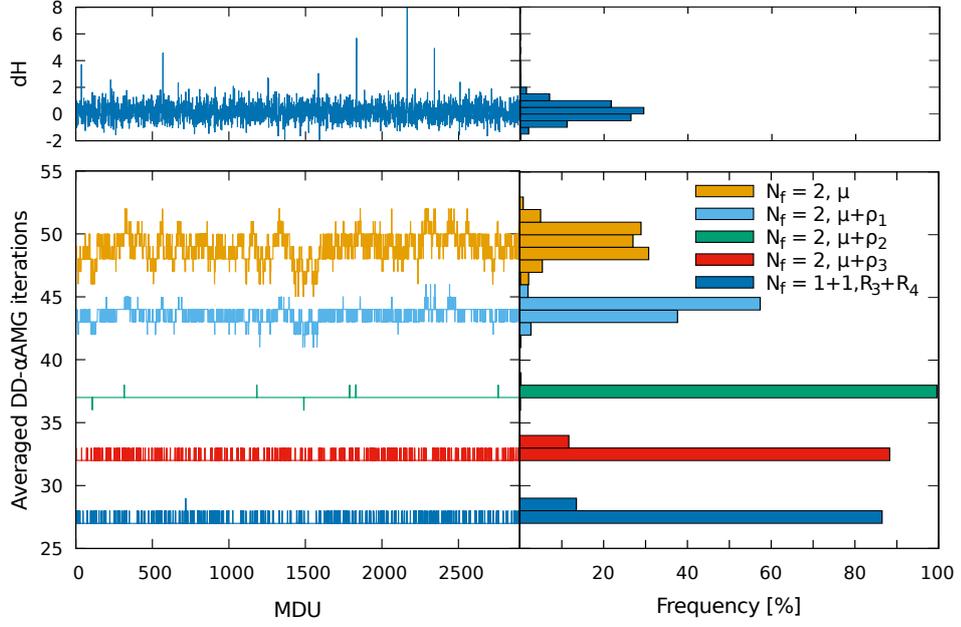
Figure 6.3: Top: The energy violation of the numerical integrator used during the MD integration of the $cB211.64$ ensemble plotted in units of MD trajectories (left) and its frequency (right). Bottom: DD-αAMG iterations count averaged on MDU considering different monomials of the HMC simulation shown in units of MD trajectories (left) and its frequency (right). The shifts ρ_i are $\{0.0003, 0.0012, 0.01\}$ and $\mu = 0.00072$. The monomials R_3 and R_4 contain the four smallest shifts of the rational approximation. The iterations count corresponds to average of the sum of the counts of the outer level during the calculation of the force in the MD.
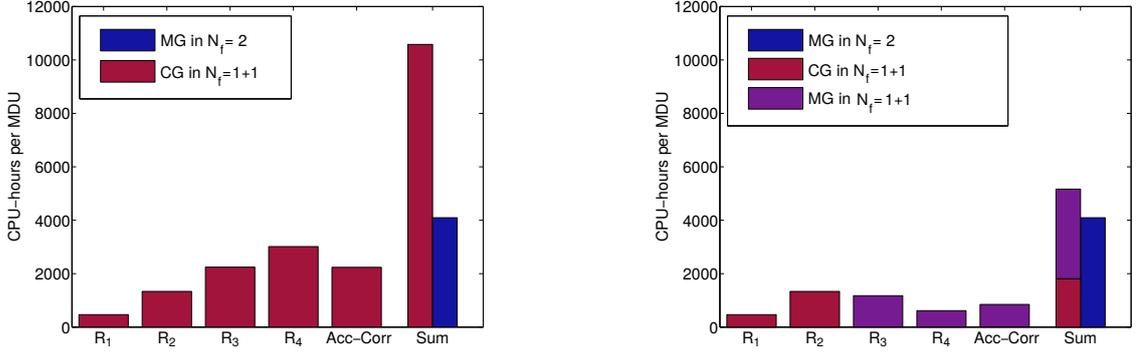
Figure 6.4: Costs per MDU of the $cB211.64$ simulation shown in units of CPU hours. In the left panel all the inversions in the RHMC are computed with a MMS-CG solver while in the right panel an hybrid approach of MMS-CG and DD-αAMG solvers is used. The last column, "Sum", reports the cost of RHMC in the $N_f = 1+1$ sector compared to HMC in $N_f = 2$ sector. The other columns regard different components of the RHMC: the "R_i"'s are the rational approximation split in different monomial – R_1 contains the largest shifts while R_4 the smallest – and "Acc-Corr" is the cost for the acceptance and correction step.

trajectory. The average time per MDU in the $N_f = 2$ simulation is slightly larger than an hour and in the $N_f = 2+1+1$ simulation below three hours. In both cases, 4096 cores employing 147 Haswell-nodes on SuperMUC are used since up to it we found that the three-level DD-αAMG method shows an almost ideal scaling up to this number of cores for a lattice with volume $V = 64^3 \times 128$.

6.3   Simulation procedure for the ensemble $cB211.64$

In this section we review the tuning of the simulation parameters for the production of the ensemble $cB211.64$. As Tab. 6.4 suggests, it has been a considerably expensive procedure involving the gen-
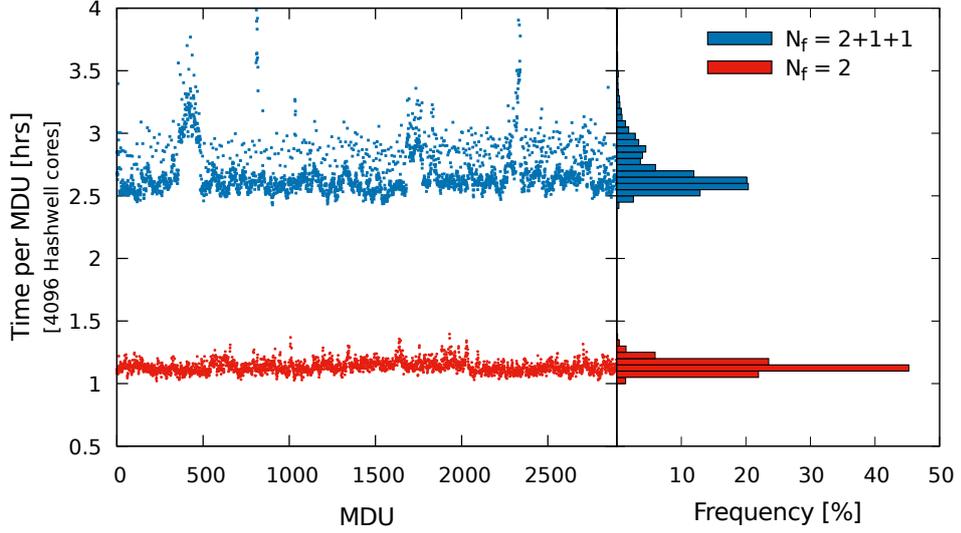
Figure 6.5: Time per MDU of the $N_f = 2$ (cA2.64) and $N_f = 2 + 1 + 1$ (cB211.64) ensembles in Table 6.1. The $N_f = 2 + 1 + 1$ timings also include the calculation of the smallest and largest eigenvalue done at least every ten trajectories. The large fluctuation around MDU 500 for the $N_f = 2 + 1 + 1$ ensemble is due to a bad allocation in the machine (it lasts exactly the duration of the allocation).

eration of other ten ensembles at smaller volume and different parameters for a precise estimation of the physical point.

Ensemble	L	μ	κ	N_{th}	$\bar{\mu}$	$\bar{\varepsilon}$
Th1.350.24.k1	24	0.0035	0.1394	755	0.1162	0.1223
Th1.350.24.k2	24	0.0035	0.13942	350	0.1162	0.1223
Th1.350.24.k3	24	0.0035	0.13945	351	0.1162	0.1223
Th1.350.24.k4	24	0.0035	0.13950	267	0.1162	0.1223
Th1.350.32.k1	32	0.0035	0.13940	88	0.1162	0.1223
Th1.200.32.k2	32	0.002	0.13942	430	0.1162	0.1223
Th2.200.32.k1	32	0.002	0.13940	178	0.1246864	0.1315052
Th2.200.32.k2	32	0.002	0.13942	439	0.1246864	0.1315052
Th2.200.32.k3	32	0.002	0.13944	392	0.1246864	0.1315052
Th2.125.32.k1	32	0.00125	0.139424	815	0.1246864	0.1315052
cB211.072.64.r1	64	0.00072	0.1394265	1647	0.1246864	0.1315052
cB211.072.64.r2	64	0.00072	0.1394265	1520	0.1246864	0.1315052

Table 6.4: Summary of the parameters of the ensembles used for the tuning and final runs: L is the lattice spatial size with the time direction taken to be 2L, μ is the twisted mass parameter of the mass degenerate light quarks, κ is the hopping parameter (common to all flavours), N_{th} are the number of thermalized trajectories in molecular dynamics units (MDU), $\bar{\mu}$ and $\bar{\varepsilon}$ are the bare twisted mass parameter of the mass non-degenerate fermion action used for the heavy quark sector. The ensembles cB211.072.64.r1 and cB211.072.64.r2 represent the targeted large volume runs at the physical point.

6.3.1 Tuning at maximal twist

Twisted mass fermions at maximal twist have $\mathcal{O}(a)$-improvement. This is the first condition to tune in a twisted mass fermions simulation. As discussed in Sec. 2.6.3, the tuning procedure to maximal twist requires a value of the hopping parameter $\kappa = \kappa_{\text{crit}}$ such that $m_{\text{PCAC}}(\kappa_{\text{crit}}) = 0$. Note that the value of m_{PCAC} has second order corrections in a^2 which depends on the twisted mass parameters μ, $\bar{\mu}$ and $\bar{\varepsilon}$. Thus determining k_{crit} at the value of interest for these parameters is important in order to

keep small the lattice artifacts which are introduced by the quark doublets [152, 153]. In practice, we allow some tolerance to the strict condition $m_{\text{PCAC}}(\kappa_{\text{crit}}) = 0$ and following Ref. [154] we require that

$$\frac{Z_A m_{\text{PCAC}}}{\mu} 0.1 \tag{6.4}$$

within errors where Z_A is the renormalization constant of the axial current. Fulfilling the condition in Eq. (6.4) is numerically consistent with $\mathcal{O}(a)$-improvement of physical observables, where it entails only an error of order $\mathcal{O}((Z_A \cdot m_{\text{PCAC}}/\mu)^2)$. Therefore for the targeted lattice spacing the error is comparable to other $\mathcal{O}([a\Lambda_{QCD}]^2)$ discretization errors. This allows an $\mathcal{O}(a^2)$ scaling of physical observables towards the continuum limit.
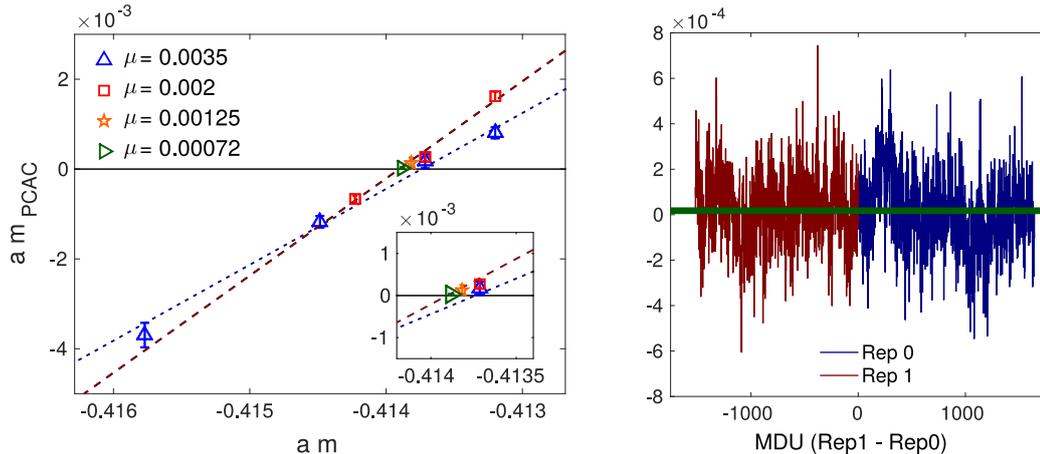
Figure 6.6: *Left:* The PCAC mass versus the bare light quark mass $am = 1/(2\kappa) - 4$ for various values of the twisted mass parameter for the tuning of the critical mass. The linear interpolations are done on the Th1.350.24 ensembles (blue, triangle points) illustrated with the blues dashed line and on the Th2.200.32 ensembles (red square points) illustrated with the red dashed line. The value of the PCAC mass for the Th2.125.32.k1 ensemble is shown by the orange star point and of the cB211.072.64 by the green right pointing triangle.
Right: The MC History of the PCAC quark mass on the large volume physical point ensembles is shown at twisted mass value $\mu = 0.00072$ and hopping parameter $\kappa = 0.1394265$.

In order to find κ_{crit}, we have generated several ensembles with fixed volumes of size $24^3 \cdot 48$ and $32^3 \cdot 64$, as listed in Table 6.4 and depicted in Fig. 6.6. At fixed twisted mass parameter μ we first scan over several values of the hopping parameter κ and we find a range close to zero (blue fit in the figure). Then we move to a smaller value of μ and scan again (red range). The dependence of the PCAC quark mass on κ at different twisted mass parameter μ differs visibly as shown by the slope of the two fits. Using simple linear fits for the $L = 32$ ensembles, we determine a critical value of κ, $\kappa_{\text{crit}} = 0.1394265$ and we employ this κ-value for our large volume ensembles.

In the $64^3 \cdot 128$ lattices we first thermalize the ensemble using 500 trajectories. We then take the latest configuration as a starting point for two replicas, each having a final statistics of about 1500 MDUs. This allows us to run in parallel two simulations and to combine them back-to-back in a unique ensemble, $cB211.64$. In Fig. 6.6 we depict the Monte Carlo history of the PCAC quark mass for these two replicas, where we separate the two replicas adding the second replica (rep 1) with a reversed history starting from zero. The PCAC quark mass fluctuates around zero and does not show particularly large autocorrelation times. Performing the average over the two replica runs, we find $Z_A m_{\text{PCAC}}/\mu = 0.02(2)$ where we used $Z_A = 0.763(2)$ and $m_{\text{PCAC}} = 0.189(114)10^{-4}$. Thus, the condition of eq. (6.4) is nicely fulfilled. We therefore conclude that the tuning to maximal twist is achieved for the c211.64 ensemble.

6.3.2 Tuning of the heavy quark sector

After fixing κ_{crit} we proceed by tuning the heavy twisted mass parameters to realize physical kaon and D_s-meson mass and decay constant. Such tuning is not a trivial task due to the mixing in flavor space introduced by the non-degenerate twisted mass operator, cf. Sec. 2.6.5. In order to tackle the problem, it is convenient to employ in an intermediate step the so-called Osterwalder Seiler (OS) fermions [155] as valence quarks avoiding the mixing effects. The OS-fermions can be used in a well defined mixed action setup having the same critical mass, m_{crit}, as determined in the maximal twist tuning [83]. In this case, strange and charm quarks are introduced as valence quark with the twisted mass operators $D(\mu_s)$ and $D(\mu_c)$ respectively, while they are simulated in the sea with the non-degenerate TM operator $D_{\text{ND}}(\bar{\mu}, \bar{\varepsilon})$. The renormalized valence quark masses for OS fermions are $m_{c,s}^{(r)} = \mu_{c,s}/Z_P$. The parameters $\mu_{c,s}$ are matched to the corresponding non-degenerate twisted mass parameters, cf. Eq. (2.113), via

$$\mu_{c,s} = \left(\bar{\mu} \pm \frac{Z_P}{Z_S} \bar{\varepsilon} \right) \tag{6.5}$$

with Z_P and Z_S denoting the non-singlet pseudoscalar and scalar renormalization constants. The correlation functions using OS or non-degenerate TM fermions are equivalent in the continuum limit. Moreover they still yield $\mathcal{O}(a)$-improved physical observables.

The parameter tuning procedure for the heavy sector can be carried out on a non-large lattice (in the present case, $32^3 \cdot 64$) and at larger than physical up/down quark mass. With a convenient choice of the physical renormalization conditions, we use for the tuning

$$C_1 \equiv \frac{\mu_c}{\mu_s} = 11.8 \qquad \text{and} \qquad C_2 \equiv \frac{m_{D_s}}{f_{D_s}} = 7.9 \tag{6.6}$$

where m_{D_s} is the D_s-meson mass and f_{D_s} is its decay constant. The condition C_2 has a strong sensitivity to the charm quark mass while C_1 fixes the strange-to-charm mass ratio. They show only small dependence on the light quark mass arising from sea quark effects. We expect these conditions to be essentially free from finite-size effects due to the heavy D_s-meson mass.
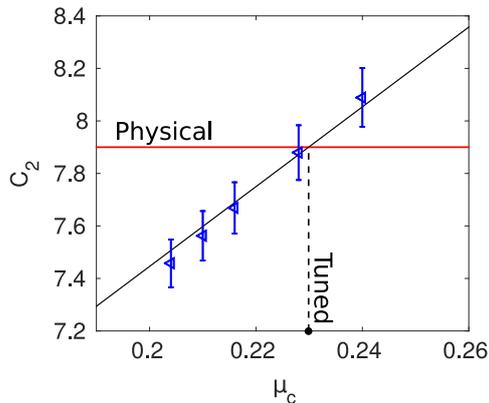
Figure 6.7: Tuning of the charm quark twisted mass parameter μ_c using the C_2 condition on the Th1.200.32.k2 ensemble. The figure shows 5 measured values of the m_{D_s}/f_{D_s} ratios with blue square points, using 5 different pairs of μ_c and μ_s which ratio is fixed by the C_1 condition. The horizontal line illustrates the physical value of C_2 and the black dashed line the tuned values for μ_c.

