

Wilson Twisted Mass Fermions: Towards Realistic Simulations of Lattice-QCD

Carsten Urbach

September 2005

**Am Fachbereich Physik
der Freien Universität Berlin
eingereichte Dissertation**

Erster Gutachter: Prof. Dr. Volkard Linke

Zweiter Gutachter: Priv.-Doz. Dr. Karl Jansen

Tag der Disputation: 12. Dezember 2005

Contents

Introduction	1
1 Theoretical basis	5
1.1 QCD in Euclidean space-time	5
1.2 Lattice regularization	9
1.2.1 Wilson lattice action	9
1.2.2 Remnant chiral symmetry on the lattice	11
1.2.3 Symanzik effective action	14
1.2.4 Twisted mass regularization	16
1.3 Observables	21
1.3.1 Wilson formulation	22
1.3.2 tmQCD formulation	24
1.3.3 Overlap formulation	26
1.3.4 Setting the scale	28
1.4 Lattice simulations	30
1.4.1 Quenched approximation	31
1.4.2 Dynamical simulations	32
1.4.3 Error estimates	32
2 Quenched scaling test	35
2.1 The definition of the critical mass	36
2.2 Scaling test set-up	39
2.2.1 Finite volume effects	40
2.3 Scaling test of mtmQCD	41
2.3.1 Scaling of the pseudo scalar decay constant with the pion definition of κ_{crit}	42
2.3.2 Critical mass from the PCAC relation	43
2.3.3 Scaling of f_{PS} with the PCAC definition of κ_{crit}	44

2.3.4	Scaling of the vector meson mass with the PCAC definition of κ_{crit}	45
2.3.5	Continuum extrapolation	46
2.3.6	Flavor breaking effects	48
2.4	Overlap versus twisted mass fermions	50
2.4.1	Cost comparison	54
2.5	Conclusion	57
3	Accelerating the HMC algorithm	59
3.1	HMC algorithm	60
3.1.1	Molecular dynamics evolution	62
3.1.2	Integration with multiple time scales	62
3.2	Mass Preconditioning	64
3.2.1	Mass preconditioning and multiple time scale integration . . .	66
3.3	Numerical results	66
3.3.1	Simulation points	66
3.3.2	Details of the implementation	67
3.3.3	Force contributions	69
3.3.4	Tuning the algorithm	71
3.3.5	Algorithm performance	73
3.3.6	Simulation cost	75
3.3.7	Scaling with the mass	77
3.4	Conclusion	79
4	Phase structure of lattice QCD	83
4.1	Effective potential model	84
4.1.1	Chiral perturbation theory on the lattice	85
4.2	Numerical results	89
4.2.1	Simulation points	89
4.2.2	Thermal cycles	89
4.2.3	Meta-stabilities	90
4.2.4	Pseudo scalar and quark masses	92
4.2.5	The phase transition as a function of the lattice spacing	94
4.2.6	Changing the gauge action	98
4.3	Conclusion	100
	Summary and Outlook	103

A Conventions	107
A.1 Dirac matrices	107
A.2 Generators of SU(3)	108
B Algorithmic details	109
B.1 Even/odd preconditioning	109
B.2 Multiple mass solver for twisted mass fermions	112
C Tables	114
Bibliography	117
List of publications	130

Introduction

The notion of local gauge symmetries as introduced in 1929 by Weyl [1] turned out to be a cornerstone of modern field theory since so-called gauge theories describe fundamental interactions very successfully. The prototype of a gauge theory is quantum electrodynamics (QED), which describes nature with a so far unknown precision. An impressive example is the anomalous magnetic moment of the muon g_μ , where the experimental results deviate from the theoretical prediction for $a_\mu \equiv (g_\mu - 2)/2$ at the order of 10^{-9} by maximally 2σ [2]. Our present understanding is that all fundamental interactions, strong interaction, electromagnetic interaction, weak interaction and gravitational interaction, are described by some form of a gauge theory.

Quantum chromodynamics (QCD), the theory of strong interactions, is based on a non-Abelian SU(3) gauge symmetry. The property of QCD that led directly to its discovery in 1973 [3, 4, 5] as a candidate theory of the strong interaction is asymptotic freedom [3, 6, 7, 8, 9], i.e. the coupling strength decreases at short distances and the quarks and gluons behave as effectively free particles. In turn, the coupling increases with the distance and at a distance of about 1 fm it assumes such large values that only bound states of quarks and gluons exist. The latter property, which is called confinement, together with asymptotic freedom imply that perturbative methods are applicable in QCD only at short distances, whereas they fail at large distances (low energies). Confinement, spontaneous breaking of chiral symmetry or the hadron mass spectrum are low-energy properties that can therefore not be described by perturbation theory and require a non-perturbative treatment of the theory.

With the framework of *lattice gauge theories* Wilson developed a non-perturbative tool to investigate the low-energy structure of QCD [10]. In this framework he was able to compute the non-relativistic quark/anti-quark potential in the static approximation showing that it increases linearly with the distance [10, 11]. Wilson proposed to regularize QCD with a discrete Euclidean space-time lattice with the inverse lattice spacing a^{-1} playing the rôle of an ultraviolet momentum cut-off. Then, in the course of renormalization the continuum is recovered by removing the cut-off, i.e. sending $a \rightarrow 0$. This approach can also be understood as replacing the continuum

INTRODUCTION

gauge theory by a discrete statistical mechanical system, an analogy that opens up the possibility to simulate QCD on computers by means of Monte Carlo methods. This framework facilitates investigations of low-energy properties of QCD from first principles with the quark masses as freely tunable parameters.

Unfortunately such computer simulations are only possible with an immense amount of computer resources. For this reason most of the current results have been obtained only in the so-called *quenched approximation*, where vacuum polarization effects of quark loops are neglected. However, even in this approximation simulations with small enough values of the quark masses are not affordable necessitating – besides the continuum extrapolation of physical quantities – an additional extrapolation in the masses to the point where the masses take their physical values. If the simulations can reach masses where chiral perturbation theory (χ PT) is valid, then this last step of the extrapolation can be performed using the analytical results derived in χ PT.

In QCD exist six flavors of quarks, whose (bare) quark masses are parameters that need to be tuned. But, aiming at investigations of the low energy structure of QCD, charm, bottom and top quarks can be considered as static to a good approximation due to their large mass values. Moreover, the two lightest quarks (up and down) are – when compared to the characteristic scale of QCD – to a good approximation mass degenerate. Hence, the targets of lattice QCD are simulations with a light doublet of mass degenerate quarks and one heavier quark, the strange quark.

While for the extrapolation in the quark mass χ PT can be of essential help, the continuum extrapolation can only be performed by using small values of the lattice spacing a , such that one is close enough to the continuum limit. However, the computational costs increase approximately proportional to a^{-7} making it often infeasible to work at small enough lattice spacing. The way out is the fact that lattice QCD formulations are not unique and fortunately a formulation with genuine small lattice artifacts can be constructed by means of an effective field theory as worked out by Symanzik [12, 13, 14]. In this concept the lattice theory at finite values of a is mapped to an effective continuum theory. Lattice expectation values are then given by the corresponding continuum value plus correction terms proportional to powers of the lattice spacing. The *Symanzik improvement programme* then means to construct a discretization where at least the largest of those terms are absent. The simplest case in this approach is the $\mathcal{O}(a)$ improvement, where all terms linear in a vanish. However, lattice artifacts proportional to higher powers of a might still be large. This, together with the question for which values of a the effective theory is valid, is one of the crucial questions in lattice QCD and needs to be investigated through a detailed analysis of the scaling behavior in a of physical quantities.

All this illustrates the need for progress on the following topics:

- Formulations of lattice QCD with reduced lattice discretization errors are essential in order to control the continuum extrapolation reliably. Those formulations are available, but they need a test in practice, which concerns the size of residual lattice artifacts on physical observables. In addition, it is necessary to investigate how a lattice theory as such differs from its continuum counterpart.
- A lattice QCD formulation should allow for simulations at small enough quark masses with affordable computational effort in order to make at least contact to chiral perturbation theory possible. Ideally, it would be desirable, of course, to work directly at the physical point. Moreover, simulations with two light mass degenerate quarks (up and down quark) and one heavier quark (strange quark) should be possible.
- Improvement and development of new algorithms is needed to reach small enough quark masses and small enough lattice spacings at possibly lower computational effort.

Motivated by these demands, we consider in this work the so-called *Wilson twisted mass* formulation of lattice QCD, which is expected to satisfy the requirements formulated above by the first two items. We present a detailed scaling test of this formulation in the quenched approximation and show that in fact lattice artifacts linear in a are absent and residual lattice artifacts are small (see chapter 2).

Then, in chapter 3, we introduce an algorithm for simulations of full QCD with scaling properties towards small quark masses that are significantly better than those of other presently used algorithms. This improvement is illustrated by comparing it to other state-of-the-art algorithms available in the literature.

In chapter 4 we finally present a study of the phase structure of lattice QCD with two flavors of Wilson twisted mass fermions and several discretizations of the gauge part in the action. This investigation is an essential preparatory work for any future large scale simulation and reveals evidence for the existence of a first order phase transition. A comprehensive understanding of the phase structure was missing so far, and became only possible with the Wilson twisted mass formulation of lattice QCD.

INTRODUCTION

Chapter 1

Theoretical basis

This chapter is devoted to give a summary of the basic theoretical concepts used for this work. We start with shortly introducing quantum chromodynamics (QCD) as the theory of strong interactions followed by a description of its regularization by the standard Wilson lattice QCD approach, as it can be found in textbooks such as [15, 16, 17, 18, 19, 20]. Then we discuss in more detail several more recent formulations of lattice QCD and compare their properties. Aiming at computer simulations of lattice QCD we finally present methods and algorithms for this purpose.

1.1 QCD in Euclidean space-time

In this section we shortly derive the SU(3) gauge invariant QCD Lagrangian, describe its quantization by means of the path integral formalism and the transition from Minkowski to Euclidean space-time. The QCD Lagrangian is constructed such that the standard lattice regularization follows immediately.

Motivated by the demand of local gauge symmetry, we consider a continuous group of (gauge) transformations, represented by a set of SU(3) matrices $V(x)$, and a quark of flavor q , represented by a color-spinor field $\psi_{c,q}^\alpha(x)$ with SU(3) color index $c = 1, 2, 3$ and Dirac spinor index $\alpha = 0, 1, 2, 3$. Suppressing all these indices in the following, the fields $\psi(x)$ transform like

$$\psi(x) \rightarrow V(x)\psi(x). \quad (1-1)$$

In order to construct a gauge invariant kinetic energy term for these quark fields, we define the gauge *covariant* derivative in direction μ by the limiting procedure:

$$D_\mu\psi(x) = \lim_{a \rightarrow 0} \frac{1}{a} [U(x, x + a\hat{\mu})\psi(x + a\hat{\mu}) - \psi(x)], \quad (1-2)$$

where $\hat{\mu}$ is a unit vector in direction μ^* . For later purposes we extract the forward covariant difference operator

$$\nabla_{\mu}\psi(x) = \frac{1}{a}[U(x, x + a\hat{\mu})\psi(x + a\hat{\mu}) - \psi(x)], \quad (1-3)$$

and define also the backward covariant difference operator

$$\nabla_{\mu}^*\psi(x) = \frac{1}{a}[\psi(x) - U(x, x - a\hat{\mu})\psi(x - a\hat{\mu})]. \quad (1-4)$$

In definition (1-2) we have introduced a so-called parallel transporter $U(x, y)$, a unitary SU(3) matrix, which obeys the transformation law

$$U(x, y) \rightarrow V(x)U(x, y)V^{\dagger}(y), \quad (1-5)$$

and we set $U(x, x) = 1$. Since $U(x, y)$ is a SU(3) matrix and a continuous function of its arguments, for infinitesimal a it can be represented as

$$U(x, x + a\hat{\mu}) = \exp\left(-igaA_{\mu}^i(x + \frac{a}{2}\hat{\mu})\lambda^i + \mathcal{O}(a^3)\right). \quad (1-6)$$

Here g is the bare gauge coupling, introduced for later convenience, and A_{μ}^i is a vector field for each generator λ^i of the transformation group. Expanding (1-6) in a and inserting it into definition (1-2) yields in the limit $a \rightarrow 0$ for the continuum covariant derivative associated with the local SU(3) gauge symmetry

$$D_{\mu} = \partial_{\mu} - igA_{\mu}^i\lambda^i, \quad (1-7)$$

which has the correct gauge transformation properties and where A_{μ}^i is the continuum *gauge potential*. With what we have defined so far we can already write a locally gauge invariant Lagrangian containing a kinetic energy term for the quark field $\psi(x)$ and a quark mass term. In order to complete the Lagrangian we have to find a kinetic term for the field A_{μ} itself. A term like this can be constructed by defining

$$U_{\square}(x; \mu, \nu) = U(x, x + a\hat{\mu})U(x + a\hat{\mu}, x + a\hat{\nu} + a\hat{\mu}) \\ \times U(x + a\hat{\nu} + a\hat{\mu}, x + a\hat{\nu})U(x + a\hat{\nu}, x) \quad (1-8)$$

as the product of the four parallel transporters at the corners of a small square in space-time, the *plaquette*. However, $U_{\square}(x; \mu, \nu)$ itself is not gauge invariant, and in order to convert it to a locally gauge invariant expression we have to take the trace. Inserting (1-6) into (1-8) and using the Campbell-Baker-Hausdorff formula we achieve by expanding to order a^6

$$\text{Tr} U_{\square}(x; \mu, \nu) = \text{Tr}[1] - \frac{1}{4}g^2a^4(F_{\mu\nu}^i)^2 + \mathcal{O}(a^6), \quad (1-9)$$

*We use the non-standard notation a (instead of for instance ϵ), because a will be identified with the lattice spacing later on.

where we introduced the field strength

$$F_{\mu\nu}^i = \partial_\mu A_\nu^i - \partial_\nu A_\mu^i + gf^{ijk}A_\mu^j A_\nu^k. \quad (1-10)$$

f^{ijk} are the *structure constants* of the symmetry group (see appendix A). By construction $(F_{\mu\nu}^i)^2$ is locally gauge invariant. Now we can write down the QCD Lagrangian that is renormalizable, conserves parity and is invariant under time reversal [3, 4, 5]

$$\mathcal{L} = -\frac{1}{4}(F_{\mu\nu}^i)^2 + \bar{\psi}(i\gamma^\mu D_\mu - m_q)\psi, \quad (1-11)$$

with $\bar{\psi} = \psi^\dagger \gamma_0$ being the anti-quark field. The fermionic part of Eq. (1-11) contains the covariant Dirac operator $M \equiv i\gamma^\mu D_\mu - m_q$ with quark mass m_q [†].

Eq. (1-11) exhibits the classical QCD Lagrangian. In order to quantize the theory we will use the path integral formalism introduced by Feynman [21]. The fundamental quantity in the path integral formalism is the classical action S , which is explicitly given for QCD with one quark flavor as the space-time integral over the Lagrangian density

$$S[\psi, \bar{\psi}, A] = \int d^4x \left\{ -\frac{1}{4}(F_{\mu\nu}^i)^2 + \bar{\psi}M\psi \right\}. \quad (1-12)$$

The expectation value of a physical observable O is then formally given by the following functional integral:

$$\langle O \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}A \mathcal{D}\psi \mathcal{D}\bar{\psi} O[\psi, \bar{\psi}, A] e^{iS[\psi, \bar{\psi}, A]}, \quad (1-13)$$

with the *partition function* \mathcal{Z}

$$\mathcal{Z} = \int \mathcal{D}A \mathcal{D}\psi \mathcal{D}\bar{\psi} e^{iS[\psi, \bar{\psi}, A]}. \quad (1-14)$$

One advantage of the path integral quantization is that it deals only with classical fields and not with operators. However, in case of fermions the (classical) fields ψ must be represented by anti-commuting variables. This is realized by the use of so called *Grassmann variables*. The integral over such Grassmann valued numbers can be defined in a sensible way and we adopt here the standard conventions (see for instance [15]).

Formally we can perform the integral over the Grassmann valued fields ψ and $\bar{\psi}$ analytically due to its Gaussian structure

$$\int \mathcal{D}\psi \mathcal{D}\bar{\psi} \exp \left\{ i \int d^4x \bar{\psi}M\psi \right\} \propto \det M, \quad (1-15)$$

[†]For simplicity we consider only one quark flavor. The generalization to N_f flavors of quarks is straightforward.

leaving us to calculate the functional determinant of the operator M^\ddagger . Nevertheless, we can define an effective action

$$S_{\text{eff}}[A] = \int d^4x \left\{ -i \log \det M - \frac{1}{4} (F_{\mu\nu}^i)^2 \right\}. \quad (1-16)$$

in terms of which we can now express the expectation value of an operator O

$$\langle O \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}A O[A] e^{iS_{\text{eff}}[A]}. \quad (1-17)$$

This is possible because due to Wick's theorem for the contraction of fields one can eliminate the time-ordered product of quark fields in the operator O by suitable factors of M^{-1} which do no longer depend on the quark fields $\bar{\psi}, \psi$.

For a numerical treatment of the functional integral in Eq. (1-17) it is useful to analytically continue the time component of the 4-vectors to purely imaginary values[§]. This rotation of the time coordinate $x^0 \rightarrow -ix^0$ leads to a Euclidean 4-vector product:

$$x^2 = (x^0)^2 - |\vec{x}|^2 \quad \rightarrow \quad -(x^0)^2 - |\vec{x}|^2 = -|x_E|^2. \quad (1-18)$$

It was shown by Osterwalder and Schrader [22] that under certain conditions one can reconstruct the whole quantum field theory in Minkowski space from the Euclidean field theory. The most important condition is the so called *Osterwalder-Schrader positivity* or *reflection positivity*, which replaces Hilbert space positivity and the spectral condition of the Minkowskian formulation [22].

In Euclidean space the action is mapped to its Euclidean version

$$S \rightarrow iS \equiv -S_E = - \int d^4x \left[\frac{1}{4} (F_{\mu\nu}^i)^2 + \bar{\psi} (\gamma_\mu D_\mu + m_q) \psi \right], \quad (1-19)$$

where the Euclidean Dirac matrices γ_μ are hermitian and satisfy the anti-commutation relation

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}. \quad (1-20)$$

For the explicit representation see appendix A.1. Consequently the exponential weight, e.g. in the partition function \mathcal{Z} , can now be interpreted as a Boltzmann factor, as the Euclidean action is real and bounded from below (if the functional determinant $\det M$ is real)

$$\mathcal{Z}_E = \int \mathcal{D}\bar{\psi} \mathcal{D}\psi \mathcal{D}A e^{-S_E}. \quad (1-21)$$

Since we will work almost solely in Euclidean space-time we will skip the subscript _E in the following and use, for convenience, t instead of x_0 .

[‡]One can show that the functional determinant is equivalent to the sum of vacuum diagrams.

[§]This is often called Wick rotation.

1.2 Lattice regularization

In the following section we will describe how QCD as a quantum field theory can be regularized in Euclidean space-time by means of a discrete space-time lattice. As a regularization scheme it can be used equally well as e.g. dimensional regularization. But it has the crucial advantage to allow for non-perturbative computations by means of Monte Carlo methods.

To this end, we introduce a hyper-cubic space-time lattice with spacing a and extension aL (aT) in the spatial directions (time direction). The boundary conditions can be chosen in different ways, and we will use periodic boundary conditions throughout this work. Only for the quark fields in time direction we use either periodic or anti-periodic boundary conditions. The parallel transporters $U(x, y)$ are SU(3)-valued and defined as in Eq. (1-6), where we use the notation $U_{x,\mu} \equiv U(x, x + a\hat{\mu})$. Note that $U_{x,-\mu} \equiv U(x, x - a\hat{\mu}) = U_{x-a\hat{\mu},\mu}^{-1}$. The set of all parallel transporters on the lattice $U \equiv \{U_{x,\mu}\}$ we call a *gauge field configuration*.

The finite lattice spacing provides an ultra-violet cutoff for the momenta. In the finite volume the allowed momenta are then given by

$$k = \pm \frac{2\pi n}{La}, \quad n = 1, \dots, L/2. \quad (1-22)$$

On lattice we are now able to specify what is meant by the functional integral over the gauge fields:

$$\int \mathcal{D}A \equiv \int \prod_{x,\mu} dU_{x,\mu}, \quad (1-23)$$

where the product is over all lattice points x and directions μ . Unlike in the continuum, the lattice gauge fields are SU(3) matrices with elements that are bounded in the range $[0, 1]$. Therefore, as proposed by Wilson, we use the invariant *Haar measure* for the integration and adopt the standard definitions (cf. e.g. [18]).

The path integral of the Grassmann valued fermionic fields $\bar{\psi}$ and ψ is discretized similarly by

$$\int \mathcal{D}\bar{\psi} \mathcal{D}\psi \equiv \int \prod_x d\bar{\psi}(x) d\psi(x). \quad (1-24)$$

1.2.1 Wilson lattice action

The next step is to discretize the action. Recalling the definition of the plaquette variable (1-8) the definition

$$S_G[U] = \sum_x \sum_{1 \leq \mu < \nu} \beta \left\{ 1 - \frac{1}{3} \text{Re Tr}(U_{\square}(x; \mu, \nu)) \right\} \quad (1-25)$$

is locally gauge invariant, real and positive for $\beta \equiv 6/g^2 > 0$. It is obvious from Eq. (1-9) that it has the correct limit as $a \rightarrow 0$. S_G is called the Wilson plaquette gauge action [10] and in the following we will call β the bare coupling constant.

We are left to discretize the fermionic part of the action. With the definitions (1-3) and (1-4) of the forward and backward difference operators ∇_μ and ∇_μ^* , respectively, the lattice Dirac operator D_W proposed by Wilson in Ref. [10] can be read off from

$$\bar{\psi} D_W[U] \psi = a^4 \sum_{x,\mu} \bar{\psi}(x) \frac{1}{2} \left[\gamma_\mu (\nabla_\mu + \nabla_\mu^*) - ar \nabla_\mu^* \nabla_\mu \right] \psi(x). \quad (1-26)$$

The so-called Wilson term $-ar \nabla_\mu^* \nabla_\mu$ has been introduced in order to cure the problem that a naive discretization gives in the continuum limit rise to $2^{d-4} = 16$ fermion excitations rather than one (the so-called fermion doubling problem). With the Wilson term the extra 15 species pick up a mass proportional to $2r/a$ and decouple in the limit $a \rightarrow 0$. The Wilson parameter r can be chosen as $-1 \leq r \leq 1$. With a bare mass m_0 the Wilson lattice action then reads

$$S[\psi, \bar{\psi}, U] = S_G[U] + \bar{\psi} (D_W[U] - m_0) \psi. \quad (1-27)$$

However, while curing the fermion doubling problem, the Wilson term explicitly breaks chiral symmetry at finite values of the lattice spacing, i.e. even the massless Wilson-Dirac operator no longer anti-commutes with γ_5 . Although chiral symmetry is expected to be recovered in the continuum, its breaking at finite values of a has important consequences, among others:

- It introduces discretization errors of $\mathcal{O}(a)$. With a naive discretization the lattice artifacts in the action appear only at $\mathcal{O}(a^2)$.
- The quark mass renormalizes both additively and multiplicatively. Hence, it is useful to define a subtracted bare quark mass parameter by

$$m_q = m_0 - m_{\text{crit}}, \quad (1-28)$$

where m_{crit} is called the critical mass parameter. The value of m_{crit} needs to be determined such that the chiral point is properly defined, for instance by demanding that the lightest pseudo scalar state becomes massless at this point.

Both of these two consequences will play an important rôle in this work.

We note in passing that usually in the simulations the *hopping parameter representation* of the Wilson lattice action (1-27) is used. It is obtained from (1-27) by re-scaling the fermionic fields as follows:

$$\psi \rightarrow \frac{\sqrt{2\kappa}}{a^{3/2}} \psi, \quad \bar{\psi} \rightarrow \frac{\sqrt{2\kappa}}{a^{3/2}} \bar{\psi}, \quad \kappa = \frac{1}{2am_0 + 8r}, \quad (1-29)$$

with the so called *hopping parameter* κ . To m_{crit} corresponds the critical hopping parameter

$$\kappa_{\text{crit}} = \frac{1}{2am_{\text{crit}} + 8r}, \quad (1-30)$$

and the fermionic part of the action reads

$$\begin{aligned} S[\psi, \bar{\psi}, U] &= \sum_x \left\{ \bar{\psi}(x)\psi(x) - \kappa \bar{\psi}(x) \sum_{\mu=1}^4 \left[U_{x,\mu}(r + \gamma_\mu)\psi(x + a\hat{\mu}) \right. \right. \\ &\quad \left. \left. + U_{x-a\hat{\mu},\mu}^\dagger(r - \gamma_\mu)\psi(x - a\hat{\mu}) \right] \right\} \\ &\equiv \sum_{x,y} \bar{\psi}(x)M_{xy}\psi(y). \end{aligned} \quad (1-31)$$

1.2.2 Remnant chiral symmetry on the lattice

If we consider massless continuum QCD with two quark flavors $u \equiv \psi_{c,u}^\alpha$ and $d \equiv \psi_{c,d}^\alpha$, the Lagrangian obeys isospin symmetry, the symmetry of an SU(2) unitary transformation mixing the u and d fields. But, since left- and right-handed quarks do not couple, this Lagrangian is actually symmetric under the separate unitary transformations

$$\begin{pmatrix} u \\ d \end{pmatrix}_L \equiv \frac{1 - \gamma_5}{2} \begin{pmatrix} u \\ d \end{pmatrix} \rightarrow U_L \begin{pmatrix} u \\ d \end{pmatrix}_L, \quad \begin{pmatrix} u \\ d \end{pmatrix}_R \equiv \frac{1 + \gamma_5}{2} \begin{pmatrix} u \\ d \end{pmatrix} \rightarrow U_R \begin{pmatrix} u \\ d \end{pmatrix}_R. \quad (1-32)$$

We can separate the U(1) and the SU(2) parts of these transformations: the symmetry of the classical Lagrangian is $U_V(1) \times U_A(1) \times SU_V(2) \times SU_A(2)$. The vector part of this symmetry is a manifest symmetry of strong interactions, but the $U_A(1)$ is broken by instanton contributions and the $SU_A(2)$ symmetry is spontaneously broken. These two statements imply that the flavor singlet axial current

$$j^{\mu 5} = (\bar{u} \bar{d})\gamma^\mu\gamma^5 \begin{pmatrix} u \\ d \end{pmatrix} \quad (1-33)$$

is anomalous (Adler-Bell-Jackiw anomaly) and that the quark condensate acquires a non-zero expectation value in the QCD vacuum. Due to Goldstone's theorem we expect three massless particles associated with the spontaneously broken symmetries, which are in nature realized as the pion triplet.

Of course, if the quarks are not exactly massless, the isotriplet axial currents are no longer exactly conserved. Then, the quark mass terms give the pions masses of the form (see e.g. Ref. [15])

$$m_\pi^2 = \frac{M^2}{f_\pi}(m_u + m_d), \quad (1-34)$$

with a mass parameter M and f_π the *pion decay constant* with dimension of a mass. The value of M has been estimated to be of the order of 400 MeV, the value of f_π is 93 MeV [23]. To give the pions the observed 140 MeV [23] one therefore needs $(m_u + m_d) \sim 10$ MeV, which is indeed a small, but important perturbation to the massless case.

Chiral symmetry plays an important rôle in continuum QCD and it is therefore desirable to preserve chiral symmetry also in lattice QCD at finite values of the lattice spacing a . But, we have seen already in the discussion around Eq. (1-26) that for the standard Wilson operator one has to live either with the doubling problem or with explicitly broken chiral symmetry.

In fact, it was proven by Nielsen and Ninomiya [24, 25, 26, 27] that for a massless lattice Dirac operator D it is not possible to achieve the following properties at the same time:

1. $D(x)$ is an essentially local operator (bounded by $e^{-\gamma|x|}$).
2. The Fourier transform of $D(x)$ fulfills $\tilde{D}(p) = i\gamma_\mu p_\mu + \mathcal{O}(ap^2)$ for $p \ll \pi/a$.
3. $\tilde{D}(p)$ is invertible for $p \neq 0$. (no massless doublers).
4. The lattice action is invariant under continuous chiral transformations, i.e. $\gamma_5 D + D\gamma_5 = 0$.

In order to circumvent this theorem it was proposed by Ginsparg and Wilson [28] to replace the property that D anti-commutes with γ_5 by the so-called Ginsparg-Wilson relation:

$$\gamma_5 D + D\gamma_5 = aD\gamma_5 R D, \quad (1-35)$$

where R is a local matrix that may have a nontrivial γ -matrix dependence but must have a chirally non-invariant piece.

The Ginsparg-Wilson relation implies a continuous symmetry of the fermionic lattice action, as was shown by Lüscher [29]. The infinitesimal variation corresponding to this symmetry reads

$$\delta\psi = \gamma_5 \left(1 - \frac{1}{2}aD\right)\psi, \quad \delta\bar{\psi} = \bar{\psi} \left(1 - \frac{1}{2}aD\right)\gamma_5, \quad (1-36)$$

which is a flavor singlet chiral transformation on the lattice. The flavor non-singlet transformations can be defined similarly by including a group generator in Eq. (1-36). An operator D fulfilling the Ginsparg-Wilson relation (1-35) is called a *Ginsparg-Wilson operator*.

Since the flavor singlet chiral symmetry in QCD is anomalous in the presence of gauge fields, it is interesting to see whether this anomaly is correctly reproduced

when the lattice Dirac operator fulfills the relation (1-35). In fact it was shown in Ref. [29] that the Ward identities associated with the global flavor singlet chiral transformations on the lattice have the correct anomaly. Moreover, the discussion in Ref. [29] reveals that flavor non-singlet chiral rotations are exact symmetries of the lattice theory. In addition the Atiyah-Singer index theorem [30] holds at finite values of the lattice spacing [31] with the important consequence that the difference of the number of positive chirality zero modes and negative chirality zero modes of a Ginsparg-Wilson Dirac operator in an external color gauge field is equal to the topological charge.

Furthermore, like in continuum QCD, the lattice chiral symmetry prohibits mixing between operators with different chirality. The latter proves to be rather useful in the course of renormalization, in particular in the calculation of matrix elements of the weak interactions Hamiltonian needed for instance for the extraction of B_K .