As first step, we use the ensemble Th1.200.32.k2 in Tab. 6.4 which has a μ around three times larger than the physical up-down quark mass and it is generated with guessed values of $\bar{\mu}$ and $\bar{\varepsilon}$. We chose the OS bare quark masses μ_c and μ_s such that the condition C_1 is always fulfilled. As depicted in Fig. 6.7, we vary the OS quark masses over a broad range such that we find the value where also the condition C_2 is satisfied. When matching the non-degenerate twisted mass parameters via Eq. (6.5),

we obtain directly the value of $\bar{\mu}$ from the tuned μ_s and μ_c, while $\bar{\varepsilon}$ requires the ratio Z_P/Z_S. The latter can be estimated by adjusting $\bar{\varepsilon}$ such that the kaon mass evaluated in the twisted mass formulation (m_K^{tm}) is the same as its counterpart computed with valence OS fermions (m_K^{OS}). Although the kaon mass value can be unphysical due to the larger value of μ, we use this matching condition only for relating heavy quark action parameters. In such way, from the values tuned in Fig. 6.7,

$$\mu_s = 0.01948 \quad \text{and} \quad \mu_c = 0.2299, \tag{6.7}$$

we find

$$\bar{\mu} = 0.12469 \quad \text{and} \quad \bar{\varepsilon} = 0.13151. \tag{6.8}$$

Since the matching steps described so far are performed on an unphysical gauge ensembles, we still need to check these conditions on a new gauge ensemble generated with the tuned values of $\bar{\mu}$ and $\bar{\varepsilon}$. Therefore, several other ensembles with volumes of $32^3 \times 64$ at light twisted mass values of $\mu = 0.002$ and $\mu = 0.00125$ have been generated to check again these conditions and to determine precisely the critical hopping parameter $\kappa_{crit} = 0.1394265$. In this procedure turned out that the previous tuning of the heavy quark sector was already good enough and we did not adjust again the parameters $\bar{\mu}$ and $\bar{\varepsilon}$.

6.3.3 Heavy meson sector at the physical point

Repeating the tuning procedure on the physical ensemble cB211.64 we find

$$\mu_s = 0.01892(13) \quad \text{and} \quad \mu_c = 0.2233(16). \tag{6.9}$$

The OS valence quark parameters are lower by around 2.4% compared to the values determined using the Th1.200.32.k2 ensemble, cf. Eq. (6.7). If we impose an exact matching between m_K^{OS} and the unitary m_K^{tm} ensembles we find the ratio of the pseudoscalar to the scalar renormalization constants to be

$$\frac{Z_P}{Z_S} = 0.813(1). \tag{6.10}$$

Using this value of Z_P/Z_S and the OS quark masses in Eq. (6.9) we obtain

$$\bar{\mu}_{phys} = 0.1211(12) \quad \text{and} \quad \bar{\varepsilon}_{phys} = 0.126(13). \tag{6.11}$$

Thus the actually employed sea quark mass parameters correspond to a *sea* strange and charm quark mass 3% heavier than those derived *a posteriori* from imposing the same tuning and matching conditions on the physical point ensembles.

In Fig. 6.8 we show the decay constants of the kaon and the D-meson and compare them with the results extracted from the $N_f = 2$ clover ensembles [76]. We employ 244 measurements on the cB211.64 and 100 on the Th2.200.32.k1 ensemble. The ratios of the kaon and D_s-meson masses to decay constants for the cB211.64 ensembles are found to be

$$\frac{m_K}{f_K} = 3.188(7) \quad \text{and} \quad \frac{m_D}{f_D} = 8.88(11), \tag{6.12}$$

where the former ratio has a central value slightly larger than the physical ratio $m_K^{phys}/f_K^{phys} = 3.162(18)$ [156], while the latter agrees well within errors with the value $m_D^{phys}/f_D^{phys} = 9.11(22)$ [17]. These results in-
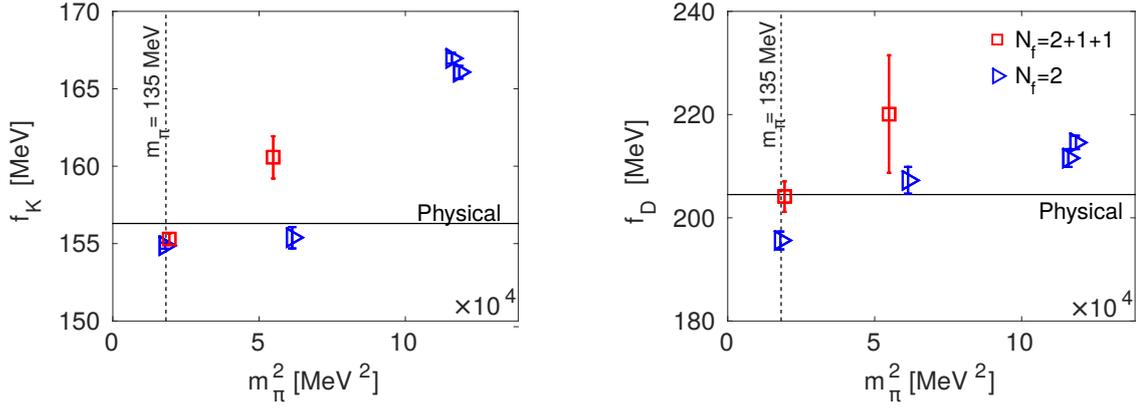
Figure 6.8: The pseudoscalar decay constants in the heavy quark sector. The left panel shows the kaon decay constant, while the right panel shows the D-meson decay constant both versus the squared pion mass. The dashed vertical line indicates the physical value of the pion mass. The red squares are the measurements for the Th2.200.32.k1 and cB211.64 ensembles, while the blue triangles are for the $N_f = 2$ clover twisted mass ensembles [76]. The scale is set via the pion decay constant.

dicate that discretization effects for our setup are small in the heavy quark sector. For a more rigorous check, a direct calculation at different values of the lattice spacing needs to be carried out.

6.3.4 Light meson sector at the physical point

The main goal of this section is to determine the value of the lattice spacing within the pion sector. In principle, the lattice spacing could be determined directly from the cB211.64 target ensembles requiring a vanishing pion mass in the chiral limit. The lattice spacing and the physical twisted mass value can be fixed by assuming a linear dependence of μ on $a^2 m_\pi^2$ and m_π^2/f_π^2. The so determined lattice spacing is

$$a = 0.0801(2) \, \text{fm}. \tag{6.13}$$

However, it is helpful to also use the other available ensembles, listed in Table 6.4, which are all tuned to maximal twist, namely Th1.350.24.k2, Th2.200.32.k2, Th2.150.32.k2 in addition to cB211.64. By employing chiral perturbation theory (χPT) to describe the quark mass dependence of the pion decay constant and pion mass, we obtain a robust result for the value of the lattice spacing. We depict in Fig. 6.9 the ratio m_π^2/f_π^2 and the pion decay constant itself as function of the light bare twisted quark mass.

In the figure we show the fits to NLO χPT [159–161], which for the ratio m_π^2/f_π^2 reads as

$$\frac{m_\pi^2}{f_\pi^2} = 16\pi^2 \xi \left(1 + P\xi + 5\xi \log(\xi)\right) \frac{F_{f_\pi}^{FVE\,2}}{F_{m_\pi}^{FVE\,2}} \tag{6.14}$$

and for the pion decay constant is

$$af_\pi = af_0 \left(1 + R\xi - 2\xi \log(\xi)\right) 1/F_{f_\pi}^{FVE}. \tag{6.15}$$

Since the ensembles out of the physical point have a relatively small volume, we have included finite volume corrections, $F_{f_\pi}^{FVE}, F_{m_\pi}^{FVE}$, from chiral perturbation theory [157, 158]. The variable ξ is given
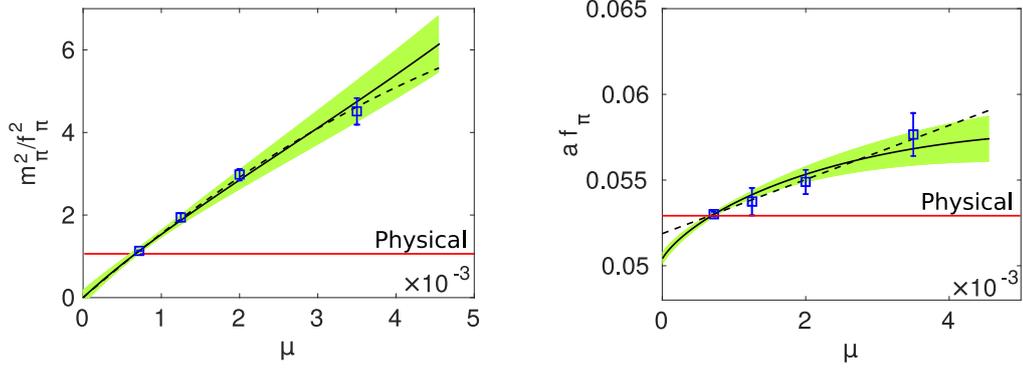
Figure 6.9: *Left:* The ratio m_π^2/f_π^2 is plotted against μ. *Right:* The pion decay constant af_π is plotted against the light twisted mass math parameter μ.

The solid lines are fits to NLO chiral perturbation theory with the error as shaded band, see eq. (6.14) and eq. (6.15). The dotted lines are fits for which the chiral logs are neglected. The pion mass and decay constant are corrected with finite volume correction terms on the left from Ref. [157] and on the left from Ref. [158].

by

$$\xi = \frac{2B_0}{Z_P(4\pi f_0)^2}\mu \tag{6.16}$$

where B_0 and f_0 are low energy constants. The fitting constants P, R are related to the NLO low energy constants by

$$P = -\bar{l}_3 - 4\bar{l}_4 - 5\log\left(\frac{m_\pi^{phys}}{4\pi f_0}\right)^2 \quad \text{and} \quad R = 2\bar{l}_4 + 2\log\left(\frac{m_\pi^{phys}}{4\pi f_0}\right)^2. \tag{6.17}$$

From the fits, we determine the values of $2B_0/Z_P = 4.52(6)$ and $af_0 = 0.0502(3)$ and we determine the finite volume correction terms by fixing the low energy constants using the results of Ref. [153]. For our target ensemble cB211.64 with $m_\pi L = 3.62$ we find that the finite volume effects yield corrections of less than 0.5% for the pion mass and less than 0.5% for the pion decay constant. By using the fit functions from χPT and fixing the ratio $m_{\pi,phys}^2/f_{\pi,phys}^2 \equiv 1.034$ we find a more correct light twisted mass parameter

$$\mu_{phys} = 0.00067(1), \tag{6.18}$$

which is 7% smaller than the one used. We then use this value in Eq. (6.15) to determine the lattice spacing. We get

$$a_{f_\pi} = 0.07986(15)(35) \text{ fm}, \tag{6.19}$$

with the first error the statistical and the second the systematic by using the physical value of the pion decay constant, $f_{\pi,phys} = 130.41(20)$ MeV [17]. We follow the procedure adopted in Ref. [154] for determining a systematic error by performing several different fits, adding or neglecting finite volume terms. Such fits employ the finite volume corrections of Ref. [158] using the calculated low energy constant c_2 of Eq. (6.23) different orders in chiral perturbation theory and including or excluding the ensemble Th2.150.32.k2 due to larger finite size effects. The systematic error is then given by the deviations of these different fits from the central value given in Eq. (6.19). Although we include ensembles like Th2.150.32.k2 or Th1.350.24.k2 which have large finite size effect up to 8% in the pion decay constant, the systematic uncertainties are suppressed due to the fact that we are using ensembles close to physical quark masses which stabilize the fits. Thus this demonstrates the

importance of working at physical quark masses. Moreover this is confirmed by the estimation of the lattice spacing which takes only the pion mass and decay constant from cB211.64 into account, cf. Eq. (6.13), which compatible with the final value in Eq. 6.19 within errors.

6.3.5 $\mathscr{O}(a^2)$ isospin-breaking lattice artifacts in the pion sector

An important aspect when working with twisted mass fermions at maximal twist is to keep the size of isospin violations small, cf. Sec. 2.6.1. This isospin breaking manifests itself by the fact that the neutral pion mass becomes lighter than the one of the charged pion. In leading order (LO) of chiral perturbation theory this effect is described by

$$a^2(m_\pi^2 - m_{\pi^0}^2) = -4c_2 a^2 \sin^2(\omega) \qquad (6.20)$$

with the twisted mass angle given by $\omega = \text{atan}(\mu/Z_A m_{\text{PCAC}})$ and c_2 a low energy constant characterizing the strength of $\mathscr{O}(a^2)$-effects of twisted mass fermions. As shown in Ref. [74, 76], a clover term reduces the value of the low energy constant c_2. Indeed, only employing a clover term, simulations at physical quark masses at large enough lattice spacing become possible as demonstrated in Ref. [76].
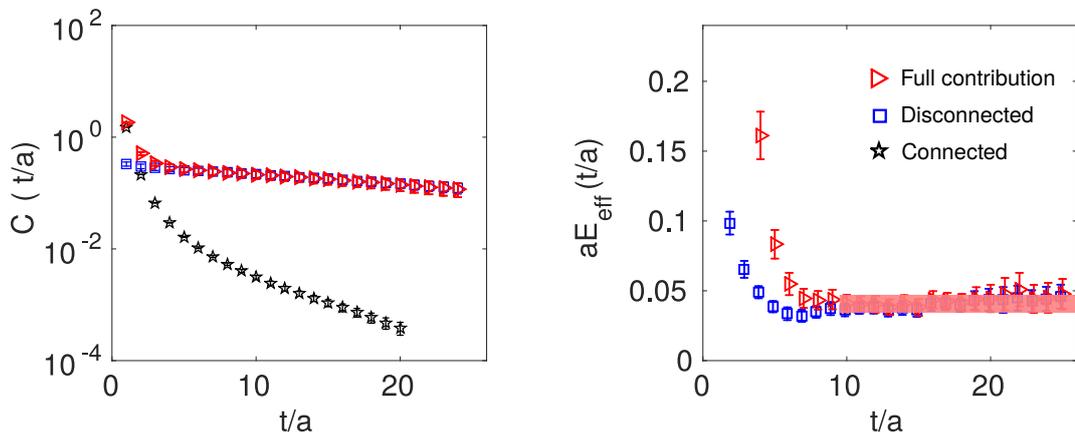
Figure 6.10: *Left:* The correlator of the neutral pion versus t/a. *Right:* The effective mass of the neutral pion. The shaded band shows the constant fit in the plateau range. The red triangle shows the data of the full correlator, while the blue squares the disconnected and the the black stars the connected contribution.

In order to calculate the neutral pion mass one needs to compute disconnected two-point functions that are notoriously noisy. To suppress the noise in the computation of the two-point functions we use a combination of exact deflation, projecting out the 200 lowest lying eigenvalues, and 6144 stochastic volume sources corresponding to an eight-distance hierarchical probing [162, 163]. The disconnected correlator needed is given by

$$C_{disc}(t_0) = \langle \hat{O}(0)\hat{O}(t_0)\rangle \quad \text{with} \quad \hat{O}(t_0) = D^{-1}(t_0,t_0) - \langle D^{-1}(t,t)\rangle \qquad (6.21)$$

where the ensemble and time average of the vacuum contribution is subtracted from the disconnected operator. Note that we used global volume noise sources to extract the disconnected contribution, however methods which do not subtract the vacuum expectation value explicitly could be more effective as pointed out in [9, 76, 164]. We find that the disconnected contribution dominates the correlator for time distances $t/a > 10$, as can be seen in Fig. 6.10. However we include the connected contribu-

tion in the plateau average, leading to a neutral pion mass given by

$$am_{\pi^{(0)}} = 0.044(9) \ . \tag{6.22}$$

Note, that for the connected contribution small statistics of around 250 measurements is used, which results in a relatively large statistical error. The charged pion mass is straight forward to compute and we find for the charged pion mass $am_\pi = 0.05658(6)$. This gives an isospin splitting in the pion mass of $22(16)\%$ and the low energy constant c_2 of Eq. (6.20) reads as

$$4c_2 a^2 = -0.0013(8) \tag{6.23}$$

assuming $\omega = \pi/2$. Thus, introducing a clover term for $N_f = 2+1+1$ twisted mass fermions suppresses isospin breaking effects effectively, i.e. by a factor of 6 compared to an $N_f = 2+1+1$ ensembles with twisted mass fermions without a clover term and a pion mass of 260 MeV at a similar lattice spacing of $a = 0.078(1)$ fm [73, 165], where it was found that the mass splitting is given by $(am_{\pi^{(0)}})^2 - (am_\pi)^2 = -0.0077(4)$. The suppression of the pion isospin breaking effects, thanks to the use of the clover term, is the underlying reason why we can perform our simulations at the physical point with $N_f = 2+1+1$ flavours of quarks.

6.3.6 Baryon sector at the physical point

As another test, whether we are in the desired physical condition, we analyze the nucleon mass which can also provide an independent determination of the lattice spacing, which can be compared to the one found in the meson sector. We measured the nucleon mass on the cB211.64 ensemble by using interpolating fields containing the operator

$$J_p = \varepsilon_{abc} \left(u_a^T C \gamma_5 d_b \right) u_c \,, \tag{6.24}$$

with $C = \gamma_4 \gamma_2$ the charge conjugation matrix. We then constructed the two point correlation function

$$C_p(t) = \frac{1}{2} \text{Tr}(1 \pm \gamma_4) \sum_{\mathbf{x}} \langle J_p(\mathbf{x}, t) \bar{J}_p(0, 0) \rangle \tag{6.25}$$

which provides the nucleon mass in the large time limit. We used 50 APE smearing steps with $\alpha_{APE} = 0.5$ [166] in combination with 125 Gaussian smearing steps with $\alpha_{gauss} = 0.2$ [167, 168] to enhance the overlap of the used point sources with the lowest state.

We extracted the nucleon mass for $t \gg 0$ by a plateau average over the effective mass $aE_{eff} = \log(C_p(t+a)/C_p(t))$ shown in Fig. 6.11. The plateau average of the nucleon mass, given by $am_N = 0.3864(9)$ on the cB211.64 ensemble, which is in agreement with a two-state fit having $am_{N,2st} = 0.3850(12)$.

As an alternative way to determine the lattice spacing, one can use the nucleon mass. A direct way would be to use the physical ratio from which, by using the pion mass determined above, the lattice spacing can be estimated directly by the value of the lattice nucleon mass. Indeed, with the the pion mass $am_\pi = 0.05658(6)$ the nucleon to pion mass ratio $0.3864(9)/0.05658(6) = 6.83(2)$ is close to its physical value of $m_N^{\text{phys}}/m_\pi^{\text{phys}} = 0.9389/0.1348 = 6.965$ where we take the average of neutron and proton mass [17] and the pion mass in the isospin symmetric limit [156]. However, as in the case of the meson sector, using more data points at heavier pion masses and χPT to describe their
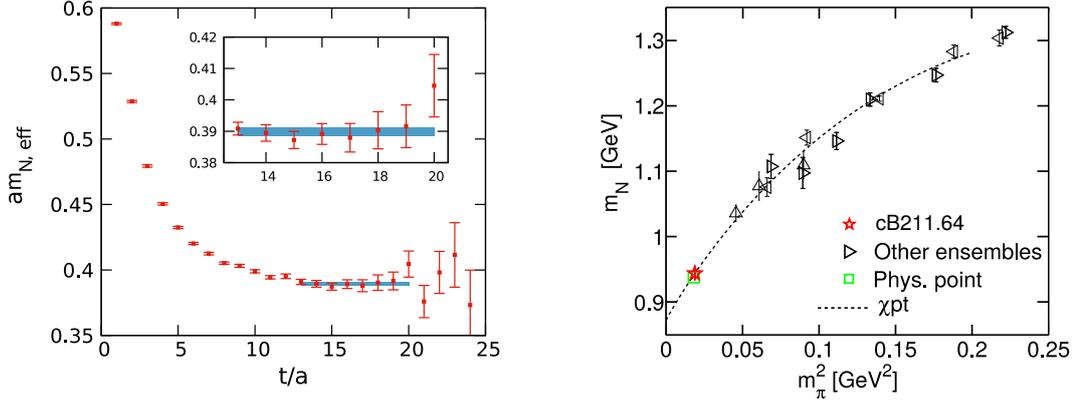
Figure 6.11: *Left:* The time dependence of the effective mass extracted from the nucleon correlator is shown. The blue band shows the fitted plateau average.
Right: The squared pion mass dependence of the nucleon mass is shown by comparing the nucleon mass from our target lattice cB211.64 to the values determined on the $N_f = 2+1+1$ ETMC ensembles. The dotted line shows the fit by employing chiral perturbation theory at $\mathcal{O}(p^3)$.

quark mass dependence, a more robust result can be obtained. More concretely, we have employed chiral perturbation theory at $\mathcal{O}(p^3)$ [169, 170] for the nucleon mass dependence on the pion mass, i.e.

$$m_N = m_N^{phys} - 4b_1 m_\pi^2 - \frac{3g_A^2}{16\pi f_\pi^2} m_\pi^3 \ . \tag{6.26}$$

Similarly to Ref. [43] we use the nucleon masses of the $N_f = 2+1+1$ ETMC ensembles without a clover term, determined in Ref. [40] to perform the chiral fit of Eq. (6.26). In our analysis we neglect cutoff effects, which appear to be small and not visible within our statistical errors. The same holds for finite volume effects, see Ref.[43]. We fixed $f_\pi = 0.1304(2)$ GeV and $g_A = 1.2723(23)$ [17] in Eq. (6.26). The resulting fit to Eq. (6.26) is shown in the right panel of Fig. 6.11 and allows to determine the lattice spacing as

$$a_{m_N}(\beta = 1.778) = 0.08087(20)(37) \text{ fm} \ . \tag{6.27}$$

The first error is statistical while the second error is the deviation between the estimate obtained from Eq. (6.26) and taken the mass from the two-state fit. Note that the statistical error in the nucleon mass is comparable with pion sector due to three orders of magnitude larger number of inversions.

6.3.7 $\mathcal{O}(a^2)$ isospin splitting in the baryon sector

The finite twisted mass value can result into a mass splitting of hadrons which are symmetric under the isospin symmetry of the light flavor doublet. As pointed out in sec. 6.3.1 this indeed leads to a sizable effect in the neutral-charged pion mass splitting. Here, we want to discuss the splitting in the baryon sector in case of the Δ-baryon employing the cB211.64 ensemble. Note that for the used lattice size the lowest decay channel of the Delta baryon, which is a nucleon+pion state with correct parity, is heavier than the Delta baryon itself. Thus, for the simulations performed here, the Delta can be treated as a stable state.
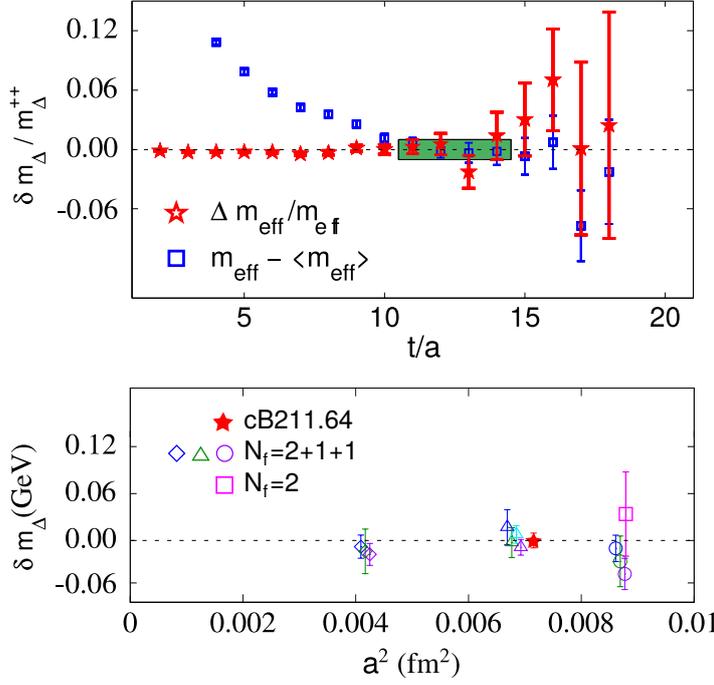
Figure 6.12: Relative differences of the effective Delta baryon masses.

Top: The relative difference given by $\delta m_{\Delta,eff}(t)/m_{\Delta^+}$ (blue squares). To illustrate the beginning of the plateau we added the relative effective mass $m_{eff}(t) - m_{eff}^{plateau}$ with $m_{eff}^{plateau}$ is the plateau value m_{Δ^+} (blue squares).

Bottom: Isospin breaking for the Delta baryons as a function of the lattice spacing squared for non-clover improved N_f=2+1+1 ensembles, N_f=2 clover improved ensemble and cB211.64 (red star). The value of cB211.64 is the plateau extracted in the top panel; the other values have been computed in Ref. [40, 43].

We measured the Δ-baryon correlator by using the following interpolating fields

$$J_{\Delta^+}^\mu = \frac{1}{\sqrt{3}}\varepsilon_{abc}\left[2\left(u_a^T C\gamma^\mu d_b\right)u_c + \left(u_a^T C\gamma^\mu u_b\right)d_c\right], \tag{6.28}$$

$$J_{\Delta^{++}}^\mu = \varepsilon_{abc}\left(u_a^T C\gamma^\mu u_b\right)u_c. \tag{6.29}$$

Note that $J_{\Delta^+}^\mu$ and $J_{\Delta^{++}}^\mu$ is symmetric under $u \to d$ to $J_{\Delta^0}^\mu$ and $J_{\Delta^-}^\mu$ respectively. We neglect the potential mixing of Δ with the spin-1/2 component which is suppressed [171]. Thus the correlators for the Δ^{++} is given by $C_\Delta = \text{Tr}[C]/3$ with $C_{ij} = \text{Tr}[(1+\gamma_4)/2\langle J_{\Delta^{++}}^i(t)\bar{J}_{\Delta^{++}}^j(0)\rangle]$ and gives an average value of $am_\Delta = 0.5251(72)$ by using a plateau average over the effective mass. Now we define the splitting in the mass by

$$\delta m_{\Delta,eff} = \log\left\{\frac{C_R(t)}{C_R(t+a)}\right\} \quad \text{with} \quad C_R = \frac{C_{\Delta^+}(t) + C_{\Delta^0}(t)}{C_{\Delta^{++}}(t) + C_{\Delta^-}(t)} \tag{6.30}$$

where we average over the symmetric parts. In Fig. 6.12 we show the effective relative mass splitting given by $\delta m_{\Delta,eff}/m_{\Delta^+}$. In addition we plot the relative effective mass $m_{eff}(t)$ of the Δ^+ particle subtracted from its plateau average to illustrate where the plateau of the Δ-baryon starts. We find that the relative splitting in the Δ mass is $\delta m_\Delta/m_{\Delta^+} = 0.0098(65)$ and hence close to zero within errors. This result is in agreement with Ref. [76] where it was found that the isospin splitting of the twisted mass action in the baryon section is suppressed.

6.3.8 Final determination of the lattice spacing

As we saw, the lattice spacing can be evaluated by matching lattice observables to their physical counterparts. This has been done, as described in sections 6.3.4 and 6.3.6, in the meson sector by employing the pion decay constant and in the baryonic sector using the nucleon mass, respectively. Differences in the values obtained for the lattice spacing as determined using different physical observables can shed light on cut-off effects. We discuss here an additional method to determine the lattice spacing, which is provided by the gradient flow scale setting parameters t_0 from Ref. [172] and w_0 from Ref. [173]. Following the procedure described in these articles and in
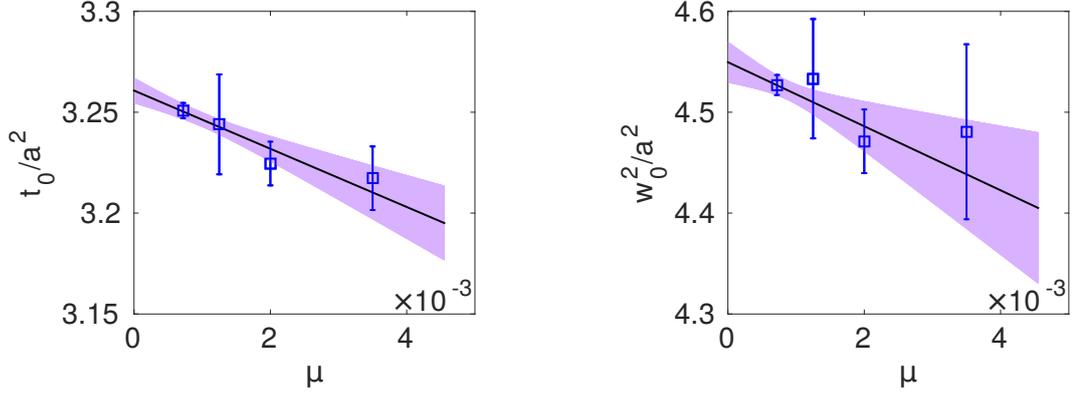
Figure 6.13: *Left:* Linear extrapolation of the gradient flow observable t_0/a^2. *Right:* Linear extrapolation of the gradient flow observable w_0^2/a^2. The solid line with the shaded violett band shows the linear extrapolation.

particular as applied to the twisted mass setup in Ref. [76], we extrapolate the gradient flow observables to the chiral limit using a fit ansatz linear in μ, which corresponds to LO χpt [174]. We employ the values computed for the ensembles Th1.350.24.k2, Th2.200.32.k2, Th2.125.32.k1, cB211.64. The resulting curves are shown in Fig. 6.13. From the fit we find we extract at $\mu = 0$ the values $t_0^{\mathrm{ch}}/a^2 = 3.261(6)$ and $w_0^{\mathrm{ch}\,2}/a^2 = 4.550(20)$. Using the phenomenological values of $\sqrt{t_0} = 0.1465(25)$ and $w_0 = 0.1755(18)$ [173] we deduce the following values for the lattice spacing

$$a_{t_0} = 0.0811(14) \text{ fm} \qquad \text{and} \qquad a_{w_0} = 0.0823(8) \text{ fm} . \qquad (6.31)$$

phys. quant.	lat. spac. [fm]	quantities in lat. units	
a_{t_0}	0.0811(14)	$t_0/a^2\big|_{\mu_\ell=0.00072} =$	3.246(7)
a_{w_0}	0.0823(8)	$w_0^2/a^2\big|_{\mu_\ell=0.00072} =$	4.512(16)
a_{f_π}	0.07986(38)	$af_\pi\big|_{\mu_\ell=0.00072} =$	0.05272(10)
a_{m_N}	0.08087(44)	$m_N/m_\pi\big|_{\mu_\ell=0.00072} =$	6.829(19)
average	0.08029(41)		

Table 6.5: We give the values of the lattice spacing determined by using different physical quantities as inputs, including in the errors the input systematic uncertainties. The final value of the lattice spacing is derived via a weighted average of a_{f_π} and a_{m_N} where for the final error a 100% correlated data is assumed [175]. The residual systematic uncertainty on the lattice spacing, which stems from higher order cutoff effects, should be of relative size $\mathcal{O}(a^2)$ and looks numerically smaller than 2%.

In Table 6.5 we summarize the values of the lattice spacing as determined from the pion mass and decay constant, the nucleon mass and the gradient flow parameters t_0 and w_0. As it can be noticed, there are small deviations of the lattice spacing between the meson and the baryons sector and in any case they are comparable to the one we have observed in the simulations with $N_f = 2$ flavours of quarks. That indicates that cutoff effects do not increase for our $N_f = 2+1+1$ flavour setup used here.