One particular solution of the Ginsparg-Wilson relation was found by Neuberger [32, 33] and is given by the so called *overlap operator*. For $R_{xy} = \delta_{xy}/\rho$ the massless operator reads

$$D_{\text{ov}}^{(0)} = \frac{\rho}{a} [1 - A(A^\dagger A)^{-1/2}], \quad A = \rho - aD_W, \quad (1-37)$$

where D_W is the Wilson-Dirac operator (1-26) and $0 < \rho < 2$ is a real parameter. A bare mass m_{ov} can be added in the following way

$$D_{\text{ov}} = \left(1 - m_{\text{ov}} \frac{a}{2\rho}\right) D_{\text{ov}}^{(0)} + m_{\text{ov}}, \quad (1-38)$$

where the somewhat un-usual form avoids $\mathcal{O}(a)$ lattice artifacts [34]. The overlap operator was shown to be manifestly gauge covariant and it has no doublers. Moreover it was shown to converge to the expected classical continuum expression (up to a finite normalization constant) and the requirement of locality is fulfilled with a certain choice of the parameter ρ [35]. Note that an equivalent formulation is provided by the *domain wall* approach [36, 37] (for a review see [38]).

However, from a feasibility point of view the overlap operator is compared to the Wilson-Dirac operator computer time demanding. The reason for this is the fact that for each application of D_{ov} an evaluation of $1/\sqrt{A^\dagger A}$ is needed. This is usually done with a polynomial in the operator $A^\dagger A$ and needs therefore a certain number of applications of the latter. This number depends naturally on the parameters under investigation and might very well be of $\mathcal{O}(100)$.

It is also part of this work to compare the computational costs for two particular lattice Dirac operators – one of those being the overlap operator.

1.2.3 Symanzik effective action

We have explained in the last paragraphs that the Wilson formulation of lattice QCD breaks chiral symmetry and shows discretization errors of $\mathcal{O}(a)$, which is not a problem of principle if one is able to simulate at small enough values of the lattice spacing a . Unfortunately, nowadays lattice QCD simulations are still restricted to rather large values of a and it is therefore worthwhile to improve the scaling with the lattice spacing towards the continuum limit.

We will present here two ways to achieve on-shell $\mathcal{O}(a)$ improvement, both of them are based on the effective action introduced by Symanzik [12, 13, 14].

Close to the continuum the lattice theory as has been defined in the previous paragraphs can be described by a local effective action

$$S_{\text{eff}} = S_0 + aS_1 + a^2S_2 + \dots \quad (1-39)$$

While the leading term, S_0 , is just the continuum action[¶], the correction terms S_k are to be interpreted as operator insertions in the continuum theory and are given by

$$S_k = \int d^4x \mathcal{L}_k(x). \quad (1-40)$$

The Lagrangians \mathcal{L}_k have mass dimension $4 + k$ and they are linear combinations of local composite fields. The list of possible fields is constraint by gauge and flavor symmetry and the exact discrete symmetries of the lattice action, including space-time lattice symmetries and charge conjugation. Moreover partial integration can be used to further reduce the number of possible terms.

Clearly the action is not the only origin for cut-off effects, also the local composite fields, from which observables of interest are built, will be a source of those. Consider a local gauge invariant composite field $\phi(x)$ on the lattice, constructed out of gluon and fermion fields. For simplicity we assume that it does not mix with other operators under renormalization. The effective field

$$\phi_{\text{eff}}(x) = \phi_0(x) + a\phi_1(x) + a^2\phi_2(x) + \dots \quad (1-41)$$

represents in the effective theory the renormalized lattice field $Z_\phi\phi(x)$ with an appropriately chosen renormalization factor Z_ϕ . The fields ϕ_k that appear in the representation (1-41) are linear combinations of local fields with appropriate dimension and symmetry properties. To leading order in the lattice spacing a connected lattice n-point function with all points x_1, \dots, x_n kept at non zero distance from

[¶]The continuum theory also has to be regularized to make the expressions meaningful. One could think e.g. of a regularization with a much smaller lattice spacing $\tilde{a} \ll a$.

each other is given by

$$\begin{aligned}
 G_n(x_1, \dots, x_n) &= \langle \phi_0(x_1) \dots \phi_0(x_n) \rangle^{\text{cont}} \\
 &\quad - a \int d^4y \langle \phi_0(x_1) \dots \phi_0(x_n) \mathcal{L}_1(y) \rangle^{\text{cont}} \\
 &\quad + a \sum_{k=1}^n \langle \phi_0(x_1) \dots \phi_1(x_k) \dots \phi_0(x_n) \rangle^{\text{cont}} + \mathcal{O}(a^2).
 \end{aligned} \tag{1-42}$$

All the expectation values on the right hand side of Eq. (1-42) have to be evaluated in the continuum theory with action S_0 , which we indicate by the superscript cont . In the second term, which is the contribution of the $\mathcal{O}(a)$ term in the effective action, the integral over y in general diverges at the points $y = x_k$. However, each such contact term can be described as the effect of a local field placed at x_k . This local field must have the global symmetry quantum numbers of $\phi_0(x_k) \mathcal{L}(y)$ and therefore the possible fields are known from ϕ_{eff} . Hence, the contact terms lead just to a redefinition of the field $\phi_1(x)$ and the renormalization factor.

In the following we will assume that the lattice action is the Wilson action. For this case in Ref. [39] the $\mathcal{O}(a)$ effective Lagrangian $\mathcal{L}_1(x)$ was derived to be a linear combination of the following fields

$$\begin{aligned}
 O_1 &= \bar{\psi} \sigma_{\mu\nu} F_{\mu\nu} \psi, \\
 O_2 &= \bar{\psi} D_\mu D_\mu \psi + \bar{\psi} \overleftarrow{D}_\mu \overleftarrow{D}_\mu \psi, \\
 O_3 &= m \text{Tr} \{ F_{\mu\nu} F_{\mu\nu} \} \\
 O_4 &= m \{ \bar{\psi} \gamma_\mu D_\mu \psi - \bar{\psi} \overleftarrow{D}_\mu \gamma_\mu \psi \} \\
 O_5 &= m^2 \bar{\psi} \psi,
 \end{aligned} \tag{1-43}$$

where m is the quark mass and $F_{\mu\nu}$ the field-strength tensor and D_μ is here the gauge covariant partial derivative, see Ref. [39]. We do not consider the fields O_2 and O_4 , because they can be eliminated by relations derived from the classical field equations. Note that in order to apply the field equations to simplify the effective Lagrangian one has to carefully treat some contact terms. However, these contact terms can be shown to amount only for a redefinition of the fields [39], see above.

Thus, aiming to improve the lattice action by adding a suitable counter-term of $\mathcal{O}(a)$ to the Wilson action, one has to add a counter-term of the form

$$a^5 \sum_x \left[c_1 \hat{O}_1(x) + c_3 \hat{O}_3(x) + c_5 \hat{O}_5(x) \right], \tag{1-44}$$

where \hat{O}_k is some lattice representation of the field O_k . Using the leftover ambiguity of $\mathcal{O}(a^2)$ we may represent the fields $\text{Tr} F_{\mu\nu} F_{\mu\nu}$ and $\bar{\psi} \psi$ by the Wilson plaquette field and the local scalar density, respectively. Since these two already appear in

the Wilson action the counter-terms proportional to \hat{O}_3 and \hat{O}_5 amount to a renormalization of the bare coupling and the bare mass. Let us remark here that the latter prescription to absorb the counter-terms of O_3 and O_5 leads to the additional complication of a mass dependent renormalization scheme, which is usually not favorable. One can cope with this complication [39], but we will not go into these details since we are not going to use this improvement programme further.

Therefore, for the on-shell $\mathcal{O}(a)$ improved action one needs a counter-term

$$\delta S[U, \bar{\psi}, \psi] = a^5 \sum_x c_{\text{sw}} \bar{\psi}(x) \frac{i}{4} \sigma_{\mu\nu} \hat{F}_{\mu\nu}(x) \psi(x), \quad (1-45)$$

where $\hat{F}_{\mu\nu}$ is a lattice representation of the field strength tensor $F_{\mu\nu}$ and c_{sw} is a tunable coefficient^{||}. The coefficient c_{sw} needs to be determined in a suitable way in order to obtain an order a improved theory on the lattice.

Depending on the possible fields ϕ_1 it might then also be necessary to determine further improvement coefficients for the field under consideration. One example is the isovector axial current

$$A_\mu^\alpha(x) = \bar{\psi}(x) \gamma_\mu \gamma_5 \frac{1}{2} \tau^\alpha \psi(x),$$

that needs one improvement coefficient c_A .

Finally, we want to mention that, since O_1 transforms like a mass term under chiral rotations [41] and therefore explicitly breaks chiral symmetry, it cannot appear if the operator obeys the Ginsparg-Wilson relation: the reason for this is that in the chiral limit there is exact chiral symmetry on the lattice and thus all correction terms in the action not proportional to some powers of the mass must be absent in the effective action [34]. Moreover, with a properly introduced mass term as in Eq. (1-38) terms proportional to am are forbidden due to the combined symmetry of $[m \rightarrow -m]$ and

$$\psi \rightarrow \gamma_5(1 - aD)\psi, \quad \bar{\psi} \rightarrow -\bar{\psi}\gamma_5.$$

Thus a lattice regularization with exact chiral symmetry shows no $\mathcal{O}(a)$ lattice artifacts. This, of course, provides an additional advantage of chirally symmetric lattice formulations compared to $\mathcal{O}(a)$ improvement by means of the Symanzik improvement programme: in the latter a (potentially large) number of improvement coefficients has to be tuned, while there is still no control about higher order lattice artifacts.

1.2.4 Twisted mass regularization

Having introduced in the last subsection the concept of Symanzik's effective action and one way to obtain on-shell $\mathcal{O}(a)$ improvement – that amounts in determining

^{||}The improved action was first obtained by Sheikholeslami and Wohlert [40].

improvement coefficients and adding of counter-terms to the lattice action – we will in this subsection follow a different approach, recently realized in Ref. [42], where only one parameter needs to be tuned in order to obtain $\mathcal{O}(a)$ improved results.

To this end we introduce the so called *twisted mass* regularization of lattice QCD (tmQCD) [43, 44]. The twisted mass fermion lattice action for $N_f = 2$ flavors of mass degenerate quarks reads

$$S_{\text{tm}} = a^4 \sum_x \bar{\chi}(x) [D_{\text{W}} + m_0 + i\mu\gamma_5\tau_3] \chi(x) = a^4 \sum_x \bar{\chi}(x) D_{\text{tm}} \chi(x), \quad (1-46)$$

where μ is referred to as the *twisted mass parameter* and τ_3 is the third Pauli matrix acting in flavor space. The twisted mass Dirac operator D_{tm} is given by

$$D_{\text{tm}} = D_{\text{W}} + m_0 + i\mu\gamma_5\tau_3. \quad (1-47)$$

We denote the fermion fields now by χ and call them the twisted basis for a reason that will become clear later. In Ref. [43] this formulation was introduced to avoid un-physically small eigenvalues of the lattice Dirac operator, since the twisted mass serves as an infrared cutoff for the eigenvalues of the operator D_{tm} , and it was shown that this particular regularization is equivalent to the standard Wilson regularization up to cut-off effects [43].

In fact the continuum limit of the action (1-46) reads

$$S_{\text{tm}}^{\text{cont}} = \int d^4x \bar{\chi}(x) [\gamma_\mu D_\mu + m_q + i\mu\gamma_5\tau_3] \chi(x), \quad (1-48)$$

which is form-invariant under axial transformations

$$\chi \rightarrow e^{i\omega\gamma_5\tau_3/2} \chi, \quad \bar{\chi} \rightarrow \bar{\chi} e^{i\omega\gamma_5\tau_3/2}. \quad (1-49)$$

The form (1-48) can be obtained from the standard continuum fermion action by rotating the fermion fields ψ with *twist angle* ω according to

$$\psi \rightarrow \chi = e^{i\omega\gamma_5\tau_3/2} \psi, \quad \bar{\psi} \rightarrow \bar{\chi} = \bar{\psi} e^{i\omega\gamma_5\tau_3/2}. \quad (1-50)$$

Note that the axial transformations (1-50) are non-anomalous: they leave the fermion measure invariant since $\text{Tr } \tau_3 = 0$. The latter rotations just transform the mass parameters according to

$$\begin{aligned} m_q &\rightarrow m_q \cos \omega + \mu \sin \omega, \\ \mu &\rightarrow -m_q \sin \omega + \mu \cos \omega, \end{aligned} \quad (1-51)$$

and the mass term in Eq. (1-48) can also be written as $m \exp(-i\omega\gamma_5\tau_3)$ with $m^2 = m_q^2 + \mu^2$. One form of particular interest – which will become clear later – is the special case with $m_q = 0$

$$S_{\text{mtm}}^{\text{cont}} = \int d^4x \bar{\chi}(x) [\gamma_\mu D_\mu + i\mu\gamma_5\tau_3] \chi(x), \quad (1-52)$$

which is related to the standard action by an axial transformation with angle $\omega = \pi/2$. It is referred to as the action with maximal twist. The corresponding lattice action has the form

$$S_{\text{mtm}} = a^4 \sum_x \bar{\chi}(x) [D_W + m_{\text{crit}} + i\mu\gamma_5\tau_3] \chi(x), \quad (1-53)$$

since the Wilson term $-ar\nabla_\mu^* \nabla_\mu$ still breaks chiral symmetry explicitly and therefore additive renormalization of m_0 is needed. We refer to this particular lattice regularization as *maximally twisted mass lattice QCD* (mtmQCD).

Since the transition from standard to twisted form of the action corresponds to a change of fermionic variables $\psi \rightarrow \chi$, the χ basis is called *twisted basis* and the ψ basis is called *physical basis*. The lattice Wilson twisted mass action (1-46) in the twisted basis can be translated to the physical basis again by an axial transformation of the form of Eq. (1-49) and it reads

$$S_{\text{tm}}^{\text{ph}} = a^4 \sum_x \bar{\psi}(x) \left[\frac{1}{2} \sum_\mu \gamma_\mu (\nabla_\mu^* + \nabla_\mu) + \left(-r \frac{a}{2} \sum_\mu \nabla_\mu^* \nabla_\mu + m_{\text{crit}} \right) e^{-i\omega\gamma_5\tau_3} + m_q \right] \psi(x). \quad (1-54)$$

In this form the fermion mass term is real and, since the Wilson term is not left invariant under the axial rotation, now the Wilson term is rotated. We remark that the tmQCD lattice action can easily be translated to the hopping parameter representation by re-scaling according to Eq. (1-29), as it was done for the Wilson lattice action with Eq. (1-31).

Finally, as in the continuum the tmQCD and the standard QCD are exactly related by the transformation (1-50), they share all the symmetries. But in the twisted basis the symmetry transformations can have a different form than usual. For instance the usual parity operation

$$\mathcal{P} : \begin{cases} U_0(\vec{x}, t) & \rightarrow U_0(-\vec{x}, t) \\ U_k(\vec{x}, t) & \rightarrow U_k^\dagger(\vec{x} - a\hat{k}, t) \\ \chi(\vec{x}, t) & \rightarrow \gamma_0 \chi(-\vec{x}, t) \\ \bar{\chi}(\vec{x}, t) & \rightarrow \bar{\chi}(-\vec{x}, t) \gamma_0 \end{cases} \quad (1-55)$$

is no longer a symmetry of the continuum action in the twisted basis. But if \mathcal{P} is replaced by the following modified parity operation $\tilde{\mathcal{P}}$ [44], the symmetry is recovered:

$$\tilde{\mathcal{P}} : \begin{cases} \chi(\vec{x}, t) & \rightarrow \gamma_0 \exp(i\omega\gamma_5\tau_3) \chi(-\vec{x}, t) \\ \bar{\chi}(\vec{x}, t) & \rightarrow \bar{\chi}(-\vec{x}, t) \exp(i\omega\gamma_5\tau_3) \gamma_0, \end{cases} \quad (1-56)$$

where ω is defined by Eq. (1-51) and the gauge fields transform under $\tilde{\mathcal{P}}$ as for \mathcal{P} . However, for the lattice action in the twisted basis \mathcal{P} is only a symmetry if it is combined with $[\mu \rightarrow -\mu]$ or a flavor exchange [44]. The former combination we will denote with \mathcal{P}_μ and the latter with \mathcal{P}_F . Moreover, on the lattice the isospin symmetry is explicitly broken in the twisted mass formulation even in the mass degenerate case.