6.3.9 Autocorrelation

The autocorrelation of the HMC algorithm, cf. Sec. 5.1.1 slows down critically for very fine lattice spacings with $a < 0.05$ fm. This can be seen in the freezing of the topological charge [123]. For our lattice with $a \sim 0.08$ fm we found that the topological charge can fluctuate between the different sectors leading to small autocorrelation times of $\tau_{int}(Q) = 13(5)$ [MDU]. As pointed out in [176] the energy density at finite flow times develops larger autocorrelation times in the regime with $a \gtrsim 0.05$ fm. Although we have relative small statistics we calculated integrated autocorrelation time for the plaquette and the gradient flow observables t_0/a^2 as shown in figure 6.14 by using the Γ-method [177]. We find a pion mass dependence given by

$$\tau_{int}(m_\pi) = A \frac{1}{m_\pi^b} \tag{6.32}$$

with $b = 2.2(5)$ case of the plaquette while $b = 2.0(7)$ in case of the gradient flow observable t_0. A possible explanation for the quark mass dependence of the autocorrelation time τ_{int} is a phase transition in case of finite twisted mass term for vanishing neutral pion masses. Although the isospin splitting is suppressed in our case, observables like the gradient flow observables shows an increase with inverse of the squared pion mass. This behavior is also seen in the PCAC mass, where moderate integrated autocorrelation times were found which can be clearly seen in the Monte Carlo history (see right panel of Fig. 6.6). However in other quantities like the pseudoscalar mass, the pseudoscalar decay constant or nucleon observables τ_{int} is very small and a quark mass dependence can be not observed.
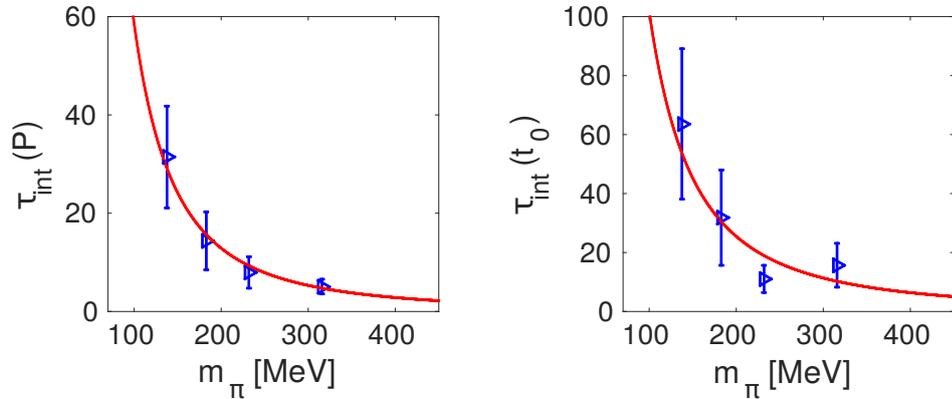
Figure 6.14: The figure shows the integrated autocorrelation τ_{int} of the ensembles at maximal twist (from right to left Th1.350.24.k1, Th2.200.32.k2, Th2.125.32.k1, cB211.072.64 for both cases). Note that we used here the ensemble Th1.350.24.k1 instead of Th1.350.24.k2 which not at maximal twist but has a larger statistics.

6.3.10 Important remarks

We have presented here our first successful simulation of maximally twisted mass fermions with $N_f = 2+1+1$ quark flavours at the physical values of the pion, the kaon and the D-meson masses. By having a lattice spacing of $a = 0.08029(41)$ fm, we find that the simulation is stable also when performed with physical values of the quark mass parameters. In particular, we have been able to carry out a demanding but smooth tuning procedure to maximal twist and to find the values of the light, strange and charm bare quark masses, which correspond to the physical ones for the first two quark generations.

In our setup, which employs a clover term, the cutoff effects appear to be small. Several observations corroborate this conclusion:

- as already mentioned above, the simulations themselves are very stable;

- as computed in the Secs. 6.3.3 and 6.3.4 and here summarized in Tab. 6.6 our tuned parameters are very close to the physical point with deviations of few percent;

- when fixing the quark mass parameters through the selected physical observables, other physical quantities, as collected in Table 6.7 come out to be consistent with their physical counterparts;

- the $\mathcal{O}(a^2)$ effects originating from the isospin breaking of twisted mass fermions are small and significantly reduced compared to our earlier simulations without clover term with $N_f = 2+1+1$ flavors at non-physical pion masses;

- deviations of the lattice spacing from the meson sector, the baryon sector and gradient flow observables, as listed in Table 6.5, are small and of the same size as in our former $N_f = 2$ flavor simulations;

	β	c_{SW}	μ	$\bar{\mu}$	$\bar{\varepsilon}$	κ_{crit}
cB211.64	1.778	1.69	0.00072	0.1246864	0.1315052	0.1394265
physical point a posteriori			0.00067(1)	0.1211(12)	0.126(13)	T.B.D.

Table 6.6: The table shows the simulations parameters for the ensembles cB211.64 compared to the values determine *at posteriori* on this ensemble for being exactly at physical point with $\beta = 1.778$ and $c_{SW} = 1.69$. An improved values of κ_{crit} needs to be determined with reweighting factors.

$am_\pi =$	0.05658(6)	$am_K =$	0.2014(4)	$am_D =$ 0.738(3)
$m_\pi/f_\pi =$	1.0731(30)	$m_K/f_K =$	3.188(7)	$m_D/f_D =$ 8.88(11)

$am_N =$	0.3864(9)	$am_{PCAC} =$	$0.189(114)10^{-4}$
$m_N/m_\pi =$	6.829(19)	$P =$	0.5543008(60)

Table 6.7: The masses and the decay constants of the charged pseudoscalar mesons as well as the plaquette P, the PCAC m_{PCAC} and the nucleon mass m_N are presented.

We thus conclude that we have given a successful demonstration that simulations of maximally twisted mass fermions with $N_f = 2+1+1$ quark flavors can be carried out with all quarks of the first two generations tuned to their physical values. This clearly opens the path for the ETM collaboration to perform simulations towards the continuum limit with a rich research program being relevant for phenomenology and ongoing and planned experiments.

7. Pion and nucleon electromagnetic form factor

We discuss results on the pion and nucleon electromagnetic form factors. In particular, we discuss nucleon observables that unlike the ones for the pion have larger statistical uncertainties: although we perform three orders of magnitude more measurements compared to pion form factors, the signal we achieve is not yet precise to shed light on the discrepancy between the muonic determination of the proton charged radius and the one extracted from electron scattering experiments, cf. Sec 1.2.1. The analysis of these ensembles also heavily relied on the application of the multigrid approaches developed.

7.1 Pion electromagnetic form factor

The electromagnetic pion form factor, $F_\pi(Q^2)$ is investigated using small values of the four-momentum transfer $Q^2 \leq 0.25$ GeV^2 [11]. We combined several $N_f = 2$ ensembles produced by the ETM collaboration at fixed lattice spacing, $a \simeq 0.09$ fm and several pion masses, from the physical value up to about 340 MeV, as well as for several lattice volumes. The study included the two physical point ensembles cA2.48 and cA2.64 reported in Tab. 6.1.

The pion form factor is computed from the matrix elements of the electromagnetic vector current

$$V_\mu(x) = \frac{2}{3}\bar{u}(x)\gamma_\mu u(x) - \frac{1}{3}\bar{d}(x)\gamma_\mu d(x) \tag{7.1}$$

between pion states, yielding

$$\left\langle \pi^+(\vec{p}')|V_\mu(0)|\pi^+(\vec{p})\right\rangle = (p'_\mu + p_\mu)F_\pi(Q^2), \tag{7.2}$$

where $q_\mu = (p_\mu - p'_\mu)$ is the 4-momentum transfer and

$$Q^2 \equiv -q^2 = \left|\vec{p} - \vec{p}'\right|^2 - \left[E_\pi(\vec{p}) - E_\pi(\vec{p}')\right]^2 = 4\left|\vec{p}\right|^2,$$

where the right hand side is given in the Breit frame, i.e $\vec{p}' = -\vec{p}$. For the discussion of the pion form factor, we consider only the Breit frame. As detailed in Ref. [178], up to discretization effects of order $\mathcal{O}(a^2)$, it is enough to compute only the connected contributions in Eq. (7.2) considering the single flavor current $\bar{u}(x)\gamma_\mu u(x)$. Since the spatial matrix elements of the vector current are vanishing, i.e. $p'_i + p_i = 0$ with $i = 1, 2, 3$ in Eq. (7.2), we compute the following correlation functions

$$C^{2\text{pt}}(t, \vec{p}) = \sum_{x,z}\left\langle J_\pi(x)J_\pi^\dagger(z)\right\rangle \delta_{t,t_x-t_z}e^{-i\vec{p}\cdot(\vec{x}-\vec{z})} \tag{7.3}$$

$$C_0^{3\text{pt}}(t, t', \vec{p}, -\vec{p}) = \sum_{x,y,z}\left\langle J_\pi(y)V_0(x)J_\pi^\dagger(z)\right\rangle \delta_{t,t_x-t_z}\delta_{t',t_y-t_z}e^{-i\vec{p}\cdot(\vec{x}-\vec{z})-i\vec{p}\cdot(\vec{x}-\vec{y})}, \tag{7.4}$$

where $V_0(x) = \bar{u}(x)\gamma_0 u(x)$ is the temporal component of the local vector current, $J_\pi(x) = \bar{d}(x)\gamma_5 u(x)$ is the interpolating operator annihilating the π^+, t is the time distance between the vector current insertion and the source and t' is the time distance between the sink and the source.
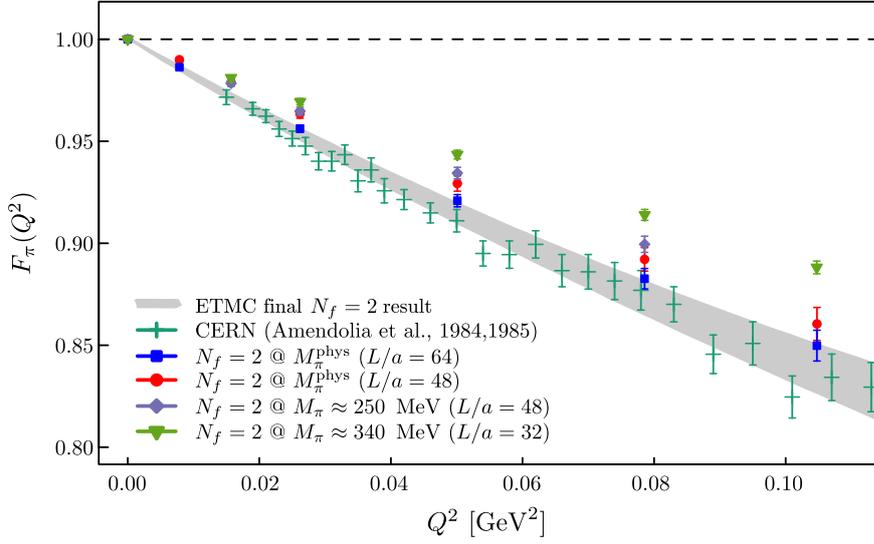
Figure 7.1: Data for the electromagnetic form factor F_π as a function of Q^2 for the $N_f = 2$ ensembles [150], where the two at M_π^{phys} with $L/a = 48$ and 64 are respectively cA2.48 and cA2.64 in Tab. 6.1. The gray band shows the extrapolated form factor at the physical pion point and after taking the infinite volume limit. In addition we show experimental results from CERN [179].

Taking the appropriate limits in the time extent we obtain

$$\lim_{\substack{t \to \infty \\ T \to \infty}} C^{2\mathrm{pt}}(t, \vec{p}) \to \frac{G_\pi^2}{2E_\pi(\vec{p})} e^{-E_\pi(\vec{p})t} \tag{7.5}$$

$$\lim_{\substack{t \to \infty \\ (t'-t) \to \infty \\ T \to \infty}} C_0^{3\mathrm{pt}}(t, t', \vec{p}, -\vec{p}) \to \frac{G_\pi^2}{2E_\pi(\vec{p})2E_\pi(\vec{p})} \left\langle \pi^-(\vec{p}) | V_0 | \pi^+(\vec{p}) \right\rangle e^{-E_\pi(\vec{p})t} e^{-E_\pi(\vec{p})(t'-t)}, \tag{7.6}$$

where G_π^2 is the amplitude of the 2-point correlation function. The ratio

$$R(t, t', \vec{p}) = \frac{C_0^{3\mathrm{pt}}(t, t', \vec{p})}{C^{2\mathrm{pt}}(t', \vec{p})}, \tag{7.7}$$

has the following limit

$$\lim_{\substack{t \to \infty \\ (t'-t) \to \infty \\ T \to \infty}} R(t, t', \vec{p}) \to \frac{\langle \pi^+(-\vec{p}) | V_0 | \pi^+(\vec{p}) \rangle}{2E_\pi(\vec{p})} = \frac{1}{Z_V} F_\pi(Q^2).$$

where the renormalization constant of the vector current, Z_V, is fixed by the normalization condition $F_\pi(0) = 1$. We get $Z_V = 0.6679 \, (1)_{stat} \, (1)_{syst}$ for the lattice spacing considered here.

As depicted in Fig. 7.1, we combine the measurements of $F_\pi(Q^2)$ from multiple twisted mass ensemble obtaining a final result, the gray band in the figure, which takes into account mass and finite volume corrections. For doing so, we used SU(2) Chiral Perturbation Theory (ChPT) at next to leading order (NLO) writing

$$F_\pi(Q^2, m_\pi, L) = 1 + \Delta F_\pi^{\mathrm{NLO}}(Q^2, m_\pi, \infty) + K_{\mathrm{FVE}}^{\mathrm{NLO}}(Q^2, m_\pi, L), \tag{7.8}$$

where $\Delta F_\pi^{\mathrm{NLO}}(Q^2, m_\pi, \infty)$ is the continuum ChPT correction while $K_{\mathrm{FVE}}^{\mathrm{NLO}}(Q^2, m_\pi, L)$ takes into account of the finite size effects. The continuum ChPT corrections are known at several order [160, 180] and

the NLO term is given in powers of $(Q/M_\pi)^2$ as

$$\Delta F_\pi^{\text{NLO}}(Q^2, m_\pi, \infty) = -s_\pi^{\text{NLO}} \frac{Q^2}{M_\pi^2} + c_\pi^{\text{NLO}} \frac{Q^4}{M_\pi^4} + \mathcal{O}\left(\frac{Q^6}{M_\pi^6}\right), \tag{7.9}$$

where the slope and curvature are respectively

$$s_\pi^{\text{NLO}} = \frac{1}{3}\xi\left[\bar{\ell}_6 - \log\frac{\xi}{\xi^{phys}} - 1\right], \tag{7.10}$$

$$c_\pi^{\text{NLO}} = \frac{1}{30}\xi. \tag{7.11}$$

with $\bar{\ell}_6$ being an SU(2) low-energy constant and $\xi \equiv M_\pi^2/(4\pi f_\pi)^2$. The SU(2) ChPT prediction of $K_{\text{FVE}}^{\text{NLO}}(Q^2, m_\pi, L)$ derived at NLO in the Breit frame is available in Refs. [181, 182]. The ensemble $cA2.64$ played an important role in the infinite volume extrapolation showing quantifiable effects also at the physical pion mass, as seen by comparing the results for $cA2.48$ and $cA2.64$ in Fig. 7.1.

An important observable, which can be extracted from the electromagnetic form factor is the mean squared charged radius. It is obtained from

$$\langle r^2 \rangle = -\frac{6}{F(0)} \frac{\partial F(Q^2)}{\partial Q^2}\bigg|_{Q^2=0}. \tag{7.12}$$

Our final value at the physical point is $\langle r^2 \rangle_\pi = 0.443(29)$ fm^2, where the error includes several sources of systematic errors. Discretization effects still need to be determined, since in this study we employed one lattice spacing. The extracted radius is consistent with the experimental value of $\langle r^2 \rangle_\pi^{exp.} = 0.452\ (11)$ fm^2 [17]. This suggests that discretization effects on our result are small and within the current statistical accuracy but this conclusion needs to be verified by repeating the calculation at a smaller value of the lattice spacing. The calculations of $\langle r^2 \rangle_\pi$ from lattice QCD have been examined recently by FLAG and are collected in Table 22 of Ref. [45]. Three results satisfy the FLAG quality criteria, namely: $\langle r^2 \rangle_\pi = 0.456\ (38)$ fm^2 [178] ($N_f = 2$), $\langle r^2 \rangle_\pi = 0.481\ (35)$ fm^2 [183] ($N_f = 2$) and $\langle r^2 \rangle_\pi = 0.403\ (19)$ fm^2 [184] ($N_f = 2+1+1$). Our finding is nicely consistent with all the these lattice results.

7.2   Nucleon electromagnetic form factor

The nucleon matrix element of the electromagnetic form factor is parameterized in terms of the Dirac (F_1) and Pauli (F_2) form factors given in Minkowski space by

$$\langle N(p',s')|V_\mu|N(p,s)\rangle = \sqrt{\frac{m_N^2}{E_N(\vec{p}')E_N(\vec{p})}}\ \bar{u}_N(p',s')\left[\gamma_\mu F_1(Q^2) + \frac{i\sigma_{\mu\nu}q^\nu}{2m_N}F_2(Q^2)\right]u_N(p,s), \quad (7.13)$$

where $N(p,s)$ is the nucleon state with initial (final) momentum p (p') and spin s (s'), with energy $E_N(\vec{p})$ ($E_N(\vec{p}')$) and mass m_N. $Q^2 \equiv q_\mu q^\mu$ is the momentum transfer squared with $q_\mu = (p'_\mu - p_\mu)$, u_N is the nucleon spinor and V_μ is the vector current. The electric and magnetic Sachs form factors $G_E(Q^2)$ and $G_M(Q^2)$ are alternative Lorentz invariant quantities and are expressed in terms of $F_1(Q^2)$ and $F_2(Q^2)$ via the relations,

$$G_E(Q^2) = F_1(Q^2) + \frac{Q^2}{4m_N^2}F_2(Q^2), \tag{7.14}$$

$$G_M(Q^2) = F_1(Q^2) + F_2(Q^2). \tag{7.15}$$

The nucleon electromagnetic form factors are investigated in Ref. [13] for the $cA2.64$ and on the $cB211.64$ ensembles. Here we summarize the results for the $cB211.64$ ensemble considering only the isovector form factors, which do not require disconnected contributions. The isovector current is

$$V_\mu^{u-d} = j_\mu^u - j_\mu^d, \tag{7.16}$$

which gives the difference between the proton and neutron form factors. In contrast to our study for the pion form factor, where we used the local vector current $j_\mu^f = \bar{q}_f \gamma_\mu q_f$, for the nucleon electromagnetic form factors we employ the symmetric lattice conserved vector current given by

$$
\begin{aligned}
j_\mu^f(x) = \frac{1}{4} [& \bar{q}_f(x+\hat{\mu}) U_\mu^\dagger(x)(1+\gamma_\mu) q_f(x) - \bar{q}_f(x) U_\mu(x)(1-\gamma_\mu) q_f(x+\hat{\mu}) \\
& + \bar{q}_f(x) U_\mu^\dagger(x-\hat{\mu})(1+\gamma_\mu) q_f(x-\hat{\mu}) - \bar{q}_f(x-\hat{\mu}) U_\mu(x-\hat{\mu})(1-\gamma_\mu) q_f(x)],
\end{aligned} \tag{7.17}
$$

which does not need renormalization.
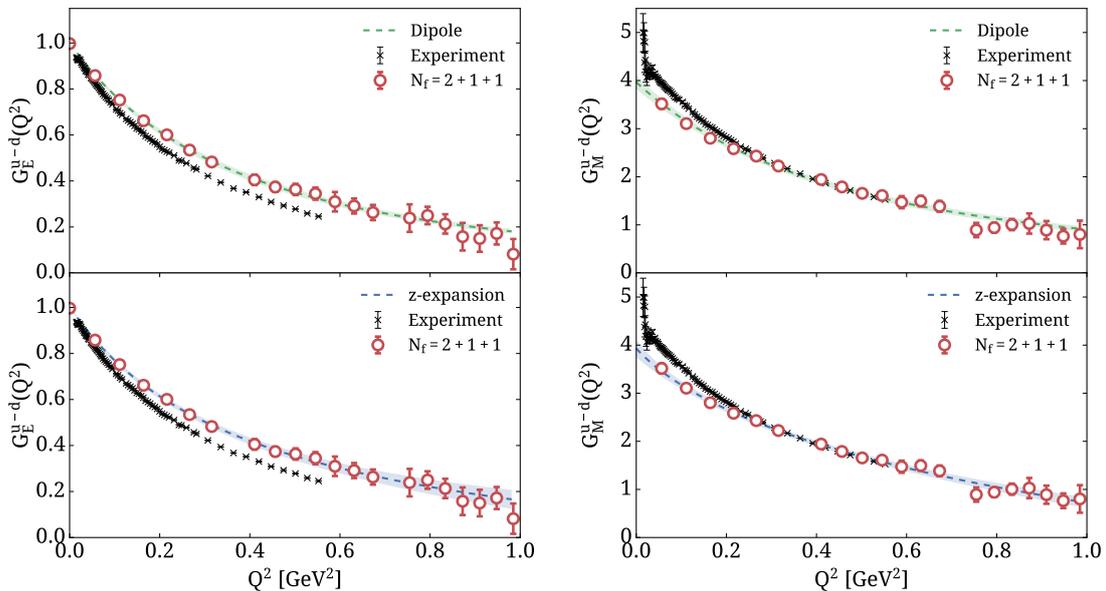
Figure 7.2: *Left:* The isovector electric form factor as a function of Q^2 (circles). We show fits to our results using a dipole form (top) and using the z-expansion (bottom) for $k_{max}{=}4$. Black crosses are experimental results taken from the A1 collaboration [51] for the proton and from Refs. [185–199] for the neutron.
Right: The isovector magnetic form factor fitted using a dipole form (top) and using the z-expansion (bottom).

In Fig. 7.2 we depict the electric and magnetic Sachs form factors for the isovector current measured on the $cB211.64$ ensemble. We use two fitting Ansätze for the Q^2-dependence. Since one expects that for small $Q^2 > 0$ the behavior is dominated by the poles in the time-like region then one can use a dipole form given by [200]

$$G(Q^2) = \frac{G(0)}{(1+\frac{Q^2}{M^2})^2}, \tag{7.18}$$

where M is the mass of the vector meson that parameterizes the Q^2 dependence. A model independent fit, which has been applied recently to experimental data of both electromagnetic and axial form

factors, is the z-expansion [201]. In this case, the form factor is expanded as a series

$$G(Q^2) = \sum_{k=0}^{k_{\max}} a_k z^k,\tag{7.19}$$

where

$$z = \frac{\sqrt{t_{\mathrm{cut}} + Q^2} - \sqrt{t_{\mathrm{cut}}}}{\sqrt{t_{\mathrm{cut}} + Q^2} + \sqrt{t_{\mathrm{cut}}}}\tag{7.20}$$

and t_{cut} is the time-like cut-off. We take $t_{\mathrm{cut}}{=}4m_\pi^2$ for the isovector combination $G^u - G^d$ [201].
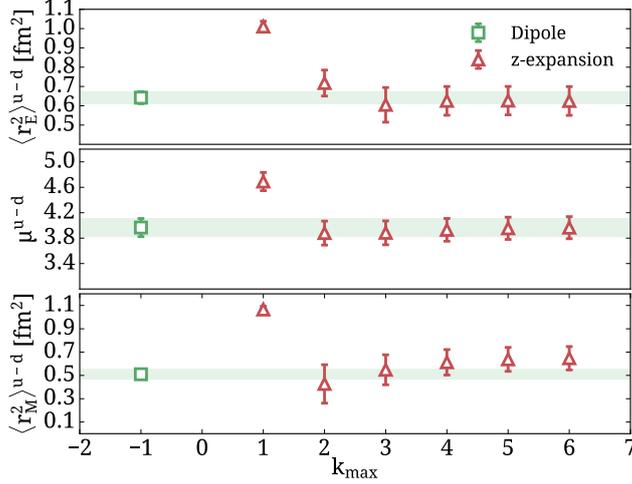
Figure 7.3: Results for the isovector charge radius $\langle r_E^2 \rangle^{u-d}$, magnetic moment μ^{u-d} and magnetic radius $\langle r_M^2 \rangle^{u-d}$ from the plateau method using $t_s/a{=}20$ as extracted from a dipole fit (green square) and z-expansion (red triangles). The latter are shown as a function of $k_{\max}$. The green band is the statistical error on the value extracted from the dipole fit.

The value of the form factor at zero momentum transfer gives the electric charge in the case of the electric form factor and the magnetic moment in the case of the magnetic form factor. We can investigate finite volume effects by comparing the two $N_f{=}2$ twisted mass ensembles with pion mass of 130 MeV with the same lattice spacing but $Lm_\pi{\simeq}3$ and $Lm_\pi{\simeq}4$. We observe consistent results between these two volumes, but we cannot exclude finite volume effects of the order of our statistical errors that may affect the magnetic form factor for small Q^2 values as the volume increases. Further studies are required to take the infinite volume limit and make definite conclusions on the small Q^2 behavior of the magnetic form factor. In Fig. 7.3 we compare the extracted values of the electric and magnetic radius defined in Eq. (7.12) and of the magnetic moment $\mu = G_M(0)$. The values extracted from the dipole fit are all compatible with the z-expansion for $k_{\max} \geq 3$. We take as reference value the one obtained from the dipole fit. In Fig. 7.4 we compare our value with others extracted using our other ensembles or from other collaborations. As can be seen, our results are one of the most precise.

Comparing results calculated using $N_f{=}2$ and $N_f{=}2{+}1{+}1$ twisted mass ensembles we observe no quenching effects. Our values for the electric and magnetic radii as well as the magnetic moment are all compatible within errors. Overall the values extracted within lattice QCD underestimate the electric and magnetic radii. This is due to the slower decay of $G_E^{u-d}(Q^2)$, which needs to be investigated further.
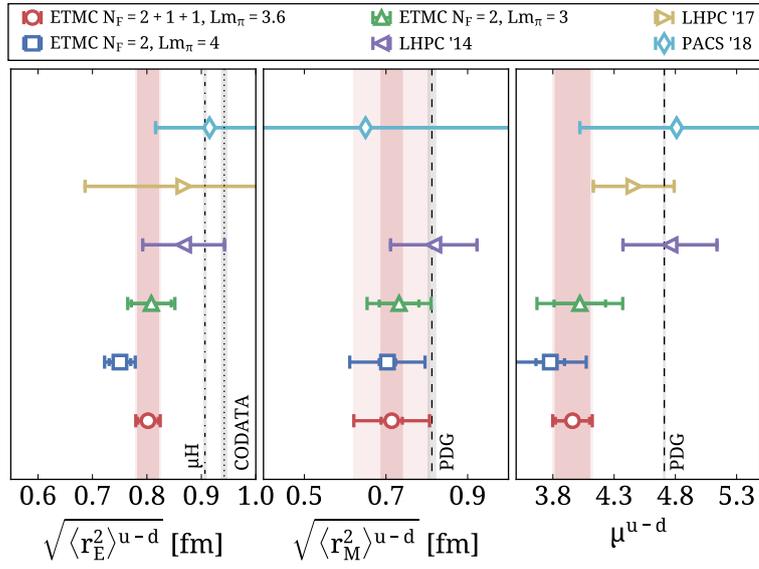
Figure 7.4: Isovector $\sqrt{\langle r_E^2 \rangle^{u-d}}$, $\sqrt{\langle r_M^2 \rangle^{u-d}}$ and μ^{u-d} with lattice QCD results from the $cB211.64$ ensemble, from the $cA2.64$ ensemble with $m_\pi L \simeq 4$ (blue squares), the $cA2.48$ ensemble with $m_\pi L \simeq 3$ from Ref. [149] (green triangles), LHPC using N_f=2+1 stout-smeared clover fermions from Ref. [202] (left purple triangles) and Ref. [203] (right yellow triangles) and from PACS using N_f=2+1 stout-smeared clover fermions from Ref. [204] (cyan rhombuses). The experimental result extracted from muonic hydrogen [47] is shown by the vertical dashed-dotted line and from CODATA [50] by the dotted vertical line. The PDG value [17] is shown with the dashed vertical line. The red vertical inner band denotes the statistical error extracted using the N_f=2+1+1 twisted mass ensemble and the outer lighter band is the total error adding statistical and systematic errors in quadrature.

Conclusions

Improvements to the simulation algorithms have been one of the main focus of this thesis. We have implemented several state-of-the-art approaches achieving two order of magnitude of speed-up. Namely we have employed the DD-αAMG solver for the inversion of the degenerate and non-degenerate twisted mass fermions presented in Chapter 4; we have optimized the solution of shifted linear systems via multigrid methods employing suitable initial guesses for the rational approximation; and accelerated N_f=2 and N_f=2+1+1 twisted mass simulations at the physical point. Additionally our production codes have been optimized for current computing architectures.
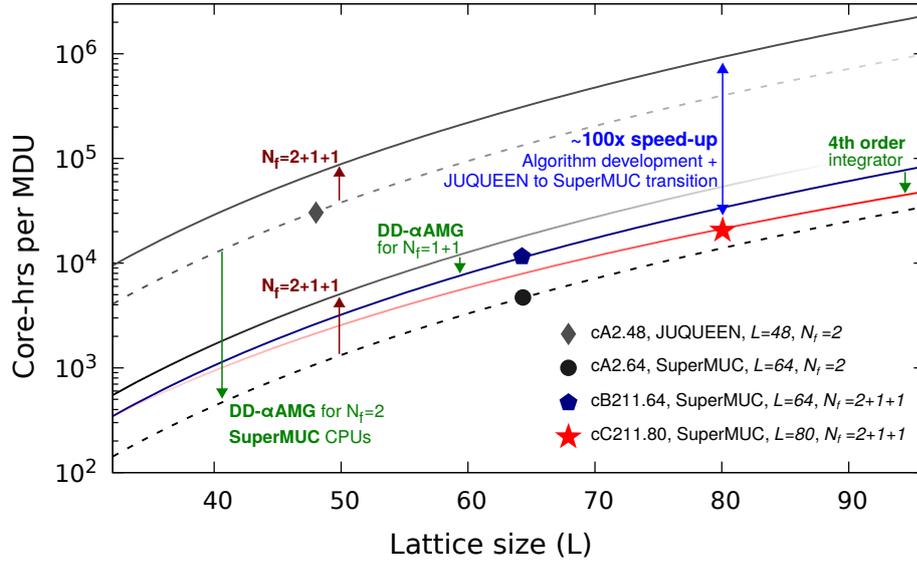
Figure 7.5: Improvements over the years in the ensemble production, in core-hours per molecular dynamics update (MDU) as a function of the lattice length L. Dashed curves are for N_f=2 simulations, while the solid curves are for N_f=2+1+1 simulations. All the curves show the volume scaling according to the standard second order integration schemes, namely L^5. The recently developed fourth order integration scheme is shown with the red curve, following a $L^{4.5}$ dependence. The rhombus, circle, pentagon and star show the data points used to determine the curves. The blue arrow shows the speedup obtained when considering all improvements developed so far for the simulation of the cC211.80 ensemble compared to our first ensemble at the physical point cA2.48. The rest of the arrows show, from left to right: the improvement obtained using DD-αAMG for $N_f = 2$ (green arrow), the cost increase when going from $N_f = 2$ to $N_f = 2 + 1 + 1$ both with and without DD-αAMG (brown arrows), the improvement when using DD-αAMG in the heavy strange-charm sector $(1 + 1)$ of the simulation (green arrow), and the improvement when using a 4th order integration scheme compared to a 2nd order scheme (green arrow).

The various improvements are illustrated in Fig. 7.5, where one can see the large speed-up obtained by transitioning to multigrid, namely almost two orders of magnitude, as well as the improved scaling obtained by employing fourth order integration schemes. N_f=2 simulations are accelerated by employing and tuning the DD-αAMG method within the simulation algorithm. It resulted in a speed-up of a factor of about one order of magnitude for simulations at the physical pion mass. The speed-up is smaller than the two order of magnitudes achieved for the single inversion at the physical pion mass. This is because the multigrid method requires a setup phase the cost of which cannot be avoided in case of simulations since it is configuration dependent and in HMC simulations the configuration evolves during the integration following a molecular dynamic trajectory. Moreover, many

other calculations are included in a HMC step and the calculation of physical quark propagators is only a fraction. The N_f=1+1 sector of the simulation was also accelerated achieving a speed-up of a factor of two for this specific sector in the simulation.

We have also implemented and analyzed fourth order integration schemes, namely the force gradient scheme (2NMFG) [130, 205] and the Omelyan-Mryglod-Folk (OMF4) integrator. Fourth order schemes have a better volume scaling compared to second order schemes: at fixed acceptance rate the number of integration steps scales with $V^{1/8}$ for the first and with $V^{1/4}$ for the latter. The drawback is that the fourth order schemes require more applications of gauge derivatives and thus more inversions of the operator D per integration steps. The scaling therefore pays back as the lattice size increases, as shown by the red curve in Fig. 7.5.

Using the aforementioned improvements, the ETM collaboration was enable to simulate three ensembles at the physical point. These are the ensembles cA2.64, cA2.64.r, cB211.64 and cC211.80 which cost per MDU is depicted in Fig. 7.5 and simulation details have been described in Chapter 6. The ensemble cA2.64, with $N_f = 2$ and lattice size $L = 64$ has been simulated for studying finite size effects by comparing results computed on cA2.48, which differ only on the lattice size having $L = 48$. The ensembles cB211.64 and cC211.80 have been simulated for taking the continuum limit in $N_f = 2 + 1 + 1$ QCD with light, strange and charm quark masses tuned to their physical value. Results on the cA2.64 and cB211.64 have been computed while the ensemble cC211.80 has not been analyzed yet. In Chapter 7, we showed results for the electromagnetic form factor on these ensembles. Analyses and publications of other quantities are under preparation.

Bibliography

[1]   C. Alexandrou et al. "Adaptive Aggregation-based Domain Decomposition Multigrid for Twisted Mass Fermions". In: *Phys. Rev.* D94.11 (2016), p. 114509. DOI: 10.1103/PhysRevD. 94.114509. arXiv: 1610.02370 [hep-lat].

[2]   S. Bacchio et al. "DDalphaAMG for Twisted Mass Fermions". In: *PoS* LATTICE2016 (2016), p. 259. arXiv: 1611.01034 [hep-lat]. URL: http://inspirehep.net/record/1495722/ files/arXiv:1611.01034.pdf.

[3]   C. Alexandrou, S. Bacchio, and J. Finkenrath. "Multigrid approach in shifted linear systems for the non-degenerated twisted mass operator". In: *Comput. Phys. Commun.* 236 (2019), pp. 51–64. DOI: 10.1016/j.cpc.2018.10.013. arXiv: 1805.09584 [hep-lat].