$\mathcal{O}(a)$ improvement at maximal twist

The mtmQCD formulation is of particular interest, because one can show that terms proportional to $\mathcal{O}(a)$ are absent in the Symanzik expansion of certain physical observables, as was proven in Ref. [42]. In order to see this we will follow in this work a different approach than originally used in Ref. [42]. It is similar to a proof in the physical basis presented in Ref. [45] and based on the following observations:

First, it is important to notice that all the expectation values in the Symanzik expansion of an operator O on the right hand side of Eq. (1-42) are to be taken in the continuum theory with action S_0 . This implies that any operator not obeying the symmetries of S_0 must have zero expectation value and therefore be absent in the expansion. The symmetry we will use is the just introduced modified parity operation $\tilde{\mathcal{P}}$.

Second, all the operators in the expansion of the operator O must share all its lattice symmetries, otherwise they must be absent. The particular symmetry of the mtmQCD lattice action is $\tilde{\mathcal{P}} \times \mathcal{D}_m \times \mathcal{D}_d$. The transformation according to $\mathcal{D}_m \times \mathcal{D}_d$ is multiplying each term with $(-1)^{d_m+d_d}$. d_m represents the mass dimension of the term and d_d its normal dimension.

And last, the fields contributing to \mathcal{L}_1 are restricted to those that obey the symmetries of the lattice action, as for instance \mathcal{P}_μ and \mathcal{P}_F .

By using these arguments we will now show that in the expansion of an operator O even under $\tilde{\mathcal{P}}$, which means O goes exactly into itself under a $\tilde{\mathcal{P}}$ transformation, terms linear in a are absent. To this end we first have to accumulate the action counter-terms needed for \mathcal{L}_1 : in addition to the fields listed in Eq. (1-43) one finds among others the following:

$$\begin{aligned}
 O_6 &= \mu^2 \bar{\chi} \chi, \\
 O_7 &= \mu \text{Tr}\{F_{\mu\nu} F_{\mu\nu}\}, \\
 O_8 &= im\mu \bar{\chi} \gamma_5 \tau_3 \chi, \\
 O_9 &= m\mu \bar{\chi} \chi, \\
 O_{10} &= \mu \text{Tr}\{F_{\mu\nu} \tilde{F}_{\mu\nu}\}.
 \end{aligned}
 \tag{1-57}$$

We listed some more fields than are actually important in order to show how they can be removed by the help of the symmetries \mathcal{P}_μ and \mathcal{P}_F . For instance O_{10} is parity

odd, but not affected by a flavor exchange and is therefore absent in $\mathcal{L}_1(x)$ due to \mathcal{P}_F . O_9 and O_7 are also absent, because they are odd under \mathcal{P}_μ . At maximal twist m is identical zero and hence we are left with O_6 in addition to O_1 (with ψ replaced by χ), where we again used the classical equations of motion to reduce the number of possible fields.

Consider now an operator O even under $\tilde{\mathcal{P}}$. In the second term of its Symanzik expansion (1-42), containing an insertion of \mathcal{L}_1 , we have to evaluate the expectation value of the $\tilde{\mathcal{P}}$ even operator O and the fields O_i listed above: $\langle OO_i \rangle^{\text{cont}}$. O_1 and O_6 are odd under $\tilde{\mathcal{P}}$, O is even and thus these terms are absent in the expansion.

Since this means that lattice artifacts originating from the lattice action are absent we can conclude already that quantities extracted from the transfer matrix, like for instance hadron masses, are not affected by $\mathcal{O}(a)$ cut-off effects at maximal twist.

In order to show that also the third term on the right hand side of Eq. (1-42) is absent for $\tilde{\mathcal{P}}$ even operators we use the aforementioned symmetry $\tilde{\mathcal{P}} \times \mathcal{D}_m \times \mathcal{D}_d$. The operator O is even under $\tilde{\mathcal{P}}$ and must therefore be even under $\mathcal{D}_m \times \mathcal{D}_d$. As the third term in Eq. (1-42) is multiplied with one power of a it must be odd under $\mathcal{D}_m \times \mathcal{D}_d$ and therefore odd under $\tilde{\mathcal{P}}$. Since O is even under $\tilde{\mathcal{P}}$ also this term must be absent in the expansion.

This is a remarkable result, because with tmQCD at maximal twist it is possible to avoid un-physically small eigenvalues and to obtain $\mathcal{O}(a)$ improved expectation values needing no improvement coefficients**. Quantities even in $\tilde{\mathcal{P}}$ are for instance hadron masses and on-shell matrix elements at zero three-momentum. More examples can be found in Ref. [42].

Overlap versus mtmQCD

We will close this section with a short comparison of the overlap lattice approach to the twisted mass lattice formulation at maximal twist. Both formulations allow to extract $\mathcal{O}(a)$ improved results and the two corresponding lattice Dirac operators are protected against un-physically small eigenvalues [43]. Of course, one has to keep in mind that for the overlap formulation $\mathcal{O}(a)$ improvement comes automatic while for mtmQCD the tuning of the twist angle is required. The first difference of principle is the fact that the overlap operator exhibits exact chiral symmetry on the lattice, while for mtmQCD chiral properties are only improved, since the symmetry is still explicitly broken, even if the effects become visible only at $\mathcal{O}(a^2)$. But the price for exact chiral symmetry on the lattice is that the cost for one application of

**One can show [42] that $\mathcal{O}(a)$ improvement for parity even operators can be achieved also for the Wilson formulation, when one averages over independent simulations with positive and negative value of the Wilson parameter r or with positive and negative value of m_q .

the overlap Dirac operator is significantly larger than the cost for the application of the twisted mass operator.

We note in passing that also with mtmQCD it is possible to avoid unwanted mixing of operators with opposite chirality, although chiral symmetry is broken. For details we refer to Ref. [46].

1.3 Observables

In the preceding sections we have discussed several formulations of lattice QCD and in particular, how one can obtain $\mathcal{O}(a)$ improvement. So far we did not yet explain how physical quantities like masses and decay constants can be extracted in lattice simulations, which is the content of the following section. Since we will work with three different formulations, the pure Wilson, the tmQCD and the overlap formulation, we have to consider in this section all of the three formulations. The composite fields and correlation functions for the Wilson formulation can equivalently be used in the tmQCD formulation, they only differ by the fact that for tmQCD we work in the twisted basis, where all the fermion fields are rotated. There are, however, some specialties in the case of maximal twist. We will discuss the three formulations in separate subsections.

Before coming to the two point functions needed in this work it is useful to discuss the relation between Euclidean and Minkowski quantities. Consider the following Euclidean time ordered two point function

$$\langle 0|T[O_1(x)O_2(0)]|0\rangle, \quad (1-58)$$

with operators O_1 and O_2 , representing the amplitude for the creation of a state with quantum numbers of operator O_2 at space-time point 0, its propagation to space-time point $x = (\vec{x}, t)$ and finally its annihilation by O_1 . By integrating over space-like coordinates we project to zero momentum states. If we then insert a complete set of energy eigenstates we obtain

$$\langle 0|\int d^3x O_1(x)O_2(0)|0\rangle = \sum_n \frac{\langle 0|O_1|n\rangle\langle n|O_2|0\rangle}{2E_n} e^{-E_n t}. \quad (1-59)$$

If there is a stable single-particle state $|n\rangle$ with the corresponding quantum numbers, then its Energy E_n is equal to the mass M_n of the particle, since we have projected to zero momentum. In general there might be several such states, but asymptotically for large enough values of t the correlation function will be dominated by the state with the lowest mass, e.g. M_1 . Thus one gets

$$\langle 0|\int d^3x O_1(x)O_2(0)|0\rangle \xrightarrow{t \rightarrow \infty} \frac{\langle 0|O_1|1\rangle\langle 1|O_2|0\rangle}{2M_1} e^{-M_1 t}, \quad (1-60)$$

and the mass of the lightest particle can be extracted from the exponential fall-off of the correlation function at large Euclidean times without analytical continuation to Minkowski space. This simple connection also holds for matrix elements. These arguments, however, break down as soon as the sum over states is not just over stable single-particle states. For instance the ρ meson mass can no longer be extracted by simply measuring the ground state energy (cf. [47, 48]).

1.3.1 Wilson formulation

Hadron and quark masses are – in the simplest cases – extracted from two point correlation functions of suitable composite fields. The most common bilinears are the scalar and the pseudo scalar densities, which read in the physical basis ψ :

$$S^0(x) = \bar{\psi}(x)\psi(x), \quad P^\alpha(x) = \bar{\psi}(x)\gamma_5\frac{\tau_\alpha}{2}\psi(x), \quad (1-61)$$

and the vector and the axial currents

$$\begin{aligned} A_\mu^\alpha(x) &= \bar{\psi}(x)\gamma_\mu\gamma_5\frac{\tau_\alpha}{2}\psi(x), \\ V_\mu^\alpha(x) &= \bar{\psi}(x)\gamma_\mu\frac{\tau_\alpha}{2}\psi(x). \end{aligned} \quad (1-62)$$

Here τ_α , $\alpha = 1, 2, 3$ are the usual Pauli matrices acting in isospin space. Moreover the local bilinear with tensor structure is of interest

$$T_k^\alpha(x) = \bar{\psi}(x)\gamma_0\gamma_k\frac{\tau_\alpha}{2}\psi(x). \quad (1-63)$$

All local interpolating field operators for mesons in a Wilson like theory can be found in table 1.1. In practice a general two point correlator with three-momentum \mathbf{p} can be re-written in terms of the quark propagators as follows

$$\begin{aligned} C_{AB}(\mathbf{p}, t) &= \sum_{\mathbf{x}} e^{-i\mathbf{p}\mathbf{x}} \langle \bar{\psi}_2(x)\Gamma_B\psi_1(x)\bar{\psi}_1(0)\Gamma_A\psi_2(0) \rangle \\ &= - \sum_{\mathbf{x}} e^{-i\mathbf{p}\mathbf{x}} \langle \text{Tr}(S_2(0, x)\Gamma_B S_1(x, 0)\Gamma_A) \rangle. \end{aligned} \quad (1-64)$$

Here $\Gamma_{A,B}$ represent the Dirac structure, the quark propagators are denoted with $S_{1,2}$ and the trace is taken over spin and color indices. The sum over the space-like points \mathbf{x} on a given time slice projects to zero momentum states. In case the quark propagators fulfill the hermiticity property $S(0, x) = \gamma_5 S(x, 0)^\dagger \gamma_5$ the computation of both $S(0, x)$ and $S(x, 0)$ can be avoided in favor of computing only one of the two.

As explained at the beginning of this section the meson masses can then be extracted from the exponential fall-off of suitable correlation functions at large Euclidean times. In terms of the local interpolating fields introduced above for instance

State	$I^G(J^{PC})$	Operator
Scalar	$1^-(0^{++})$	$\bar{u}(x)d(x)$
	$1^-(0^{++})$	$\bar{u}(x)\gamma_0d(x)$
Pseudo-scalar	$1^-(0^{-+})$	$\bar{u}(x)\gamma_5d(x)$
	$1^-(0^{-+})$	$\bar{u}(x)\gamma_5\gamma_0d(x)$
Vector	$1^+(1^{--})$	$\bar{u}(x)\gamma_id(x)$
	$1^+(1^{--})$	$\bar{u}(x)\gamma_0\gamma_id(x)$
Axial	$1^+(1^{++})$	$\bar{u}(x)\gamma_5\gamma_id(x)$
Tensor	$1^+(1^{+-})$	$\bar{u}(x)\gamma_i\gamma_jd(x)$

Table 1.1: The local interpolating field operators for light mesons in a Wilson like theory. The quark fields corresponding to the two quark flavors *up* and *down* are labeled with *u* and *d*.

the pseudo scalar mass m_{PS} and the vector meson mass m_{V} can be extracted from the following two point correlation functions:

$$C_{PP}^\alpha(t) = a^3 \sum_{\mathbf{x}} \langle P^\alpha(x) P^\alpha(0) \rangle \quad \alpha = 1, 2, \quad (1-65)$$

$$C_{VV}^\alpha(t) = \frac{a^3}{3} \sum_{k=1}^3 \sum_{\mathbf{x}} \langle V_k^\alpha(x) V_k^\alpha(0) \rangle \quad \alpha = 1, 2, \quad (1-66)$$

$$C_{TT}^\alpha(t) = \frac{a^3}{3} \sum_{k=1}^3 \sum_{\mathbf{x}} \langle T_k^\alpha(x) T_k^\alpha(0) \rangle \quad \alpha = 1, 2. \quad (1-67)$$

In a similar way the correlation functions $C_{AA}^\alpha(t)$, $C_{AP}^\alpha(t)$ and others can be defined. Then, due to periodic boundary conditions, for example the correlation function C_{PP}^α is expected to behave for large Euclidean times t like $f(t) = A \cosh(m_{\text{PS}}(t - T/2))$, where T is the lattice time extent, m_{PS} is the corresponding mass and A the amplitude. By fitting the functional form $f(t)$ to the data in a certain window $t_{\min} \leq t \leq t_{\max}$, where all the excited states have vanished, the values for the masses are actually extracted. Moreover, the amplitude A gives an estimate for the matrix element $\langle 0 | P^\alpha(0) | \pi \rangle$, where we denote with $|\pi\rangle$ the pseudo scalar state with the lightest mass.

Besides the meson masses, the quark mass and the pseudo scalar decay constant

$$f_{\text{PS}} \equiv m_{\text{PS}}^{-1} \langle 0 | A_0^\alpha(0) | \pi \rangle \quad (1-68)$$

are quantities of interest. For calculating f_{PS} the following two methods are possible. The first one is to obtain the amplitude $\langle 0 | A_0^\alpha(0) | \pi \rangle$ from the asymptotic behavior of the correlation function $C_{AA}^\alpha(t)$ while the pseudo scalar mass is extracted from $C_{PP}^\alpha(t)$. The second method [49] is to obtain the ratio (we skip the flavor index α

in the following)

$$r_{AP} = \frac{\langle 0|A_0(0)|\pi\rangle}{\langle 0|P(0)|\pi\rangle} \quad (1-69)$$

from the asymptotic behavior of

$$\frac{C_{AP}(t)}{C_{PP}(t)} = r_{AP} \tanh[m_{\text{PS}}(T/2 - t)], \quad (1-70)$$

where again m_{PS} is extracted from $C_{PP}(t)$. The value for f_{PS} can then be obtained from

$$f_{\text{PS}} = m_{\text{PS}}^{-1} r_{AP} \langle 0|P(0)|\pi\rangle. \quad (1-71)$$

Clearly, both methods agree asymptotically for large Euclidean times. With one of these definitions for f_{PS} also the bare current quark mass can be defined by the PCAC relation

$$m^{\text{PCAC}} = \frac{f_{\text{PS}}}{2\langle 0|P(0)|\pi\rangle} m_{\text{PS}}^2. \quad (1-72)$$

Note that this quantity – in contrast to the masses extracted from the exponential fall-off – requires multiplicative renormalization, which is also the case for f_{PS} determined in the way explained above. In the pure Wilson case m^{PCAC} serves as an estimate for the bare quark mass.