[4]   S. Bacchio, C. Alexandrou, and J. Finkenrath. "Multigrid accelerated simulations for Twisted Mass fermions". In: *35th International Symposium on Lattice Field Theory (Lattice 2017) Granada, Spain, June 18-24, 2017*. 2017. arXiv: 1710.06198 [hep-lat]. URL: http:// inspirehep.net/record/1631185/files/arXiv:1710.06198.pdf.

[5]   S. Bacchio and J. Finkenrath. *DDalphaAMG, twisted mass version.* Available at https:// github.com/sbacchio/DDalphaAMG.

[6]   J. Finkenrath et al. "Simulation of an ensemble of $N_f = 2 + 1 + 1$ twisted mass clover-improved fermions at physical quark masses". In: *EPJ Web Conf.* 175 (2018), p. 02003. DOI: 10.1051/epjconf/201817502003. arXiv: 1712.09579 [hep-lat].

[7]   S. Bacchio and J. Finkenrath. *tmLQCD interfaced to DDalphaAMG.* Available at https: //github.com/finkerath/tmLQCD/tree/DDalphaAMG.

[8]   C. Alexandrou et al. "Simulating twisted mass fermions at physical light, strange and charm quark masses". In: *Phys. Rev.* D98.5 (2018), p. 054518. DOI: 10.1103/PhysRevD.98. 054518. arXiv: 1807.00495 [hep-lat].

[9]   L. Liu et al. "Isospin-0 $\pi\pi$ s-wave scattering length from twisted mass lattice QCD". In: *Phys. Rev.* D96.5 (2017), p. 054516. DOI: 10.1103/PhysRevD.96.054516. arXiv: 1612.02061 [hep-lat].

[10]   L. Liu et al. "Isospin-0 $\pi\pi$ scattering from twisted mass lattice QCD". In: *PoS* LATTICE2016 (2017), p. 119. DOI: 10.22323/1.256.0119. arXiv: 1701.08961 [hep-lat].

[11]   C. Alexandrou et al. "Pion vector form factor from lattice QCD at the physical point". In: *Phys. Rev.* D97.1 (2018), p. 014508. DOI: 10.1103/PhysRevD.97.014508. arXiv: 1710. 10401 [hep-lat].

[12]   C. Alexandrou et al. "Computation of parton distributions from the quasi-PDF approach at the physical point". In: *EPJ Web Conf.* 175 (2018), p. 14008. DOI: 10.1051/epjconf/ 201817514008. arXiv: 1710.06408 [hep-lat].

[13]   C. Alexandrou et al. "Proton and neutron electromagnetic form factors from lattice QCD". In: (2018). arXiv: 1812.10311 [hep-lat].

[14] M. Lüscher. "Local coherence and deflation of the low quark modes in lattice QCD". In: *JHEP* 07 (2007), p. 081. DOI: 10.1088/1126-6708/2007/07/081. arXiv: 0706.2298 [hep-lat].

[15] A. Einstein. "On the general theory of relativity". In: *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* 1915 (1915), pp. 799–801.

[16] M. E. Peskin. *An introduction to quantum field theory*. CRC Press, 2018.

[17] C. Patrignani et al. "Review of Particle Physics". In: *Chin. Phys.* C40.10 (2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001.

[18] S. L. Glashow. "Partial-symmetries of weak interactions". In: *Nuclear Physics* 22.4 (1961), pp. 579–588.

[19] A. Salam. "Weak and electromagnetic interactions". In: *Elementary particle physics: relativistic groups and analyticity, N. Svartholm, ed* (1968), p. 367.

[20] S. Weinberg. "A model of leptons". In: *Physical review letters* 19.21 (1967), p. 1264.

[21] M. Gell-Mann and E. P. Rosenbaum. "Elementary particles". In: *Scientific American* 197.1 (1957), pp. 72–92.

[22] S. Chatrchyan et al. "Evidence for the direct decay of the 125 GeV Higgs boson to fermions". In: *Nature Phys.* 10 (2014), pp. 557–560. DOI: 10.1038/nphys3005. arXiv: 1401.6527 [hep-ex].

[23] R. Feynman, M. Gell-Mann, and G. Zweig. "Group U (6) ⊗ U (6) Generated by Current Components". In: *Physical Review Letters* 13.22 (1964), p. 678.

[24] R. P. Feynman, R. B. Leighton, and M. Sands. "The feynman lectures on physics; vol. i". In: *American Journal of Physics* 33.9 (1965), pp. 750–752.

[25] H. Fritzsch, M. Gell-Mann, and H. Leutwyler. "Advantages of the color octet gluon picture". In: *Physics Letters B* 47.4 (1973), pp. 365–368. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(73)90625-4. URL: http://www.sciencedirect.com/science/article/pii/0370269373906254.

[26] C.-N. Yang and R. L. Mills. "Conservation of isotopic spin and isotopic gauge invariance". In: *Physical review* 96.1 (1954), p. 191.

[27] D. Barkai, M. Creutz, and K. Moriarty. "Monte Carlo study of SU (3) gauge theory on a 12 4 lattice". In: *Physical Review D* 29.6 (1984), p. 1207.

[28] H. Georgi and H. D. Politzer. "Electroproduction scaling in an asymptotically free theory of strong interactions". In: *Physical Review D* 9.2 (1974), p. 416.

[29] H. Georgi and H. D. Politzer. "Freedom at moderate energies: Masses in color dynamics". In: *Physical Review D* 14.7 (1976), p. 1829.

[30] K. G. Wilson. "KG Wilson, Phys. Rev. D 10, 2445 (1974)." In: *Phys. Rev. D* 10 (1974), p. 2445.

[31] M. Creutz. "Monte Carlo Study of Quantized SU(2) Gauge Theory". In: *Phys. Rev.* D21 (1980), pp. 2308–2315. DOI: 10.1103/PhysRevD.21.2308.

[32] S. Durr et al. "Ab-Initio Determination of Light Hadron Masses". In: *Science* 322 (2008), pp. 1224–1227. DOI: 10.1126/science.1163233. arXiv: 0906.3599 [hep-lat].

[33] R. Aaij et al. "Observation of the doubly charmed baryon Ξ_{cc}^{++}". In: *Phys. Rev. Lett.* 119.11 (2017), p. 112001. DOI: 10.1103/PhysRevLett.119.112001. arXiv: 1707.01621 [hep-ex].

[34] H. Na and S. A. Gottlieb. "Charm and bottom heavy baryon mass spectrum from lattice QCD with 2+1 flavors". In: *PoS* LATTICE2007 (2007), p. 124. DOI: 10.22323/1.042.0124. arXiv: 0710.1422 [hep-lat].

[35] H. Na and S. Gottlieb. "Heavy baryon mass spectrum from lattice QCD with 2+1 dynamical sea quark flavors". In: *PoS* LATTICE2008 (2008), p. 119. DOI: 10.22323/1.066.0119. arXiv: 0812.1235 [hep-lat].

[36] S. Aoki et al. "2+1 Flavor Lattice QCD toward the Physical Point". In: *Phys. Rev.* D79 (2009), p. 034503. DOI: 10.1103/PhysRevD.79.034503. arXiv: 0807.1661 [hep-lat].

[37] L. Liu et al. "Singly and Doubly Charmed J=1/2 Baryon Spectrum from Lattice QCD". In: *Phys. Rev.* D81 (2010), p. 094505. DOI: 10.1103/PhysRevD.81.094505. arXiv: 0909.3294 [hep-lat].

[38] W. Bietenholz et al. "Flavour blindness and patterns of flavour symmetry breaking in lattice simulations of up, down and strange quarks". In: *Phys. Rev.* D84 (2011), p. 054509. DOI: 10.1103/PhysRevD.84.054509. arXiv: 1102.5300 [hep-lat].

[39] R. A. Briceno, H.-W. Lin, and D. R. Bolton. "Charmed-Baryon Spectroscopy from Lattice QCD with $N_f = 2+1+1$ Flavors". In: *Phys. Rev.* D86 (2012), p. 094504. DOI: 10.1103/PhysRevD.86.094504. arXiv: 1207.3536 [hep-lat].

[40] C. Alexandrou et al. "Baryon spectrum with $N_f = 2+1+1$ twisted mass fermions". In: *Phys. Rev.* D90.7 (2014), p. 074501. DOI: 10.1103/PhysRevD.90.074501. arXiv: 1406.4310 [hep-lat].

[41] Z. S. Brown et al. "Charmed bottom baryon spectroscopy from lattice QCD". In: *Phys. Rev.* D90.9 (2014), p. 094507. DOI: 10.1103/PhysRevD.90.094507. arXiv: 1409.0497 [hep-lat].

[42] P. Pérez-Rubio, S. Collins, and G. S. Bali. "Charmed baryon spectroscopy and light flavor symmetry from lattice QCD". In: *Phys. Rev.* D92.3 (2015), p. 034504. DOI: 10.1103/PhysRevD.92.034504. arXiv: 1503.08440 [hep-lat].

[43] C. Alexandrou and C. Kallidonis. "Low-lying baryon masses using $N_f = 2$ twisted mass clover-improved fermions directly at the physical pion mass". In: *Phys. Rev.* D96.3 (2017), p. 034511. DOI: 10.1103/PhysRevD.96.034511. arXiv: 1704.02647 [hep-lat].

[44] K. A. Olive et al. "Review of Particle Physics". In: *Chin. Phys.* C38 (2014), p. 090001. DOI: 10.1088/1674-1137/38/9/090001.

[45] S. Aoki et al. "FLAG Review 2019". In: (2019). arXiv: 1902.08191 [hep-lat].

[46] R. Frezzotti and G. C. Rossi. "Chirally improving Wilson fermions. 1. O(a) improvement". In: *JHEP* 08 (2004), p. 007. DOI: 10.1088/1126-6708/2004/08/007. arXiv: hep-lat/0306014 [hep-lat].

[47] R. Pohl et al. "The size of the proton". In: *Nature* 466 (2010), pp. 213–216. DOI: 10.1038/nature09250.

[48] J. Arrington. "An examination of proton charge radius extractions from e-p scattering data". In: *J. Phys. Chem. Ref. Data* 44 (2015), p. 031203. DOI: 10.1063/1.4922414. arXiv: 1506.00873 [nucl-ex].

[49] A. Antognini et al. "Proton Structure from the Measurement of $2S - 2P$ Transition Frequencies of Muonic Hydrogen". In: *Science* 339 (2013), pp. 417–420. DOI: 10.1126/science.1230016.

[50] P. J. Mohr, D. B. Newell, and B. N. Taylor. "CODATA Recommended Values of the Fundamental Physical Constants: 2014". In: *Rev. Mod. Phys.* 88.3 (2016), p. 035009. DOI: 10.1103/RevModPhys.88.035009. arXiv: 1507.07956 [physics.atom-ph].

[51] J. C. Bernauer et al. "Electric and magnetic form factors of the proton". In: *Phys. Rev.* C90.1 (2014), p. 015206. DOI: 10.1103/PhysRevC.90.015206. arXiv: 1307.6227 [nucl-ex].

[52] X. Zhan et al. "High-Precision Measurement of the Proton Elastic Form Factor Ratio $\mu_p G_E/G_M$ at low Q^2". In: *Phys. Lett.* B705 (2011), pp. 59–64. DOI: 10.1016/j.physletb.2011.10.002. arXiv: 1102.0318 [nucl-ex].

[53] R. P. Feynman. "RP Feynman, Rev. Mod. Phys. 20, 367 (1948)." In: *Rev. Mod. Phys.* 20 (1948), p. 367.

[54] R. P. Feynman, A. R. Hibbs, and D. F. Styer. *Quantum mechanics and path integrals*. Courier Corporation, 2010.

[55] I. Montvay and G. Münster. *Quantum fields on a lattice*. Cambridge University Press, 1997.

[56] G. Münster and P. Weisz. "On the roughening transition in abelian lattice gauge theories". In: *Nuclear Physics B* 180.1 (1981), pp. 13–22. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(81)90152-8. URL: http://www.sciencedirect.com/science/article/pii/0550321381901528.

[57] S. R. Sharpe. "Quenched chiral logarithms". In: *Phys. Rev.* D46 (1992), pp. 3146–3168. DOI: 10.1103/PhysRevD.46.3146. arXiv: hep-lat/9205020 [hep-lat].

[58] B. Orth, T. Lippert, and K. Schilling. "Finite-size effects in lattice QCD with dynamical Wilson fermions". In: *Phys. Rev.* D72 (2005), p. 014503. DOI: 10.1103/PhysRevD.72.014503. arXiv: hep-lat/0503016 [hep-lat].

[59] Y. Iwasaki. "Renormalization Group Analysis of Lattice Theories and Improved Lattice Action: Two-Dimensional Nonlinear O(N) Sigma Model". In: *Nucl. Phys.* B258 (1985), pp. 141–156. DOI: 10.1016/0550-3213(85)90606-6.

[60] D. J. Gross and F. Wilczek. "Ultraviolet Behavior of Nonabelian Gauge Theories". In: *Phys.Rev.Lett.* 30 (1973), pp. 1343–1346. DOI: 10.1103/PhysRevLett.30.1343.

[61] H. D. Politzer. "Reliable Perturbative Results for Strong Interactions?" In: *Phys.Rev.Lett.* 30 (1973), pp. 1346–1349. DOI: 10.1103/PhysRevLett.30.1346.

[62] W. E. Caswell. "Asymptotic Behavior of Nonabelian Gauge Theories to Two Loop Order". In: *Phys.Rev.Lett.* 33 (1974), p. 244. DOI: 10.1103/PhysRevLett.33.244.

[63] D. Jones. "Two Loop Diagrams in Yang-Mills Theory". In: *Nucl.Phys.* B75 (1974), p. 531. DOI: 10.1016/0550-3213(74)90093-5.

[64] M. Creutz. "Quarks, Guons and Lattices". In: (1984).

[65] K. G. Wilson. "Quarks and Strings on a Lattice". In: (1975), p. 99.

[66] B. Sheikholeslami and R. Wohlert. "Improved Continuum Limit Lattice Action for QCD with Wilson Fermions". In: *Nucl. Phys.* B259 (1985), p. 572. DOI: 10.1016/0550-3213(85)90002-1.

[67] M. Bochicchio et al. "Chiral Symmetry on the Lattice with Wilson Fermions". In: *Nucl.Phys.* B262 (1985), p. 331. DOI: 10.1016/0550-3213(85)90290-1.

[68] C. Gattringer and C. Lang. "Quantum Chromodynamics on the Lattice: An Introductory Presentation". In: Lecture Notes in Physics. Springer, 2009. Chap. 11, pp. 267–279. ISBN: 9783642018497. URL: https://books.google.it/books?id=l2hZKnlYDxoC.

[69] R. Frezzotti et al. "A Local formulation of lattice QCD without unphysical fermion zero modes". In: *Nucl. Phys. Proc. Suppl.* 83 (2000), pp. 941–946. DOI: 10.1016/S0920-5632(00)91852-8. arXiv: hep-lat/9909003 [hep-lat].

[70] R. Frezzotti et al. "Lattice QCD with a chirally twisted mass term". In: *JHEP* 08 (2001), p. 058. arXiv: hep-lat/0101001 [hep-lat].

[71] R. Frezzotti, S. Sint, and P. Weisz. "O(a) improved twisted mass lattice QCD". In: *JHEP* 07 (2001), p. 048. DOI: 10.1088/1126-6708/2001/07/048. arXiv: hep-lat/0104014 [hep-lat].

[72] S. R. Sharpe and J. M. S. Wu. "Twisted mass chiral perturbation theory at next-to-leading order". In: *Phys. Rev.* D71 (2005), p. 074501. DOI: 10.1103/PhysRevD.71.074501. arXiv: hep-lat/0411021 [hep-lat].

[73] G. Herdoiza et al. "Determination of Low-Energy Constants of Wilson Chiral Perturbation Theory". In: *JHEP* 05 (2013), p. 038. DOI: 10.1007/JHEP05(2013)038. arXiv: 1303.3516 [hep-lat].

[74] D. Becirevic et al. "Exploring twisted mass lattice QCD with the Clover term". In: *Phys. Rev.* D74 (2006), p. 034501. DOI: 10.1103/PhysRevD.74.034501. arXiv: hep-lat/0605006 [hep-lat].

[75] P. Dimopoulos, H. Simma, and A. Vladikas. "Quenched B(K)-parameter from Osterwalder-Seiler tmQCD quarks and mass-splitting discretization effects". In: *JHEP* 07 (2009), p. 007. DOI: 10.1088/1126-6708/2009/07/007. arXiv: 0902.1074 [hep-lat].

[76] A. Abdel-Rehim et al. "Simulating QCD at the physical point with $N_f = 2$ Wilson twisted mass fermions at maximal twist". In: (2015). arXiv: 1507.05068 [hep-lat].

[77] S. Aoki, R. Frezzotti, and P. Weisz. "Computation of the improvement coefficient c(SW) to one loop with improved gluon actions". In: *Nucl. Phys.* B540 (1999), pp. 501–519. DOI: 10.1016/S0550-3213(98)00742-1. arXiv: hep-lat/9808007 [hep-lat].

[78] R. Frezzotti et al. "Reducing cutoff effects in maximally twisted lattice QCD close to the chiral limit". In: *JHEP* 04 (2006), p. 038. DOI: 10.1088/1126-6708/2006/04/038. arXiv: hep-lat/0503034 [hep-lat].

[79] K. Jansen et al. "Quenched scaling of Wilson twisted mass fermions". In: *JHEP* 09 (2005), p. 071. DOI: 10.1088/1126-6708/2005/09/071. arXiv: hep-lat/0507010 [hep-lat].

[80] A. Frommer et al. "Adaptive aggregation based domain decomposition multigrid for the lattice Wilson Dirac operator". In: *SIAM J. Sci. Comput.* 36 (2014), A1581–A1608. DOI: `10.1137/130919507`. arXiv: `1303.1377` `[hep-lat]`.

[81] L. Del Debbio et al. "Stability of lattice QCD simulations and the thermodynamic limit". In: *JHEP* 02 (2006), p. 011. DOI: `10.1088/1126-6708/2006/02/011`. arXiv: `hep-lat/0512021` `[hep-lat]`.

[82] R. Frezzotti and G. C. Rossi. "Twisted mass lattice QCD with mass nondegenerate quarks". In: *Nucl. Phys. Proc. Suppl.* 128 (2004). [,193(2003)], pp. 193–202. DOI: `10.1016/S0920-5632(03)02477-0`. arXiv: `hep-lat/0311008` `[hep-lat]`.

[83] R. Frezzotti and G. C. Rossi. "Chirally improving Wilson fermions. II. Four-quark operators". In: *JHEP* 10 (2004), p. 070. DOI: `10.1088/1126-6708/2004/10/070`. arXiv: `hep-lat/0407002` `[hep-lat]`.

[84] T. Chiarappa et al. "Numerical simulation of QCD with u, d, s and c quarks in the twisted-mass Wilson formulation". In: *Eur. Phys. J.* C50 (2007), pp. 373–383. DOI: `10.1140/epjc/s10052-006-0204-4`. arXiv: `hep-lat/0606011` `[hep-lat]`.

[85] Y. Saad. *Iterative methods for sparse linear systems*. Vol. 82. siam, 2003.

[86] Y. Saad. "Chebyshev Acceleration techniques for solving nonsymmetric eigenvalue problems". In: 42.166 (Apr. 1984), pp. 567–588. ISSN: 0025-5718 (paper), 1088-6842 (electronic). DOI: `http://dx.doi.org/10.2307/2007602`.

[87] H. Neff et al. "On the low fermionic eigenmode dominance in QCD on the lattice". In: *Phys. Rev.* D64 (2001), p. 114509. DOI: `10.1103/PhysRevD.64.114509`. arXiv: `hep-lat/0106016` `[hep-lat]`.

[88] A. Stathopoulos and K. Orginos. "Computing and deflating eigenvalues while solving multiple right hand side linear systems in quantum chromodynamics". In: *SIAM J. Sci. Comput.* 32 (2010), pp. 439–462. DOI: `10.1137/080725532`. arXiv: `0707.0131` `[hep-lat]`.

[89] R. Lehoucq and J. Scott. "An evaluation of software for computing eigenvalues of sparse nonsymmetric matrices". In: *Preprint MCS-P547* 1195 (1996), p. 5.

[90] K. Jansen and C. Urbach. "tmLQCD: A program suite to simulate Wilson twisted mass lattice QCD". In: *Comput. Phys. Commun.* 180 (2009), pp. 2717–2738. DOI: `10.1016/j.cpc.2009.05.016`. arXiv: `0905.3331` `[hep-lat]`.

[91] P. Wesseling. *An introduction to multigrid methods*. Pure and applied mathematics. John Wiley & Sons Australia, Limited, 1992. ISBN: 9780471930839. URL: `https://books.google.com.cy/books?id=ywOzQgAACAAJ`.

[92] R. C. Brower et al. "The multigrid method for fermion calculations in quantum chromodynamics". In: (1987).

[93] R. Ben-Av et al. "Fermion simulations using parallel transported multigrid". In: *Phys. Lett.* B253 (1991), pp. 185–192. DOI: `10.1016/0370-2693(91)91382-6`.

[94] J. C. Vink. "Multigrid inversion of staggered and Wilson fermion operators with SU(2) gauge fields in two-dimensions". In: *Phys. Lett.* B272 (1991), pp. 81–85. DOI: `10.1016/0370-2693(91)91016-0`.

[95] T. Kalkreuter. "Multigrid methods for propagators in lattice gauge theories". In: *J. Comput. Appl. Math.* 63 (1995), p. 57. DOI: 10.1016/0377-0427(95)00049-6. arXiv: hep-lat/9409008 [hep-lat].

[96] M. Lüscher. "Deflation acceleration of lattice QCD simulations". In: *JHEP* 12 (2007), p. 011. DOI: 10.1088/1126-6708/2007/12/011. arXiv: 0710.5417 [hep-lat].

[97] J. Brannick et al. "Adaptive multigrid algorithm for lattice QCD". In: *Phys. Rev. Lett.* 100 (2008), p. 041601. DOI: 10.1103/PhysRevLett.100.041601. arXiv: 0707.4018 [hep-lat].

[98] R. Babich et al. "Adaptive multigrid algorithm for the lattice Wilson-Dirac operator". In: *Phys. Rev. Lett.* 105 (2010), p. 201602. DOI: 10.1103/PhysRevLett.105.201602. arXiv: 1005.3043 [hep-lat].

[99] J. Brannick et al. "Adaptive multigrid algorithm for the QCD Dirac-Wilson Operator". In: *PoS* LAT2007 (2007), p. 029. arXiv: 0710.3612 [hep-lat].

[100] M. A. Clark et al. "The removal of critical slowing down". In: *PoS* LATTICE2008 (2008), p. 035. arXiv: 0811.4331 [hep-lat].

[101] R. Babich et al. "The role of multigrid algorithms for LQCD". In: *PoS* LAT2009 (2009), p. 031. arXiv: 0912.2186 [hep-lat].

[102] J. C. Osborn et al. "Multigrid solver for clover fermions". In: *PoS* LATTICE2010 (2010), p. 037. arXiv: 1011.2775 [hep-lat].

[103] I. Gohberg, P. Lancaster, and L. Rodman. *Indefinite linear algebra and applications*. Springer Science & Business Media, 2006.

[104] M. Brezina et al. "Adaptive smoothed aggregation (α SA) multigrid". In: *SIAM review* 47.2 (2005), pp. 317–346.

[105] A. Frommer et al. "An adaptive aggregation based domain decomposition multilevel method for the lattice wilson dirac operator: multilevel results". In: (2013). arXiv: 1307.6101 [hep-lat].

[106] M. Rottmann. "Adaptive domain decomposition multigrid for lattice QCD". PhD thesis. Bergische Universität Wuppertal, 2016.

[107] M. Lüscher. "Solution of the Dirac equation in lattice QCD using a domain decomposition method". In: *Comput. Phys. Commun.* 156 (2004), pp. 209–220. DOI: 10.1016/S0010-4655(03)00486-7. arXiv: hep-lat/0310048 [hep-lat].

[108] J. A. Vogel. "Flexible BiCG and flexible Bi-CGSTAB for nonsymmetric linear systems". In: *Appl. Math. Comput.* 188.1 (2007), pp. 226–233. ISSN: 0096-3003. DOI: 10.1016/j.amc.2006.09.116. URL: http://dx.doi.org/10.1016/j.amc.2006.09.116.

[109] S. C. Eisenstat, H. C. Elman, and M. H. Schultz. "Variational iterative methods for nonsymmetric systems of linear equations". In: *SIAM J. Numer. Anal.* 20.2 (1983), pp. 345–357. ISSN: 0036-1429. DOI: 10.1137/0720023. URL: http://dx.doi.org/10.1137/0720023.

[110] Y. Saad. "A flexible inner-outer preconditioned GMRES algorithm". In: *SIAM J. Sci. Comput.* 14.2 (1993), pp. 461–469. ISSN: 1064-8275. DOI: 10.1137/0914028. URL: http://dx.doi.org/10.1137/0914028.

[111] M. Rottmann et al. *DDalphaAMG*. Available at https://github.com/DDalphaAMG/DDalphaAMG.

[112] J. Brannick et al. "Multigrid preconditioning for the overlap operator in lattice QCD". In: *Numer. Math.* (2015). DOI: 10.1007/s00211-015-0725-6. arXiv: 1410.7170 [hep-lat].

[113] S. D. Cohen et al. "Multigrid algorithms for Domain-Wall fermions". In: *PoS* LATTICE2011 (2011), p. 030. arXiv: 1205.2933 [hep-lat].

[114] R. C. Brower et al. "Multigrid for Staggered Lattice Fermions". In: *Submitted to: Phys. Rev. D* (2018). arXiv: 1801.07823 [hep-lat].

[115] G. S. Bali et al. "Nucleon isovector couplings from $N_f = 2$ lattice QCD". In: *Phys. Rev.* D91.5 (2015), p. 054501. DOI: 10.1103/PhysRevD.91.054501. arXiv: 1412.7336 [hep-lat].

[116] J. Liesen and P. Tichý. "Convergence analysis of Krylov subspace methods". In: *GAMM-Mitt.* 27.2 (2004), pp. 153–173.

[117] R. Babich et al. *QUDA: A library for QCD on GPUsQUDA, version 0.8*. Available at http://lattice.github.io/quda/.

[118] M. Lüscher. *OpenQCD software, version 1.2 and 1.4*. Available at http://luscher.web.cern.ch/luscher/openQCD.

[119] M. Lüscher. "Stochastic locality and master-field simulations of very large lattices". In: *35th International Symposium on Lattice Field Theory (Lattice 2017) Granada, Spain, June 18-24, 2017*. 2017. arXiv: 1707.09758 [hep-lat]. URL: https://inspirehep.net/record/1613675/files/arXiv:1707.09758.pdf.

[120] S. Heybrock et al. "Adaptive algebraic multigrid on SIMD architectures". In: (2015). arXiv: 1512.04506 [physics.comp-ph]. URL: https://inspirehep.net/record/1409505/files/arXiv:1512.04506.pdf.

[121] D. Richtmann, S. Heybrock, and T. Wettig. "Multiple right-hand-side setup for the DD-αAMG". In: *PoS* LATTICE2015 (2016), p. 035. arXiv: 1601.03184 [hep-lat].

[122] N. Metropolis et al. "Equation of state calculations by fast computing machines". In: *The journal of chemical physics* 21.6 (1953), pp. 1087–1092.

[123] S. Schaefer, R. Sommer, and F. Virotta. "Critical slowing down and error analysis in lattice QCD simulations". In: *Nucl. Phys.* B845 (2011), pp. 93–119. DOI: 10.1016/j.nuclphysb.2010.11.020. arXiv: 1009.5228 [hep-lat].

[124] N. Cabibbo and E. Marinari. "A new method for updating SU(N) matrices in computer simulations of gauge theories". In: *Physics Letters B* 119.4 (1982), pp. 387–390. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(82)90696-7. URL: http://www.sciencedirect.com/science/article/pii/0370269382906967.

[125] F. Fucito et al. "A Proposal for Monte Carlo Simulations of Fermionic Systems". In: *Nucl. Phys.* B180 (1981). [,586(1980)], p. 369. DOI: 10.1016/0550-3213(81)90055-9.

[126] D. H. Weingarten and D. N. Petcher. "Monte Carlo Integration for Lattice Gauge Theories with Fermions". In: *Phys. Lett.* 99B (1981), pp. 333–338. DOI: 10.1016/0370-2693(81)90112-X.

[127] S. Duane et al. "Hybrid monte carlo". In: *Phys. Lett.* B195 (1987), pp. 216–222. DOI: 10.1016/0370-2693(87)91197-X.

[128] S. A. Gottlieb et al. "Hybrid molecular dynamics algorithms for the numerical simulation of quantum chromodynamics". In: *Phys. Rev.* D35 (1987), pp. 2531–2542. DOI: 10.1103/PhysRevD.35.2531.

[129] M. Luscher. "Computational Strategies in Lattice QCD". In: *Modern perspectives in lattice QCD: Quantum field theory and high performance computing. Proceedings, International School, 93rd Session, Les Houches, France, August 3-28, 2009*. 2010, pp. 331–399. arXiv: 1002.4232 [hep-lat].

[130] A. D. Kennedy, P. J. Silva, and M. A. Clark. "Shadow Hamiltonians, Poisson Brackets, and Gauge Theories". In: *Phys. Rev.* D87.3 (2013), p. 034511. DOI: 10.1103/PhysRevD.87.034511. arXiv: 1210.6600 [hep-lat].

[131] I. Omelyan, I. 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 (2003), pp. 272–314. ISSN: 0010-4655. DOI: https://doi.org/10.1016/S0010-4655(02)00754-3. URL: http://www.sciencedirect.com/science/article/pii/S0010465502007543.

[132] M. Hasenbusch and K. Jansen. "Speeding up lattice QCD simulations with clover improved Wilson fermions". In: *Nucl. Phys.* B659 (2003), pp. 299–320. DOI: 10.1016/S0550-3213(03)00227-X. arXiv: hep-lat/0211042 [hep-lat].

[133] R. Frezzotti and K. Jansen. "A Polynomial hybrid Monte Carlo algorithm". In: *Phys. Lett.* B402 (1997), pp. 328–334. DOI: 10.1016/S0370-2693(97)00475-9. arXiv: hep-lat/9702016 [hep-lat].

[134] M. A. Clark and A. D. Kennedy. "Accelerating dynamical fermion computations using the rational hybrid Monte Carlo (RHMC) algorithm with multiple pseudofermion fields". In: *Phys. Rev. Lett.* 98 (2007), p. 051601. DOI: 10.1103/PhysRevLett.98.051601. arXiv: hep-lat/0608015 [hep-lat].

[135] N. I. Achieser. *Theory of approximation*. Courier Corporation, 2013.

[136] E. Zolotarev. "Application of elliptic functions to questions of functions deviating least and most from zero". In: *Zap. Imp. Akad. Nauk. St. Petersburg* 30.5 (1877), pp. 1–59.

[137] M. Lüscher and S. Schaefer. "Lattice QCD with open boundary conditions and twisted-mass reweighting". In: *Comput. Phys. Commun.* 184 (2013), pp. 519–528. DOI: 10.1016/j.cpc.2012.10.003. arXiv: 1206.2809 [hep-lat].

[138] A. Frommer, S. Güttel, and M. Schweitzer. "Efficient and stable Arnoldi restarts for matrix functions based on quadrature". In: *SIAM Journal on Matrix Analysis and Applications* 35.2 (2014), pp. 661–683. ISSN: 0895-4798. DOI: 10.1137/13093491X.

[139] A. Frommer, S. Güttel, and M. Schweitzer. "Convergence of restarted Krylov subspace methods for Stieltjes functions of matrices". In: *SIAM Journal on Matrix Analysis and Applications* 35.4 (2014), pp. 1602–1624. ISSN: 0895-4798. DOI: 10.1137/140973463.

[140] H. Neuberger. "Exactly massless quarks on the lattice". In: *Phys. Lett.* B417 (1998), pp. 141–144. DOI: 10.1016/S0370-2693(97)01368-3. arXiv: hep-lat/9707022 [hep-lat].

[141] B. Jegerlehner. "Krylov space solvers for shifted linear systems". In: (1996). arXiv: hep-lat/9612014 [hep-lat].

[142] U. Glassner et al. "How to compute Green's functions for entire mass trajectories within Krylov solvers". In: *Int. J. Mod. Phys.* C7 (1996), p. 635. DOI: 10.1142/S0129183196000533. arXiv: hep-lat/9605008 [hep-lat].

[143] A. Frommer and U. Glässner. "Restarted GMRES for Shifted Linear Systems". In: *SIAM Journal on Scientific Computing* 19.1 (1998), pp. 15–26. DOI: 10.1137/S1064827596304563. eprint: https://doi.org/10.1137/S1064827596304563. URL: https://doi.org/10.1137/S1064827596304563.

[144] S. J. Smith. "Lebesgue constants in polynomial interpolation". In: *Annales Mathematicae et Informaticae*. Vol. 33. 109-123. Eszterházy Károly College, Institute of Mathematics and Computer Science. 2006, pp. 1787–5021.

[145] A. Abdel-Rehim et al. "Nucleon and pion structure with lattice QCD simulations at physical value of the pion mass". In: *Phys. Rev.* D92.11 (2015). [Erratum: Phys. Rev.D93,no.3,039904(2016)], p. 114513. DOI: 10.1103/PhysRevD.92.114513,10.1103/PhysRevD.93.039904. arXiv: 1507.04936 [hep-lat].

[146] C. Helmes et al. "The η' meson at the physical point with $N_f = 2$ Wilson twisted mass fermions". In: *EPJ Web Conf.* 175 (2018), p. 05025. DOI: 10.1051/epjconf/201817505025. arXiv: 1710.03698 [hep-lat].