1.3.2 tmQCD formulation

Introducing a twisted mass term in the Wilson-Dirac operator is equivalent to transforming the fermion fields $\psi \rightarrow \chi$ according to Eq. (1-50). Therefore also the composite fields have to be transformed. The result for the axial and vector currents is the following:

$$A_\mu'^\alpha \equiv \bar{\chi} \gamma_\mu \gamma_5 \frac{\tau_\alpha}{2} \chi = \begin{cases} A_\mu^\alpha \cos(\omega) + \epsilon^{3\alpha\beta} V_\mu^\beta \sin(\omega) & (\alpha = 1, 2), \\ A_\mu^3 & (\alpha = 3), \end{cases} \quad (1-73)$$

$$V_\mu'^\alpha \equiv \bar{\chi} \gamma_\mu \frac{\tau_\alpha}{2} \chi = \begin{cases} V_\mu^\alpha \cos(\omega) + \epsilon^{3\alpha\beta} A_\mu^\beta \sin(\omega) & (\alpha = 1, 2), \\ V_\mu^3 & (\alpha = 3), \end{cases} \quad (1-74)$$

whereas the rotated scalar and the pseudo scalar densities are given by:

$$P'^\alpha \equiv \bar{\chi} \gamma_5 \frac{\tau_\alpha}{2} \chi = \begin{cases} P^\alpha & (\alpha = 1, 2), \\ P^3 \cos(\omega) + i\frac{1}{2} S^0 \sin(\omega) & (\alpha = 3), \end{cases} \quad (1-75)$$

$$S'^0 \equiv \bar{\chi} \chi = S^0 \cos(\omega) + 2iP^3 \sin(\omega). \quad (1-76)$$

In the special case of $\omega = \pi/2$ the vector and the axial currents with $\alpha = 1, 2$ transform into each other, while the pseudo scalar densities ($\alpha = 1, 2$) do not rotate

with ω . Therefore the charged pseudo scalar (“pion”) mass can be extracted for all values of ω from the pseudo scalar density also in the twisted basis. In the following we will skip the ‘ on the quantities in the twisted basis again, since we will always work in the twisted basis if tmQCD is concerned.

The PCAC and PCVC relations assume in the twisted basis the following form [44]:

$$\begin{aligned}\partial_\mu A_\mu^\alpha &= 2m_0 P^\alpha + i\mu\delta^{3\alpha} S^0, \\ \partial_\mu V_\mu^\alpha &= -2\mu\epsilon^{3\alpha\beta} P^\beta.\end{aligned}\tag{1-77}$$

While in the pure Wilson-Dirac operator there is only one mass term, which is aligned to the Wilson term (proportional to the unit matrix in flavor space), in the twisted mass operator (1-54) there is an additional mass term in the three-direction of flavor space, which is dis-aligned to the Wilson term. Therefore, at general values of the twist angle the quark mass estimate has to contain both of them, μ and m_q

$$m = \sqrt{(Z_{m_q} m_q)^2 + (Z_\mu \mu)^2}.\tag{1-78}$$

Of course, the twisted mass lattice action can be studied in the full parameter space (μ, m_q) , but automatic $\mathcal{O}(a)$ improvement is only realized at full twist corresponding to $m_q = 0$. A sensible definition for an estimate of m_q is given by

$$m_\chi^{\text{PCAC}} = \frac{\langle \nabla_\mu^* \bar{\chi}(x) \gamma_\mu \gamma_5 \frac{\tau^\pm}{2} \chi(x) \bar{\chi}(y) \frac{\tau^\pm}{2} \gamma_5 \chi(y) \rangle}{2 \langle \bar{\chi}(x) \frac{\tau^\pm}{2} \gamma_5 \chi(x) \bar{\chi}(y) \frac{\tau^\pm}{2} \gamma_5 \chi(y) \rangle},\tag{1-79}$$

where we introduced $\tau^\pm = \tau_1 \pm i\tau_2$. Comparing to Eqs. (1-73,1-74,1-75) it is clear that m_χ^{PCAC} is identical to m^{PCAC} for $\omega = 0$. Moreover, at the special value of $\omega = \pi/2$, m_χ^{PCAC} is zero due to the exact vector symmetry in the lattice theory, and the quark mass is purely given by the twisted mass parameter μ , as it should be. Thus, m_χ^{PCAC} is a quantity that can be used to determine the parameters at which $\omega = \pi/2$.

In the twisted basis we can also define a quantity f_χ^{PS} , given by Eq. (1-68), but now in the twisted basis. We denote with the subscript χ that this is a quantity extracted in the twisted basis, and in fact it does not correspond to the physical pseudo scalar decay constant. A further estimate estimate for m_q in the twisted basis is then given by

$$m_\chi^{\text{PCAC}} = \frac{f_\chi^{\text{PS}}}{2 \langle 0 | P(0) | \pi \rangle} m_{\text{PS}}^2,\tag{1-80}$$

representing again the quark mass term aligned to the Wilson term.

Pseudo scalar decay constant at maximal twist

Of course, given the relations of the composite fields (1-73) connecting the twisted basis χ to the basis ψ , the value for f_{PS} can be extracted from f_{χ}^{PS} . But for the special choice $\omega = \pi/2$ this does not work (see above) and there is another possibility to extract a value for f_{PS} even without the need of a renormalization factor (cf. [50, 51, 52]).

Of particular interest in this context is the PCVC relation, which takes in the twisted basis the following form (1-77):

$$\partial_{\mu} \tilde{V}_{\mu}^{\alpha} = -2\mu \epsilon^{3\alpha\beta} P^{\beta}, \quad (1-81)$$

which holds exactly when the point split vector current \tilde{V} as defined in Ref. [44] is used. From this follows that the vector current is protected against renormalization – in analogy to Ginsparg-Wilson fermions, which means the renormalization factor $Z_{\tilde{V}} = 1$. Therefore, Eq. (1-81) implies that $Z_P = Z_{\mu}^{-1}$, where Z_P is the renormalization factor of the pseudo scalar density and Z_{μ} the one for the twisted mass parameter μ .

Fixing the flavor index to $\alpha = 1$ we now again start with the standard definition for the pseudo scalar decay constant in the physical basis Eq. (1-68)

$$\langle 0 | A_0^1(0) | \pi \rangle = f_{\text{PS}} m_{\text{PS}}. \quad (1-82)$$

In the twisted basis at maximal twist the rôle of the axial and vector current is just interchanged, and therefore we can write in the twisted basis

$$\partial_{\mu} \langle 0 | V_{\mu}^2(0) | \pi \rangle = f_{\text{PS}} m_{\text{PS}}^2. \quad (1-83)$$

Using then the vector Ward identity (1-81), we can finally relate the divergence of the vector current to the pseudo scalar density and obtain

$$f_{\text{PS}} m_{\text{PS}}^2 = \partial_{\mu} \langle 0 | V_{\mu}^2(0) | \pi \rangle = 2\mu \langle 0 | P^1(0) | \pi \rangle. \quad (1-84)$$

Thus, by fitting the pseudo scalar correlation function for large time separations, we can obtain m_{PS} and the amplitude $|\langle 0 | P^1(0) | \pi \rangle|^2 / m_{\text{PS}}$ from which we then compute the desired matrix element $|\langle 0 | P^1(0) | \pi \rangle|$. Hence, we have all necessary ingredients to determine f_{PS} from Eq. (1-84), without the need of any renormalization factor.

1.3.3 Overlap formulation

As we have discussed before the overlap formulation obeys exact chiral symmetry on the lattice and the theory is $\mathcal{O}(a)$ improved, as long as cut-off effects originating from the action are considered. We have seen that besides the action also the operators

are possible origins of cut-off effects. For the overlap formulation $\mathcal{O}(a)$ improved bilinears with a Dirac structure Γ can be constructed as follows ($\bar{a} = a/\rho$):

$$O_\Gamma = \bar{\psi}_q \Gamma \left(1 - \frac{\bar{a} D_{\text{ov}}^{(0)}}{2} \right) \psi_{q'} = \frac{1}{1 - \frac{\bar{a} m_{q'}}{2}} \left(\psi_q \Gamma \psi_{q'} \right), \quad (1-85)$$

where ψ_q and $\psi_{q'}$ represent two different quark flavors q and q' with masses m_q and $m_{q'}$, respectively. The pseudo scalar meson mass is then extracted from the exponential fall-off at large Euclidean times of

$$C_{PP}^{\text{ov}}(t) = a^3 \sum_{\mathbf{x}} \langle P^\dagger(x) P(0) \rangle, \quad (1-86)$$

where the pseudo scalar density now is given by

$$P(x) = \bar{\psi}_q(x) \gamma_5 \left[\left(1 - \frac{\bar{a}}{2} D_{\text{ov}}^{(0)} \right) \psi_{q'} \right](x). \quad (1-87)$$

In order to remove contributions of topological zero modes (which are finite volume effects, cf. [53, 54, 55, 56, 57, 58]) the same quantity can be extracted from

$$C_{PP-SS}^{\text{ov}}(t) = a^3 \sum_{\mathbf{x}} \langle P^\dagger(x) P(0) - S^\dagger(x) S(0) \rangle, \quad (1-88)$$

where S can be defined via Eq. (1-85) as follows

$$S(x) = \bar{\psi}_q(x) \left[\left(1 - \frac{\bar{a}}{2} D_{\text{ov}}^{(0)} \right) \psi_{q'} \right](x). \quad (1-89)$$

While C_{PP-SS}^{ov} has the advantage that contributions from topological zero modes are canceled, it has the drawback that the scalar meson appears as an excited state and can affect the extraction of the ground state mass for large quark masses. The vector meson mass m_V is obtained with the overlap operator from

$$C_{VV}^{\text{ov}} = \frac{a^3}{3} \sum_{k=1}^3 \sum_{\mathbf{x}} \langle V_k^\dagger(x) V_k(0) \rangle, \quad (1-90)$$

where the vector current is defined to be

$$V_\mu(x) = \bar{\psi}_q(x) \gamma_\mu \left[\left(1 - \frac{\bar{a}}{2} D_{\text{ov}}^{(0)} \right) \psi_{q'} \right](x). \quad (1-91)$$

For later purposes we define the axial vector current in the overlap formulation

$$A_\mu(x) = \bar{\psi}_q(x) \gamma_\mu \gamma_5 \left[\left(1 - \frac{\bar{a}}{2} D_{\text{ov}}^{(0)} \right) \psi_{q'} \right](x). \quad (1-92)$$

The extraction of the bare quark mass and the pseudo scalar decay constant from the Ward identity in case of overlap fermions is identical to the one described

here for Wilson fermions. Only the bilinears have to be replaced by the suitable counterparts as given in Eqs. (1-87) and (1-92).

We remark here that with overlap fermions the value for f_{PS} can also be obtained in the same way as for twisted mass fermions by using the PCAC relation. In fact f_{PS} can be computed from

$$f_{\text{PS}}^{\text{ov}} = \frac{2m_{\text{ov}}}{m_{\text{PS}}^2} |\langle 0 | P(0) | \pi \rangle|, \quad (1-93)$$

where as in the tmQCD case no renormalization constant is needed. The quark mass m_{ov} is the quark mass parameter in the overlap operator (1-38).

1.3.4 Setting the scale

In lattice calculations, we need to fix one dimension-full quantity in order to set the overall scale and to translate lattice units into physical units. This can be done by using a hadronic scale r_0 , which is introduced by the force $F(r)$ between static quarks at intermediate distance r [59].

The hadronic length scale r_0 defined by the implicit equation

$$r^2 F(r)|_{r=r(c)} = c, \quad r_0 = r(1.65) \quad (1-94)$$

has turned out to be a good choice to set the scale in lattice QCD calculations: it can be computed on the lattice with high precision, both statistically and systematically, and r_0 is known to have a value of about 0.5 fm in QCD.

The force can be computed from the static quark potential that in turn can be determined on the lattice from Wilson loops, which are $r \times t$ loops of gauge links. Wilson loops are defined – similar to the plaquette variable – as the traces of products of parallel transports around a closed loop. This is, as discussed before, a gauge invariant object due to cyclic invariance in the trace. They represent a static quark/anti-quark pair separated with distance r in space and propagating in time the distance t . In order to improve the overlap with the ground state usually several levels of APE smearing [60] are applied to the space like gauge links.

Following the variational approach of Ref. [61] one gets a correlation matrix $W_{ij}(r, t)$ with i, j representing the smearing levels applied at the two space like gauge link products. The correlation matrix can be used to solve the generalized eigenvalue problem

$$W_{ij}(r, t)v_j(r) = \lambda(r, t_0, t)W_{ij}(r, t_0)v_j(r). \quad (1-95)$$

The eigenvector of the largest eigenvalue can be used to project W_{ij} on the ground state and the ground state energy can then be obtained from the exponential fall-off

of the latter at large enough values of t , leading to an estimate of the potential $V(r)$. Equivalently, the generalized eigenvalue can be used as an estimate for the ground state energy.

From the potential $V(r)$ the force can be computed by numerical differentiation, which in general reads

$$F(r') = [V(r') - V(r' - a)] / a, \quad (1-96)$$

at a certain distance r' . We decided to use a distance $r' = r_{\text{I}}$ for which the lattice force does not deviate from the tree-level continuum value [59]

$$F(r_{\text{I}}) = \frac{4}{3} \frac{g^2}{4\pi r_{\text{I}}^2} + \mathcal{O}(g^4). \quad (1-97)$$

The value of r_0/a can then be determined by interpolating the force to the value of r/a where Eq. (1-94) is fulfilled. A second possibility is to fit

$$V(r) = V_0 + \sigma r + \alpha \frac{1}{r}$$

to the data and determine the force from the best fit function. σ is the so called string tension and α parameterizes the Coulomb like part proportional to the gauge coupling. The final results, however, should agree within the errors for both methods.

The time like gauge links represent a static propagator of the static quarks that can be derived from an effective static action. Here the naive choice is the Eichten-Hill action [62], but one can also use static actions that show an improved signal to noise ratio [63, 64, 65, 66]. The improved actions formally correspond to actions with APE or HYP [67] smeared time like gauge links or a differently discretized time derivative. The APE and in particular the HYP action lead to improved statistical and systematical precision and therefore to a better estimate for r_0 in lattice units [63, 68, 69].

Our experience is that using a HYP static action is superior to using a APE static action, which is then superior to the Eichten-Hill static action. This concerns in particular the signal to noise ratio at large distances, where it is only with a improved static action possible to determine a value for the force. This becomes important close to the continuum and if one wants to use the potential to investigate for instance string breaking [69, 70]. We remark that changing the static action is equivalent to a different choice of the operator. Therefore, the different lattice estimates for r_0/a deviate from each-other by lattice artifacts, while the continuum extrapolated value should of course agree.

1.4 Lattice simulations

In this section we will shortly discuss some general principles for constructing algorithms used in lattice simulations.

In order to compute the partition function (1-21) or an expectation value of an operator O in principle one has to solve a high dimensional integral over the gauge fields and the Grassmann valued fermionic fields. However, most of the contributions to the integral have only low weight and therefore stochastic integration with *importance sampling* is an effective method to evaluate such integrals.

Stochastic integration with importance sampling preferentially chooses such configurations that have a strong weight. At the same time it is assured that the sample average estimates the ensemble average. This means that the sample is representative for the ensemble. In particular, such a stochastic integration can be performed by means of Markov chains:

Markov chain

Consider a stochastic process in which a finite set of configurations $U_{\tau_1}, U_{\tau_2}, \dots$ is generated sequentially according to some transition probability $P_{ij} \equiv P(U_i \rightarrow U_j)$. The state of the system at any given simulation time τ_i will be a multi-dimensional random variable, whose distribution depends only on the preceding state, if P_{ij} depends only on the state U_i . A set of configurations generated in this way is called a *Markov chain*.

For an observable O we can define a simulation time average over a given set of configurations $\{U_{\tau_i}\}$ generated in a Markov chain by

$$\langle O \rangle_N = \frac{1}{N} \sum_{i=1}^N O(U_{\tau_i}). \quad (1-98)$$

We want to set up the transition probability in such a way that $\langle O \rangle_N$ is in the limit $N \rightarrow \infty$ equal to the ensemble average corresponding to a given Boltzmann distribution e^{-S} . In order to achieve this it is sufficient that the transition probability fulfills as a sufficient condition the so called *detailed balance* condition

$$e^{-S(U)} P(U \rightarrow U') = e^{-S(U')} P(U' \rightarrow U). \quad (1-99)$$

There are many algorithms known that correctly implement the condition of detailed balance. One of these is the *Metropolis algorithm* which is given by the following two steps:

1. Chose an arbitrary test configuration U' .

2. Accept configuration U' as successor of configuration U with probability

$$P(U \rightarrow U') = \begin{cases} \frac{e^{-S(U')}}{e^{-S(U)}} & \text{if } e^{-S(U')} < e^{-S(U)} \\ 1 & \text{else.} \end{cases} \quad (1-100)$$

This step is called a Metropolis accept/reject step.

The Metropolis algorithm is in principle applicable to a any system and it can be very efficient, if one can efficiently produce test configurations in such a way that the acceptance rate is high. One only has to take into account that the generation of test configurations must be *ergodic*. This means nothing but that the probability to generate configuration U' as a next test configuration must not be zero for any possible configuration.

1.4.1 Quenched approximation

As explained before, the dependence of expectation values on the fermionic fields ψ can be removed by the help of Wick's theorem. An expectation value of an operator O then reads

$$\langle O \rangle = \frac{1}{Z} \int \mathcal{D}U \ O[U] \ e^{-S_G[U] - \log \det M[U]} . \quad (1-101)$$

This means that only the integral over the color gauge configurations has to be performed and we need an algorithm to generate color gauge configurations U with the desired distribution. However, for instance for each accept/reject step the determinant of M needs to be computed, which is an highly non-local object.

Therefore, solely due to limited computer resources, computations in lattice QCD were often performed in a crude approximation: it consists of neglecting the fermion contribution to the path integral, i.e. setting $\det M = \text{constant}$. This approximation corresponds to neglecting vacuum polarization effects of quark loops. As a consequence for instance the string between a quark and an anti-quark does not break at any distance.

Even though one could expect this approximation to be bad, since quenched lattice QCD is confining, asymptotically free and shows spontaneous chiral symmetry breaking, it is reasonable to use it as a model of QCD. As a side effect one can also extract physical results and compare it to experiment. And in fact the quenched approximation seems to work surprisingly well: the deviations from experimental measurements are only of the order of 10%, even though the systematic errors are hard to estimate.

From a practical point of view setting $\det M = \text{constant}$ corresponds to simulating a pure Yang-Mills gauge theory [71] with for instance the Wilson gauge action (1-25) on the lattice. For pure gauge theories there are efficient Metropolis

Monte-Carlo algorithms available, like the *heat-bath* and the *overrelaxation* algorithm. While the latter of the two being not ergodic, their combination gives rise to an efficient algorithm to generate gauge configurations with the correct distribution. Details on this algorithm can be found in text books, see for instance Refs. [18, 19].

Finally we remark that, since r_0 as defined in section 1.3.4 is a quantity depending only on the gauge fields, the scale in quenched simulations does not depend on the fermionic mass, but only on the coupling β in the here discussed quenched approximation. It was computed for the Wilson plaquette gauge action in a range of coupling constants between $\beta = 5.7$ and $\beta = 6.75$ in Ref. [72] (see also [73]). Therefore, in the quenched simulations for this work we did not compute r_0 but rather used the values from Ref. [72].

1.4.2 Dynamical simulations

Of course, the aim of lattice QCD computations is to take the fermion determinant into account^{††}. The reason for full QCD simulations to be much more expensive than the one in the quenched approximation is – as mentioned above – that the computation of the determinant is rather demanding.

Nevertheless, there are two widely used algorithms, which include the determinant in the generation of gauge configurations: on the one hand Multi-Boson like algorithms [74] and on the other hand Hybrid Monte Carlo (HMC) like algorithms [75]. The latter is discussed in detail in chapter 3, while the former is not used for this work. However, a Two-Step Multi-Boson (TSMB) algorithm [76] was used in the collaboration in addition to the HMC for checks and for production runs. Naturally, we cross-checked the results and found full agreement within errors.

Details for Multi-Boson like algorithms can be found in the mentioned references.

1.4.3 Error estimates

In lattice simulations with Monte-Carlo methods there exist several sources of errors. Apart from errors due to discretization and finite volume effects, the most important contribution – and usually also the largest – is the statistical error.

The latter arises from the fact that due to finite computer and human resources it is not possible to average over an infinitely large sample. Nevertheless, it is possible to estimate the error one makes by approximating the infinite large sample with a finite one. It is the content of this subsection to discuss the effects of taking only a finite sub-sample.

^{††}Only because for a long time most of the simulations have been performed in the quenched approximation it became common to call full QCD simulations *dynamical simulations*.

Statistical error

When averaging over a sample of M configurations in order to measure the (primary) observable O , the root-mean-square deviation^{‡‡} can be computed by

$$\sigma_{O,M}^2 = \frac{1}{M-1} \left(\frac{1}{M} \sum_{i=1}^M O_i^2 - \left[\frac{1}{M} \sum_{i=1}^M O_i \right]^2 \right) = \frac{1}{M-1} (\langle O^2 \rangle - \langle O \rangle^2), \quad (1-102)$$

where O_i represents the i -th measurement of the observable O . For this estimate it is necessary that the single measurements are not correlated, which is in general not true. Therefore we call it the *naive* error estimate. In order to account for the correlation one can use the *integrated autocorrelation time* τ_{int} . It can be defined as follows

$$\tau_{\text{int}} = \frac{1}{2} \sum_{\tilde{t}=-\infty}^{\infty} \frac{\Gamma_O(|\tilde{t}|)}{\Gamma_O(0)}, \quad (1-103)$$

where we label the ‘‘Monte Carlo time’’ with \tilde{t} and we introduced the autocorrelation function Γ_O for an operator O

$$\Gamma_O(|i-j|) = \langle (O_i - \langle O \rangle)(O_j - \langle O \rangle) \rangle.$$

$\Gamma_O(|i-j|)$ depends only on the distance $\tilde{t} = |i-j|$ between measurements and decays typically exponentially fast with the so called autocorrelation time τ_c

$$\Gamma_O(\tilde{t}) \propto \exp(-\tilde{t}/\tau_c).$$

Typically τ_c and τ_{int} are found to be of the same order. There are two possibilities to incorporate the integrated autocorrelation time in the estimate of the statistical error:

On the one hand one can leave out during the course of production of configurations τ_{int} many configurations until one is used for the measurements. Then the sample of configurations is uncorrelated and the naive error (1-102) can serve as a good approximation for the real error. This approach has the disadvantage that the value of τ_{int} is in general not known before the measurements were performed and one might perform measurements which are then not used for the final result.

On the other hand one can use the value of τ_{int} for the error computation taking all the measurements into account: it is possible to show that the statistical error of correlated measurements can be computed from the naive error and the integrated autocorrelation time in the following way [77]

$$\sigma_{O,M}^2 = 2\tau_{\text{int}} (\sigma_{O,M}^{\text{naiv}})^2. \quad (1-104)$$

^{‡‡}The denominator arises from the fact that the exact mean value \overline{O} was replaced by its estimate over the sample, because the exact mean value is not known.

This means de facto a reduction of the number of independent measurements to $M/(2\tau_{\text{int}})$.

Of course, when the integrated autocorrelation time shall be used to estimate the error, a reliable estimate of τ_{int} itself and its error is needed. First of all a reliable determination of τ_{int} is only possible, if the sample is large enough. Moreover the estimation of the statistical error on τ_{int} is a delicate procedure. It is discussed in detail in Ref. [78] and we will solely perform our error analysis along the lines of this reference (cf. also [77]).

Another method to estimate the real error is to pre-average the measurements on blocks of the total sample:

$$O_{l,B} = \frac{1}{B} \sum_{i=1+(l-1)B}^{lB} O_i, \quad l = 1, \dots, M_B = \frac{M}{B}. \quad (1-105)$$

If the blocks B become large enough – in units of τ order of τ_{int} – then the block averages are uncorrelated and the error can be estimated by

$$\sigma^2 = \frac{1}{M_B(M_B - 1)} \sum_l \left(O_{l,B} - \frac{1}{M_B} \sum_{l'} O_{l',B} \right)^2. \quad (1-106)$$

The value of the latter will increase with increasing B and will, if the block-averages become uncorrelated, reach a plateau. Of course the plateau will be only reached, if the sample is large enough. This method is called *binning*.

A further method, which is similar to binning, is the so called *Jackknife binning*. Instead of using the blocks itself to pre-average the measurements, the blocks complementary to the binning blocks are used. Therefore the blocks are significantly larger:

$$O_{l,\bar{B}} = \frac{1}{M - B} \left(\sum_{i=1}^{(l-1)B} O_i + \sum_{i=lB+1}^M O_i \right).$$

One can show that the error is now estimated by

$$\sigma^2 = \frac{M_B - 1}{M_B} \sum_l \left(O_{l,\bar{B}} - \frac{1}{M_B} \sum_{l'} O_{l',\bar{B}} \right)^2.$$

Especially for derived quantities the Jackknife binning is a widely used method (cf. [18]).

In this work we solely rely on the error analysis with help of the integrated autocorrelation time. The other methods are only used for checks. The statistical error on the integrated autocorrelation time as well as the error for derived quantities is determined along the lines of [78] (cf. [77]).

Chapter 2

Scaling test in the quenched approximation

In this chapter we will derive a result that was thought to be almost impossible to reach using the standard Wilson formulation of lattice QCD: we will give *continuum* results (in the quenched approximation) for the pseudo scalar mass m_{PS} , the pseudo scalar decay constant f_{PS} and the vector mass m_{V} down to values of the pseudo scalar mass of $m_{\text{PS}} = 270$ MeV, a value that is almost a factor of 2 smaller than what could be reached with Wilson fermions so far. This result is visualized in figure 2.1, which we will discuss later in detail. But we mention already now that, while simulations with Wilson fermions [79] – represented by open squares – had to be stopped* at values of $m_{\text{PS}} \approx 600$ MeV, our data – represented by open squares – reach with controllable errors down to significantly smaller values of m_{PS} . Thus, we can finally enter a region of mass values, where contact to chiral perturbation theory (χ PT) can be made, without worrying about lattice artifacts and convergence problems of χ PT. The tool that made this possible is mtmQCD, as introduced in section 1.2.4.

However, mtmQCD is still a rather new formulation of lattice QCD and hence we had to ascertain and scrutinize this approach. Most important for our understanding of mtmQCD has been a detailed scaling test, which we performed in a wide range of lattice spacings and quark masses in the quenched approximation. In particular, the range of m_{PS} values has been between 270 and 1200 MeV and the range of lattice spacings between 0.048 and 0.17 fm. We show that indeed lattice artifacts linear in a are absent in physical observables and that residual lattice spacing artifacts are small at *all* these values of m_{PS} . We perform then continuum extrapolations for the pseudo scalar decay constant and the vector mass finding full agreement with

*In the literature one can find also simulations with Wilson fermions and smaller values of the pseudo scalar mass (cf. for instance [80]). However, in those simulations single, so-called exceptional configurations had to be removed from the ensemble “by hand”.

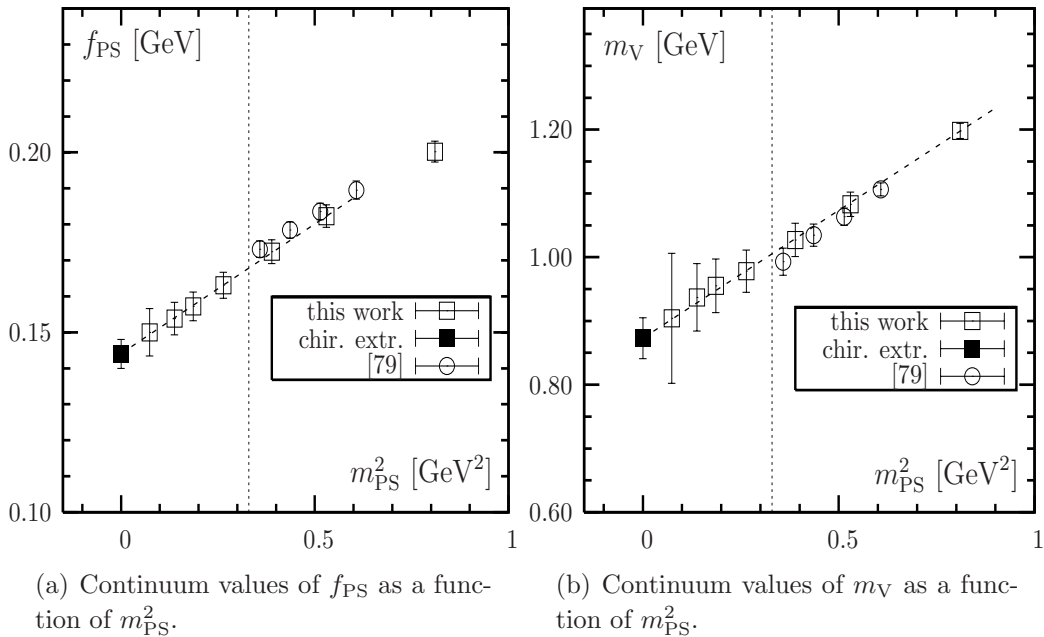


Figure 2.1: Continuum values of f_{PS} [GeV] (a) and m_V [GeV] (b) as a function of m_{PS}^2 [GeV²] and the chiral extrapolation of our data. In addition to our results, represented by squares, we also plot continuum results extracted from data published in [79] where non-perturbatively $\mathcal{O}(a)$ improved Wilson fermions were used. For the chiral extrapolation we used only our results and we indicate the linear extrapolation by the dashed lines. The dotted vertical lines roughly mark the value of m_{PS} where simulations with Wilson fermions had to be stopped.

results available in the literature. In addition we address the important question about size and scaling behavior of flavor breaking effects in mtmQCD, which vanish as a^2 , while being non-negligible of size.