[147] A. Abdel-Rehim et al. "Direct Evaluation of the Quark Content of Nucleons from Lattice QCD at the Physical Point". In: *Phys. Rev. Lett.* 116.25 (2016), p. 252001. DOI: 10.1103/PhysRevLett.116.252001. arXiv: 1601.01624 [hep-lat].

[148] C. Alexandrou et al. "Nucleon axial form factors using $N_f = 2$ twisted mass fermions with a physical value of the pion mass". In: *Phys. Rev.* D96.5 (2017), p. 054507. DOI: 10.1103/PhysRevD.96.054507. arXiv: 1705.03399 [hep-lat].

[149] C. Alexandrou et al. "Nucleon electromagnetic form factors using lattice simulations at the physical point". In: *Phys. Rev.* D96.3 (2017), p. 034503. DOI: 10.1103/PhysRevD.96.034503. arXiv: 1706.00469 [hep-lat].

[150] C. Alexandrou et al. "Nucleon Spin and Momentum Decomposition Using Lattice QCD Simulations". In: *Phys. Rev. Lett.* 119.14 (2017), p. 142002. DOI: 10.1103/PhysRevLett.119.142002. arXiv: 1706.02973 [hep-lat].

[151] C. Urbach. "Reversibility Violation in the Hybrid Monte Carlo Algorithm". In: *Comput. Phys. Commun.* 224 (2018), pp. 44–51. DOI: 10.1016/j.cpc.2017.12.005. arXiv: 1710.07526 [hep-lat].

[152] R. Baron et al. "Computing K and D meson masses with $N_f = 2+1+1$ twisted mass lattice QCD". In: *Comput. Phys. Commun.* 182 (2011), pp. 299–316. DOI: 10.1016/j.cpc.2010.10.004. arXiv: 1005.2042 [hep-lat].

[153] N. Carrasco et al. "Up, down, strange and charm quark masses with $N_f = 2+1+1$ twisted mass lattice QCD". In: *Nucl. Phys.* B887 (2014), pp. 19–68. DOI: 10.1016/j.nuclphysb.2014.07.025. arXiv: 1403.4504 [hep-lat].

[154] P. Boucaud et al. "Dynamical Twisted Mass Fermions with Light Quarks: Simulation and Analysis Details". In: *Comput. Phys. Commun.* 179 (2008), pp. 695–715. DOI: 10.1016/j.cpc.2008.06.013. arXiv: 0803.0224 [hep-lat].

[155] K. Osterwalder and E. Seiler. "Gauge Field Theories on the Lattice". In: *Annals Phys.* 110 (1978), p. 440. DOI: 10.1016/0003-4916(78)90039-8.

[156] S. Aoki et al. "Review of lattice results concerning low-energy particle physics". In: *Eur. Phys. J.* C77.2 (2017), p. 112. DOI: 10.1140/epjc/s10052-016-4509-7. arXiv: 1607.00299 [hep-lat].

[157] G. Colangelo, S. Durr, and C. Haefeli. "Finite volume effects for meson masses and decay constants". In: *Nucl. Phys.* B721 (2005), pp. 136–174. DOI: 10.1016/j.nuclphysb.2005.05.015. arXiv: hep-lat/0503014 [hep-lat].

[158] G. Colangelo, U. Wenger, and J. M. S. Wu. "Twisted Mass Finite Volume Effects". In: *Phys. Rev.* D82 (2010), p. 034502. DOI: 10.1103/PhysRevD.82.034502. arXiv: 1003.0847 [hep-lat].

[159] S. Weinberg. "Phenomenological Lagrangians". In: *Physica* A96.1-2 (1979), pp. 327–340. DOI: 10.1016/0378-4371(79)90223-1.

[160] J. Gasser and H. Leutwyler. "Chiral Perturbation Theory to One Loop". In: *Annals Phys.* 158 (1984), p. 142. DOI: 10.1016/0003-4916(84)90242-2.

[161] J. Gasser and H. Leutwyler. "Chiral Perturbation Theory: Expansions in the Mass of the Strange Quark". In: *Nucl. Phys.* B250 (1985), pp. 465–516. DOI: 10.1016/0550-3213(85)90492-4.

[162] A. Stathopoulos, J. Laeuchli, and K. Orginos. "Hierarchical probing for estimating the trace of the matrix inverse on toroidal lattices". In: (2013). arXiv: 1302.4018 [hep-lat].

[163] A. Abdel-Rehim et al. "Disconnected diagrams with twisted-mass fermions". In: *PoS* LATTICE2016 (2016), p. 155. DOI: 10.22323/1.256.0155. arXiv: 1611.03802 [hep-lat].

[164] K. Ottnad and C. Urbach. "Flavor-singlet meson decay constants from $N_f = 2+1+1$ twisted mass lattice QCD". In: *Phys. Rev.* D97.5 (2018), p. 054508. DOI: 10.1103/PhysRevD.97.054508. arXiv: 1710.07986 [hep-lat].

[165] R. Baron et al. "Light hadrons from lattice QCD with light (u,d), strange and charm dynamical quarks". In: *JHEP* 06 (2010), p. 111. DOI: 10.1007/JHEP06(2010)111. arXiv: 1004.5284 [hep-lat].

[166] M. Albanese et al. "Glueball Masses and String Tension in Lattice QCD". In: *Phys. Lett.* B192 (1987), pp. 163–169. DOI: 10.1016/0370-2693(87)91160-9.

[167] S. Gusken. "A Study of smearing techniques for hadron correlation functions". In: *Nucl. Phys. Proc. Suppl.* 17 (1990), pp. 361–364. DOI: 10.1016/0920-5632(90)90273-W.

[168] C. Alexandrou et al. "The Static approximation of heavy - light quark systems: A Systematic lattice study". In: *Nucl. Phys.* B414 (1994), pp. 815–855. DOI: 10.1016/0550-3213(94)90262-3. arXiv: hep-lat/9211042 [hep-lat].

[169] J. Gasser, M. E. Sainio, and A. Svarc. "Nucleons with Chiral Loops". In: *Nucl. Phys.* B307 (1988), pp. 779–853. DOI: 10.1016/0550-3213(88)90108-3.

[170] B. C. Tiburzi and A. Walker-Loud. "Hyperons in Two Flavor Chiral Perturbation Theory". In: *Phys. Lett.* B669 (2008), pp. 246–253. DOI: 10.1016/j.physletb.2008.09.054. arXiv: 0808.0482 [nucl-th].

[171] C. Alexandrou et al. "Light baryon masses with dynamical twisted mass fermions". In: *Phys. Rev.* D78 (2008), p. 014509. DOI: 10.1103/PhysRevD.78.014509. arXiv: 0803.3190 [hep-lat].

[172] M. Lüscher. "Properties and uses of the Wilson flow in lattice QCD". In: *JHEP* 08 (2010). [Erratum: JHEP03,092(2014)], p. 071. DOI: 10.1007/JHEP08(2010)071, 10.1007/JHEP03(2014)092. arXiv: 1006.4518 [hep-lat].

[173] S. Borsanyi et al. "High-precision scale setting in lattice QCD". In: *JHEP* 09 (2012), p. 010. DOI: 10.1007/JHEP09(2012)010. arXiv: 1203.4469 [hep-lat].

[174] O. Bar and M. Golterman. "Chiral perturbation theory for gradient flow observables". In: *Phys. Rev.* D89.3 (2014). [Erratum: Phys. Rev.D89,no.9,099905(2014)], p. 034505. DOI: 10.1103/PhysRevD.89.099905, 10.1103/PhysRevD.89.034505. arXiv: 1312.4999 [hep-lat].

[175] M. Schmelling. "Averaging correlated data". In: *Physica Scripta* 51.6 (June 1995), pp. 676–679. DOI: 10.1088/0031-8949/51/6/002. URL: https://doi.org/10.1088%2F0031-8949%2F51%2F6%2F002.

[176] M. Bruno, S. Schaefer, and R. Sommer. "Topological susceptibility and the sampling of field space in $N_f = 2$ lattice QCD simulations". In: *JHEP* 08 (2014), p. 150. DOI: 10.1007/JHEP08(2014)150. arXiv: 1406.5363 [hep-lat].

[177] U. Wolff. "Monte Carlo errors with less errors". In: *Comput. Phys. Commun.* 156 (2004). [Erratum: Comput. Phys. Commun.176,383(2007)], pp. 143–153. DOI: 10.1016/S0010-4655(03)00467-3, 10.1016/j.cpc.2006.12.001. arXiv: hep-lat/0306017 [hep-lat].

[178] R. Frezzotti, V. Lubicz, and S. Simula. "Electromagnetic form factor of the pion from twisted-mass lattice QCD at N(f) = 2". In: *Phys. Rev.* D79 (2009), p. 074506. DOI: 10.1103/PhysRevD.79.074506. arXiv: 0812.4042 [hep-lat].

[179] S. R. Amendolia et al. "A Measurement of the Space - Like Pion Electromagnetic Form-Factor". In: *Nucl. Phys.* B277 (1986), p. 168. DOI: 10.1016/0550-3213(86)90437-2.

[180] J. Bijnens, G. Colangelo, and P. Talavera. "The Vector and scalar form-factors of the pion to two loops". In: *JHEP* 05 (1998), p. 014. DOI: 10.1088/1126-6708/1998/05/014. arXiv: hep-ph/9805389 [hep-ph].

[181] F.-J. Jiang and B. C. Tiburzi. "Flavor Twisted Boundary Conditions in the Breit Frame". In: *Phys. Rev.* D78 (2008), p. 037501. DOI: 10.1103/PhysRevD.78.037501. arXiv: 0806.4371 [hep-lat].

[182] G. Colangelo and A. Vaghi. "Pseudoscalar mesons in a finite cubic volume with twisted boundary conditions". In: *JHEP* 07 (2016), p. 134. DOI: 10.1007/JHEP07(2016)134. arXiv: 1607.00916 [hep-lat].

[183] B. B. Brandt, A. Jüttner, and H. Wittig. "The pion vector form factor from lattice QCD and NNLO chiral perturbation theory". In: *JHEP* 11 (2013), p. 034. DOI: 10.1007/JHEP11(2013)034. arXiv: 1306.2916 [hep-lat].

[184] J. Koponen et al. "Size of the pion from full lattice QCD with physical u , d , s and c quarks". In: *Phys. Rev.* D93.5 (2016), p. 054503. DOI: 10.1103/PhysRevD.93.054503. arXiv: 1511.07382 [hep-lat].

[185] J. Golak et al. "Extraction of electromagnetic neutron form-factors through inclusive and exclusive polarized electron scattering on polarized He-3 target". In: *Phys. Rev.* C63 (2001), p. 034006. DOI: 10.1103/PhysRevC.63.034006. arXiv: nucl-th/0008008 [nucl-th].

[186] J. Becker et al. "Determination of the neutron electric form-factor from the reaction He-3(e,e' n) at medium momentum transfer". In: *Eur. Phys. J.* A6 (1999), pp. 329–344. DOI: 10.1007/s100500050351.

[187] T. Eden et al. "Electric form factor of the neutron from the $^2H(\vec{e},e'\vec{n})^1H$ reaction at $Q^2 = 0.255$ (GeV/c)^2". In: *Phys. Rev.* C50.4 (1994), R1749–R1753. DOI: 10.1103/PhysRevC.50.R1749.

[188] M. Meyerhoff et al. "First measurement of the electric form-factor of the neutron in the exclusive quasielastic scattering of polarized electrons from polarized He-3". In: *Phys. Lett.* B327 (1994), pp. 201–207. DOI: 10.1016/0370-2693(94)90718-8.

[189] I. Passchier et al. "The Charge form-factor of the neutron from the reaction polarized H-2(polarized e, e-prime n) p". In: *Phys. Rev. Lett.* 82 (1999), pp. 4988–4991. DOI: 10.1103/PhysRevLett.82.4988. arXiv: nucl-ex/9907012 [nucl-ex].

[190] G. Warren et al. "Measurement of the electric form-factor of the neutron at $Q^2 = 0.5$ and 1.0 GeV^2/c^2". In: *Phys. Rev. Lett.* 92 (2004), p. 042301. DOI: 10.1103/PhysRevLett.92.042301. arXiv: nucl-ex/0308021 [nucl-ex].

[191] H. Zhu et al. "A Measurement of the electric form-factor of the neutron through polarized-d (polarized-e, e-prime n)p at Q**2 = 0.5-(GeV/c)**2". In: *Phys. Rev. Lett.* 87 (2001), p. 081801. DOI: 10.1103/PhysRevLett.87.081801. arXiv: nucl-ex/0105001 [nucl-ex].

[192] B. Plaster et al. "Measurements of the neutron electric to magnetic form-factor ratio G(En) / G(Mn) via the H-2(polarized-e, e-prime,polarized-n)H-1 reaction to Q**2 = 1.45-(GeV/c)**2". In: *Phys. Rev.* C73 (2006), p. 025205. DOI: 10.1103/PhysRevC.73.025205. arXiv: nucl-ex/0511025 [nucl-ex].

[193] R. Madey et al. "Measurements of G(E)n / G(M)n from the H-2(polarized-e,e-prime polarized-n) reaction to Q**2 = 1.45 (GeV/c)**2". In: *Phys. Rev. Lett.* 91 (2003), p. 122002. DOI: 10.1103/PhysRevLett.91.122002. arXiv: nucl-ex/0308007 [nucl-ex].

[194] D. Rohe et al. "Measurement of the neutron electric form-factor G(en) at 0.67-(GeV/c)**2 via He-3(pol.)(e(pol.),e' n)". In: *Phys. Rev. Lett.* 83 (1999), pp. 4257–4260. DOI: 10.1103/PhysRevLett.83.4257.

[195] J. Bermuth et al. "The Neutron charge form-factor and target analyzing powers from polarized-He-3 (polarized-e,e-prime n) scattering". In: *Phys. Lett.* B564 (2003), pp. 199–204. DOI: 10.1016/S0370-2693(03)00725-1. arXiv: nucl-ex/0303015 [nucl-ex].

[196] D. I. Glazier et al. "Measurement of the electric form-factor of the neutron at Q**2 = 0.3-(GeV/c)**2 to 0.8-(GeV/c)**2". In: *Eur. Phys. J.* A24 (2005), pp. 101–109. DOI: 10.1140/epja/i2004-10115-8. arXiv: nucl-ex/0410026 [nucl-ex].

[197]   C. Herberg et al. "Determination of the neutron electric form-factor in the D(e,e' n)p reaction and the influence of nuclear binding". In: *Eur. Phys. J.* A5 (1999), pp. 131–135. DOI: 10.1007/s100500050268.

[198]   R. Schiavilla and I. Sick. "Neutron charge form-factor at large q**2". In: *Phys. Rev.* C64 (2001), p. 041002. DOI: `10.1103/PhysRevC.64.041002`. arXiv: `nucl-ex/0107004` `[nucl-ex]`.

[199]   M. Ostrick et al. "Measurement of the neutron electric form-factor G(E,n) in the quasifree H-2(e(pol.),e' n(pol.))p reaction". In: *Phys. Rev. Lett.* 83 (1999), pp. 276–279. DOI: `10.1103/PhysRevLett.83.276`.

[200]   C. F. Perdrisat, V. Punjabi, and M. Vanderhaeghen. "Nucleon Electromagnetic Form Factors". In: *Prog. Part. Nucl. Phys.* 59 (2007), pp. 694–764. DOI: `10.1016/j.ppnp.2007.05.001`. arXiv: `hep-ph/0612014` `[hep-ph]`.

[201]   R. J. Hill and G. Paz. "Model independent extraction of the proton charge radius from electron scattering". In: *Phys. Rev.* D82 (2010), p. 113005. DOI: `10.1103/PhysRevD.82.113005`. arXiv: `1008.4619` `[hep-ph]`.

[202]   J. R. Green et al. "Nucleon electromagnetic form factors from lattice QCD using a nearly physical pion mass". In: *Phys. Rev.* D90 (2014), p. 074507. DOI: `10.1103/PhysRevD.90.074507`. arXiv: `1404.4029` `[hep-lat]`.

[203]   N. Hasan et al. "Computing the nucleon charge and axial radii directly at $Q^2 = 0$ in lattice QCD". In: *Phys. Rev.* D97.3 (2018), p. 034504. DOI: `10.1103/PhysRevD.97.034504`. arXiv: `1711.11385` `[hep-lat]`.

[204]   K.-I. Ishikawa et al. "Nucleon form factors on a large volume lattice near the physical point in 2+1 flavor QCD". In: *Phys. Rev.* D98.7 (2018), p. 074510. DOI: `10.1103/PhysRevD.98.074510`. arXiv: `1807.03974` `[hep-lat]`.

[205]   H. Yin and R. D. Mawhinney. "Improving DWF Simulations: the Force Gradient Integrator and the Móbius Accelerated DWF Solver". In: *PoS* LATTICE2011 (2011), p. 051. DOI: `10.22323/1.139.0051`. arXiv: `1111.5059` `[hep-lat]`.