Moreover, we compare at one value of the lattice spacing mtmQCD with the overlap formulation of lattice QCD, the latter of which has exact chiral symmetry at finite values of the lattice spacing a , as explained in section 1.2.2. We show that for all the quantities investigated here both formulations reveal consistent results. In particular, it is possible to simulate with both formulations pseudo scalar masses lower than 300 MeV without practical problems. However, a cost comparison yields that the overlap formulation is a factor of 20 to 70 more expensive than mtmQCD.

2.1 The definition of the critical mass

Naturally, when a new formulation is under investigation there are a lot of subtleties to understand and to learn. One of those in case of mtmQCD is the definition of the critical mass, whose potential influence on the cut-off effects will be discussed in the following section. However, beforehand we would like to stress the following remarks:

2.1. THE DEFINITION OF THE CRITICAL MASS

the proof of $\mathcal{O}(a)$ improvement for mtmQCD is based on the Symanzik effective theory [12, 13, 14], implying that the proof is only valid if the Symanzik expansion itself is valid. In particular, the lattice spacing should be sufficiently smaller than the physical scale. Moreover, since the Symanzik expansion is a perturbative concept, it is assumed that the dependence of quantities under investigation on the lattice spacing is smooth. Especially, in the vicinity of a dis-continuous phase transition – as we will find in the case of dynamical twisted mass fermions in chapter 4 – we cannot expect the expansion to hold.

The only parameter that needs to be tuned in tmQCD in order to obtain $\mathcal{O}(a)$ improvement is the twist angle. Maximal twist is achieved by tuning the angle to a value of $\pi/2$, which is equivalent to setting the value of κ to its critical value κ_{crit} (1-30), or equivalently the value of the (untwisted) bare quark mass m_0 to its critical value m_{crit} . In other words, we are interested in a situation where the renormalized quark mass is determined *only* by the twisted mass $m_{\text{R}} \sim \mu$, i.e. the twisted mass term determines physics. However, any lattice determination of κ_{crit} is affected by errors, which in turn means that the (untwisted) bare quark mass is zero only up to an error: $m_q = 0 + \delta$. This error can have significant consequences for the $\mathcal{O}(a)$ improvement, which we exemplify in the following. Consider again the renormalized quark mass m_{R} at finite lattice spacing. Neglecting the renormalization factors for a moment, which are of order one, this quantity can be estimated by a combination of the twisted mass parameter μ and the untwisted quark mass m_q to be

$$m_{\text{R}} \sim \sqrt{\mu^2 + m_q^2} = \mu \left(1 + \frac{m_q^2}{2\mu^2} + \mathcal{O}(m_q^4/\mu^4) \right), \quad (2-1)$$

where the expansion is only valid if $m_q = \delta \ll \mu$. It is evident from Eq. (2-1) that if $\delta \ll \mu$ is fulfilled the residual value of m_q contributes only as a small quadratical correction. On the other hand, if $\delta \ll \mu$ is not fulfilled physics is no longer dominated by the twisted mass term as it should be at maximal twist. Hence, it is important for $\mathcal{O}(a)$ improvement to use a κ_{crit} determination with δ as small as possible.

Note that if $\delta \approx \Delta a + \mathcal{O}(a^2)$ with a coefficient Δ one can always find a value of a keeping μ fixed in physical units, where $\delta \ll \mu$ is fulfilled.

In practice, there are at least two ways to determine the value of κ_{crit} : the first is to determine the value of κ where the value of am_{PS} vanishes with pure Wilson fermions, i.e. $\mu = 0$. We will refer to this determination method as the *pion definition* of κ_{crit} and denote it with $\kappa_{\text{crit}}^{\text{pion}}$. This determination involves an extrapolation of $(am_{\text{PS}})^2$ in κ to the κ value where am_{PS} vanishes. The extrapolation contains usually a large uncertainty because simulations with pure Wilson fermions and small pseudo scalar masses are hardly possible and therefore, one has to extrapolate from rather large masses. Unfortunately, the size of the extrapolation error is unknown and cannot easily be parameterized in terms of the lattice spacing.

CHAPTER 2. QUENCHED SCALING TEST

β	5.7	5.85	6.0	6.1	6.2	6.45
r_0/a	2.924	4.067	5.368	6.324	7.360	10.41
$L^3 \times T$	$14^3 \times 28$	$16^3 \times 32$	$16^3 \times 32$	$20^3 \times 40$	$24^3 \times 48$	$32^3 \times 64$
$N_{\text{meas}}^{\text{pion}}$	600	378	387	300	260	182
$N_{\text{meas}}^{\text{PCAC}}$	600	500	400	-	300	-

Table 2.1: For the six values of β this table contains the value for r_0/a , the lattice size and the number of measurements $N_{\text{meas}}^{\text{pion}}$ with the pion definition and $N_{\text{meas}}^{\text{PCAC}}$ with the PCAC definition of κ_{crit} .

The second definition of κ_{crit} we consider here makes use of the PCAC relation: At fixed non-zero value of the twisted mass parameter μ the value of κ needs to be determined where the value of m_χ^{PCAC} vanishes. The resulting value $\kappa_{\text{crit}}(a\mu)$ is still depending on the value of μ . This dependence can be removed by extrapolating $\kappa_{\text{crit}}(a\mu)$ to $a\mu = 0$. The extrapolation is only short in μ since simulations with a twisted mass parameter can be safely performed also with small values of $a\mu$ and therefore the extrapolation error is assumed to be small. We will refer to this definition of κ_{crit} as the *PCAC definition* and denote it with $\kappa_{\text{crit}}^{\text{PCAC}}$. Note that one main difference between the two definitions is that for $\kappa_{\text{crit}}^{\text{pion}}$ pure Wilson fermions are used, while for $\kappa_{\text{crit}}^{\text{PCAC}}$ the tmQCD regularization is used.

In this chapter we will use both of the two definitions for κ_{crit} and compare the residual lattice artifacts in physical observables between the pion definition and the PCAC definition for various κ values of the quark mass and the lattice spacing.

We remark here that the “optimal” definition of κ_{crit} with respect to $\mathcal{O}(a)$ improvement with mtmQCD is theoretically not yet clarified. Nevertheless, a detailed discussion of lattice artifacts of mtmQCD can be found in Ref. [45]. In fact, it was shown in this reference that in the Symanzik expansion of an operator O at maximal twist there appear at order a^2 terms proportional to $1/m_{\text{PS}}^4$, which are called *leading “infra-red divergent” cut-off effects*. These terms will be strongly visible as large $\mathcal{O}(a^2)$ cut-off effects in the limit of vanishing pseudo scalar masses. In the same reference it was shown that with the PCAC definition of κ_{crit} the coefficients multiplying the $(a/m_{\text{PS}}^2)^2$ terms vanish faster in the chiral limit than $1/m_{\text{PS}}^4$ curing such the problem of those large cut-off effects. Note that in Refs. [81, 82] the same proposal was made, including also arguments from χ PT. For a recent discussion see also Refs. [83, 84].

2.2 Scaling test set-up

In order to verify the prediction of Ref. [42] of $\mathcal{O}(a)$ improvement at $\omega = \pi/2$ we have chosen up to six values of the bare coupling constant β in a range of lattice spacing between 0.048 fm and 0.17 fm. We used the Wilson plaquette gauge action (1-25) and periodic boundary conditions for gauge and fermion fields. The β values can be found together with the values for r_0/a and the lattice sizes in table 2.1. The number of measurements N_{meas} for the two κ_{crit} definitions can also be found in table 2.1. We have set the Wilson parameter to $r = 1$ and used $r_0 = 0.5$ fm to set the scale throughout this chapter.

In order to fix the physical situation in our scaling test, we decided to study physical quantities as a function of β (i.e. as a function of a) for fixed values of $r_0 m_{\text{PS}}$. For this purpose we roughly fixed the values of $r_0 \mu$ at each β value to $r_0 \mu \approx 0.02, 0.04, 0.08, 0.16, 0.24, 0.32, 0.40$. For the PCAC definition we have two additional intermediate values $r_0 \mu \approx 0.059$ and $r_0 \mu \approx 0.123$. Note that at $\beta = 6.45$ we have simulated only the two lightest quark masses with the pion definition in order to check our continuum extrapolations.

Due to the wide range of twisted mass values at each β value we could then interpolate the results when necessary to match a desired value of $r_0 m_{\text{PS}}$. For the quantities we considered here it was sufficient to perform a linear interpolation in $(r_0 m_{\text{PS}})^2$ between the two closest points. We used the ROOT and MINUIT packages from CERN (cf. [85, 86]) for these interpolations. Since we do not want to extrapolate to mass values where we have no data available, the lowest value of the pseudo scalar mass was 298 MeV with the pion definition of κ_{crit} and 270 MeV with the PCAC definition.

Since in this chapter we work in the quenched approximation the gauge configurations do not depend on the bare quark masses. We produced at each value of β an ensemble of gauge configurations using a combination of the over-relaxation and the heat-bath algorithm. One heat-bath sweep was always followed by $L/2 + 1$ over-relaxation sweeps, where L is the spatial lattice size. We skipped as many intermediate configurations as needed to obtain completely independent configurations.

Again, because the configuration at each β value are quark mass independent, we could make use of a multi mass conjugate gradient (CG-M) iterative solver as explained in appendix B.2. Such a solver allows one to invert on the lowest mass and get within the same inversion also the result for all the other masses. At $\beta = 6.45$, however, we used even/odd preconditioning, since we simulated only the two lowest values of $r_0 \mu$. This preconditioning accelerates the solvers, but prevents us to use a multi mass solver.

In order to compute the quark propagators needed in the contraction of the correlation functions, the Dirac operator needs to be inverted on a given source.

CHAPTER 2. QUENCHED SCALING TEST

$a\mu$	am_{PS}			$am_{\text{PS}}L _{(L=16)}$
	$12^3 \times 24$	$14^2 \times 32$	$16^3 \times 32$	
0.005	-	-	0.1700(25)	2.7
0.01	0.2327(70)	0.2301(37)	0.2254(19)	3.6
0.02	0.3193(48)	0.3175(30)	0.3122(16)	5.0
0.04	0.4520(40)	0.4506(23)	0.4452(14)	7.1
0.06	0.5596(35)	0.5575(19)	0.5535(12)	8.9
0.08	0.6541(31)	0.6510(17)	0.6488(11)	10.4
0.10	0.7417(26)	0.7378(16)	0.7359(11)	11.8

Table 2.2: Values of am_{PS} at $\beta = 5.85$ for three different lattice Volumes: $12^3 \times 24$, $14^2 \times 32$ and $16^3 \times 32$. In the last column we also give the value of $am_{\text{PS}}L$ for the $16^3 \times 32$ lattice.

The simplest choice here is to use a point source located at $x = 0$ and for every combination of color and spinor indices. Therefore, 12 inversions are needed per configuration. We used a point source for all the results with the pion definition of κ_{crit} . However, it is known that the overlap with the ground state can be improved if (sink) smearing techniques are applied, which turned out to be crucial in order to determine a reliable estimate for the vector meson mass m_V . We used Jacobi sink smearing [87] for the determination of m_V with the PCAC definition of κ_{crit} . Since we did not use smearing techniques with the pion definition, we were not able to reliably extract values for am_V in this case.

2.2.1 Finite volume effects

Before presenting the results of the scaling test, we show in this subsection results at one value of $\beta = 5.85$ for three different lattice volumes in order to check for finite volume effects. At this value of β we have performed 140 measurements on a $12^3 \times 24$ lattice, 140 measurements on a $14^3 \times 32$ lattice and 380 measurements on a $16^3 \times 32$ lattice, corresponding to physical spatial extends of about 1.48 fm, 1.72 fm and 1.96 fm, respectively. The hopping parameter was set to its critical value $\kappa_{\text{crit}}^{\text{pion}} = 0.161662(17)$ obtained with the pion definition of κ_{crit} . We measured the values of the pseudo scalar mass for all three volumes and collected the data in table 2.2.

Aiming for an analysis along the lines of Ref. [88], we first extrapolated the pseudo scalar masses to $L = \infty$ by fitting a functional form

$$am_{\text{PS}}(L) = am_{\text{PS}}(L = \infty) + \frac{a_1}{L^{3/2}} \exp\{-a_2 am_{\text{PS}}(L = \infty)L\} \quad (2-2)$$

with coefficient a_1 and a_2 to our data. For the three values of $a\mu$ between 0.01 and 0.04 we show the data points together with the fits in figure 2.2(a). From this figure

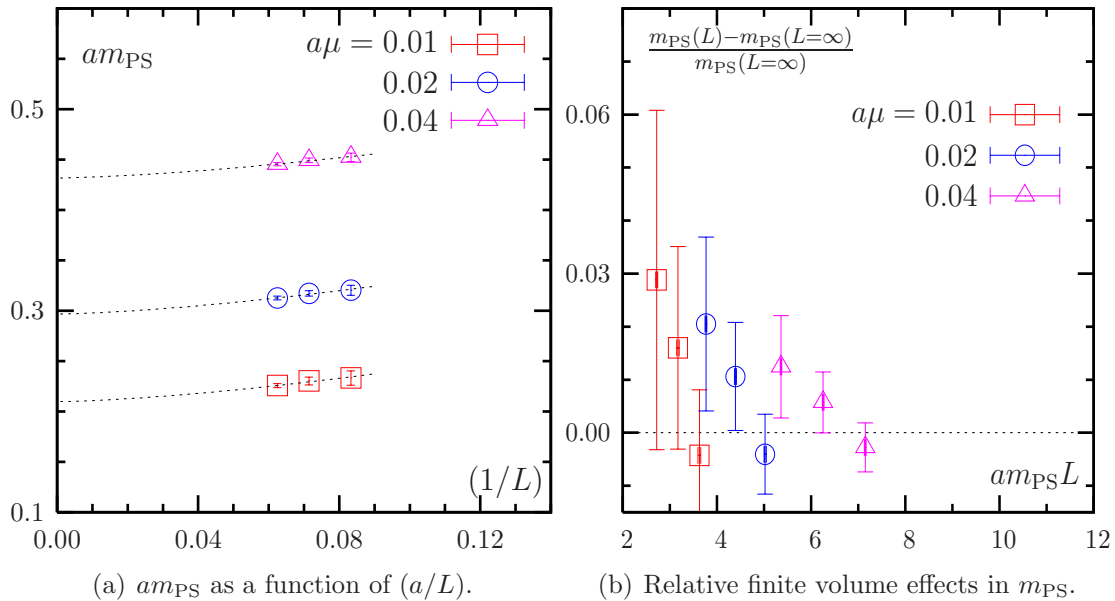


Figure 2.2: Finite volume effects in m_{PS} at $\beta = 5.85$. In (a) am_{PS} is plotted as a function of $(1/L)$ together with exponential fits to the data (see text). In (b) the relative finite volume effects are plotted as a function of $am_{\text{PS}}L$. In both graphs we plot only data for the three lowest values of $a\mu$. Note that in our notation L is dimensionless.

it is already evident that finite volume effects, at least for the quantity am_{PS} are small. An analysis of $(m_{\text{PS}}(L) - m_{\text{PS}}(L = \infty))/m_{\text{PS}}(L = \infty)$ (see figure 2.2(b)) shows that for the simulation points corresponding to the smallest values of $a\mu$ the finite volume effects are within 2 – 3 percent and at most within two standard deviations from the extrapolated infinite volume limit. In practice they are thus not relevant for the following discussion.

Other quantities than m_{PS} are affected by finite volume artifacts of qualitatively the same form as Eq. (2-2) with, of course, in general different coefficients. However, since for the following scaling test we will stay at almost constant physical volume for all values of β under investigation, the scaling test itself should not be affected by finite volume effects.

2.3 Scaling test of mtmQCD

In this section we present the numerical results of the scaling test. We first compare the scaling behavior for the two definitions of κ_{crit} , present then continuum results for f_{PS} and m_{V} and investigate the flavor breaking effects. Finally, we compare the twisted mass and the overlap formulation.

β	$\kappa_{\text{crit}}^{\text{pion}}$	$\kappa_{\text{crit}}^{\text{PCAC}}$
5.7	0.169198(48)	0.171013(160)
5.85	0.161662(17)	0.162379(93)
6.0	0.156911(35)	0.157409(72)
6.1	0.154876(10)	-
6.2	0.153199(16)	0.153447(32)
6.45	0.150009(11)	-

Table 2.3: Values of $\kappa_{\text{crit}}^{\text{pion}}$ and $\kappa_{\text{crit}}^{\text{PCAC}}$ for all values of β .

2.3.1 Scaling of the pseudo scalar decay constant with the pion definition of κ_{crit}

At all the β values of our simulations we made our own determination of the value of $\kappa_{\text{crit}}^{\text{pion}}$ from the intercept in κ at zero pseudo scalar mass. The values of $\kappa_{\text{crit}}^{\text{pion}}$ are given in table 2.3. In a first step we then computed the values of am_{PS} at each of our simulation points. The values can be found in table C.1 in appendix C.

Since we work at maximal twist at each value of β we can extract the pseudo scalar decay constant f_{PS} as explained in section 1.3 Eq. (1-84). This prescription allows the extraction of f_{PS} without the need to compute renormalization constants. The results are collected in table C.2 and they are visualized in figure 2.3, where we show $r_0 f_{\text{PS}}$ as a function of $(a/r_0)^2$. The different symbols correspond from top to bottom to about $m_{\text{PS}} = 720$ MeV, 515 MeV, 380 MeV and 300 MeV. In addition to our data represented by the open symbols we also plot our continuum extrapolation represented by filled symbols (for better visibility some points are slightly displaced). For the continuum extrapolations we used only the data for $\beta \geq 6.0$.

For the interpretation of figure 2.3 it is important to remind that – since we use a multi mass solver – the results at one lattice spacing for different masses are strongly correlated, which explains the similar fluctuations at fixed a/r_0 for the different values of $r_0 m_{\text{PS}}$. First of all it is evident from figure 2.3 that for all values of $r_0 m_{\text{PS}}$ plotted the cut-off effects are linear in $(a/r_0)^2$ for lattice spacings lower than a given bound. From figure 2.3 we can estimate this bound to be $(a/r_0)^2 < 0.04$. The slope of the continuum extrapolation, however, is strongly mass dependent: it becomes steeper the lower the pseudo scalar mass value becomes. This can only be explained with a dependence of the $\mathcal{O}(a^2)$ cut-off effects on the mass as described in Ref. [45]. The size of these cut-off effects can be significantly reduced by using the PCAC definition of κ_{crit} , as will be shown in the following subsections.

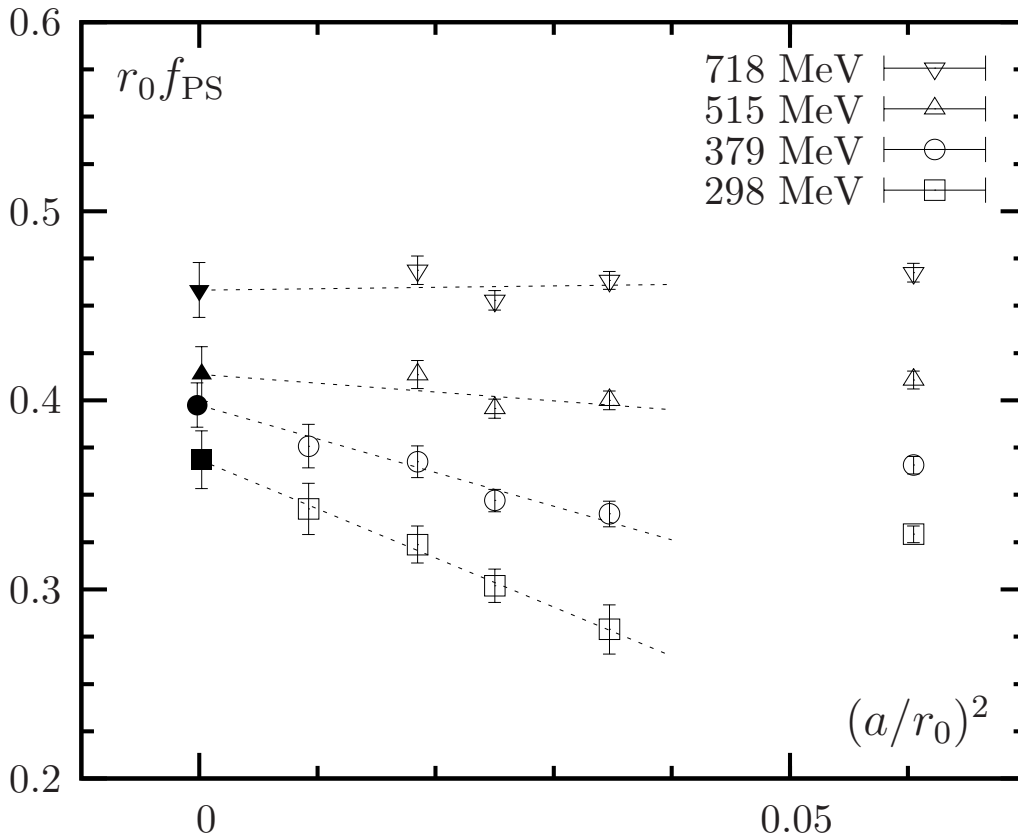


Figure 2.3: Scaling plot for $r_0 f_{PS}$ as a function of $(a/r_0)^2$ for four different values of $r_0 m_{PS}$. The different symbols correspond from top to bottom to 718 MeV, 515 MeV, 379 MeV and 298 MeV. Our data points are represented by open symbols while the continuum extrapolations are plotted with filled symbols

2.3.2 Critical mass from the PCAC relation

With the PCAC definition we considered only the β values 5.7, 5.85, 6.0 and 6.2. At each of these β values we first determined the values of $\kappa_{\text{crit}}^{\text{PCAC}}$ with the method explained above and collected them in table 2.3. An example for this determination at $\beta = 5.7$ can be found in figure 2.4, where we show in the left panel the interpolation in $1/\kappa$ to the point where $m_{\chi}^{\text{PCAC}} = 0$ and in the right panel the extrapolation of $\kappa_{\text{crit}}(a\mu)$ to $a\mu = 0$. With the straight line we indicate the linear extrapolation to $a\mu = 0$ and in addition we included the value of κ_{crit} determined by the pion definition in the figure. The difference is supposed to be of $\mathcal{O}(a)$. The errors on the values of $\kappa_{\text{crit}}^{\text{PCAC}}$ in table 2.3 stem from the necessary inter- and extrapolation.

The first quantity we investigated was again the pseudo scalar meson mass m_{PS} . The data are collected in table C.3 in appendix C. In figure 2.5(a) $(am_{PS})^2$ is plotted as a function of $a\mu$ at $\beta = 6.0$ for the two definitions of κ_{crit} . The data points for the PCAC definition show a linear behavior in $a\mu$ down to very small bare quark masses, which is not observed with the pion definition of κ_{crit} . This effect is better

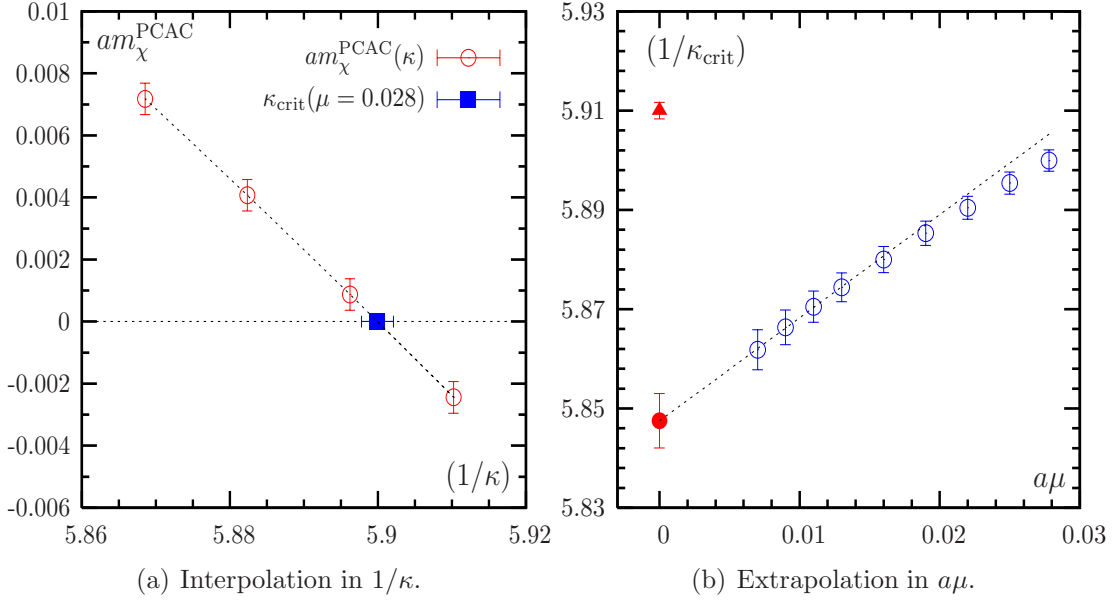


Figure 2.4: Determination of the critical hopping parameter from the PCAC definition at $\beta = 5.7$: (a) interpolation of $am_\chi^{\text{PCAC}}(1/\kappa)$ to $m_\chi^{\text{PCAC}} = 0$ for one value of the twisted mass parameter $a\mu = 0.028$ (b) $1/\kappa_{\text{crit}}$ versus $a\mu$, extrapolation to $a\mu = 0$, the triangle indicates the $1/\kappa_{\text{crit}}$ value determined by $(am_{\text{PS}})^2 \rightarrow 0$ at $\mu = 0$ for unimproved Wilson fermions.

visible in figure 2.5(b), where we plot $(am_{\text{PS}})^2/(a\mu)$ as function of $a\mu$. The same behavior is observed for the other β values. This shows that the GMOR relation [89] on the lattice

$$m_{\text{PS}}^2 = \mu \frac{2|\langle 0|P|\pi\rangle|^2}{\Sigma} + \mathcal{O}(a^2), \quad (2-3)$$

is not affected by large cut-off effects (see Ref. [45]) when the PCAC definition is used for the determination of κ_{crit} . Relation (2-3) can be derived from a vector variation of the charged pseudo scalar density, similar to the pure Wilson case [90]. It allows the extraction of the scalar condensate Σ with the need of only the renormalization factor $Z_P = 1/Z_\mu$. Our determination of Z_P and Σ is ongoing [91].

2.3.3 Scaling of f_{PS} with the PCAC definition of κ_{crit}

As a next quantity we determined f_{PS} from the data with the PCAC definition of κ_{crit} . The values can be found in table C.4 in appendix C. In order to compare to the results obtained with the pion definition we had to match the values of $r_0 m_{\text{PS}}$ by interpolating our results.

First of all, for all our values of $r_0 m_{\text{PS}}$ the observed cut-off effects in f_{PS} are linear in $(a/r_0)^2$. Moreover, if we consider at several values of $r_0 m_{\text{PS}}$ the size of the $\mathcal{O}(a^2)$ cut-off effects, we find that with the PCAC definition of κ_{crit} their size is significantly reduced for small values of $r_0 m_{\text{PS}}$ when compared to the pion definition.

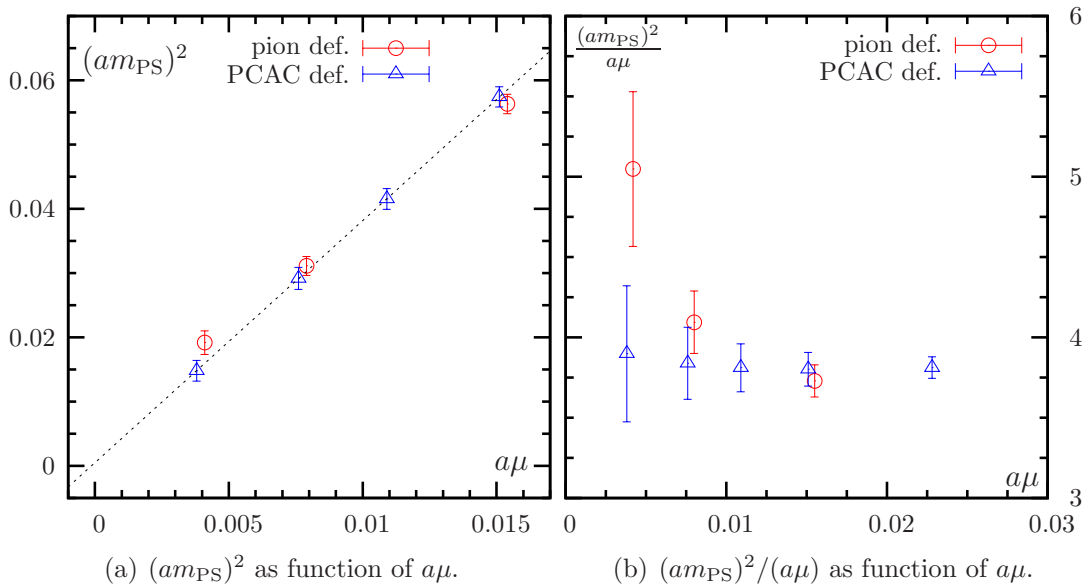


Figure 2.5: (a) Pseudo scalar mass squared as a function of $a\mu$ both, for the pion definition and the PCAC definition of κ_{crit} at $\beta = 6.0$. (b) $(am_{\text{PS}})^2 / (a\mu)$ for both κ_{crit} definitions as function of $a\mu$ at the same value of β . The pion definition data are in (a) and (b) slightly displaced for better visibility.

For the values of $m_{\text{PS}} = 298$ MeV and $m_{\text{PS}} = 515$ MeV we have plotted f_{PS} in figure 2.6 as a function of $(a/r_0)^2$ for both definitions of κ_{crit} . It is evident that down to a pseudo scalar mass of 298 MeV the extrapolation of the PCAC definition data is essentially flat. Moreover, the continuum values extrapolated separately for the pion definition and the PCAC definition data agree very well within the errors.

As expected, at small values of the pseudo scalar mass the size of the residual cut-off effects is significantly smaller if the PCAC definition is used instead of the pion definition.

2.3.4 Scaling of the vector meson mass with the PCAC definition of κ_{crit}

The second quantity we used to check the prediction of automatic $\mathcal{O}(a)$ improvement is the vector mass m_V . At maximal twist this mass can be extracted from the exponential fall of the two point correlation function C_{AA} or C_{TT} in the twisted basis as explained in section 1.3. The extraction of a value for m_V is difficult without smearing techniques, which is in particular the case for small quark masses. Therefore, we used local source and Jacobi smeared sinks to extract the vector meson mass. In addition it turned out that the tensor correlator systematically shows smaller statistical fluctuations and thus, we used exclusively the tensor correlator for the determination of values for am_V . As explained above we have for m_V only

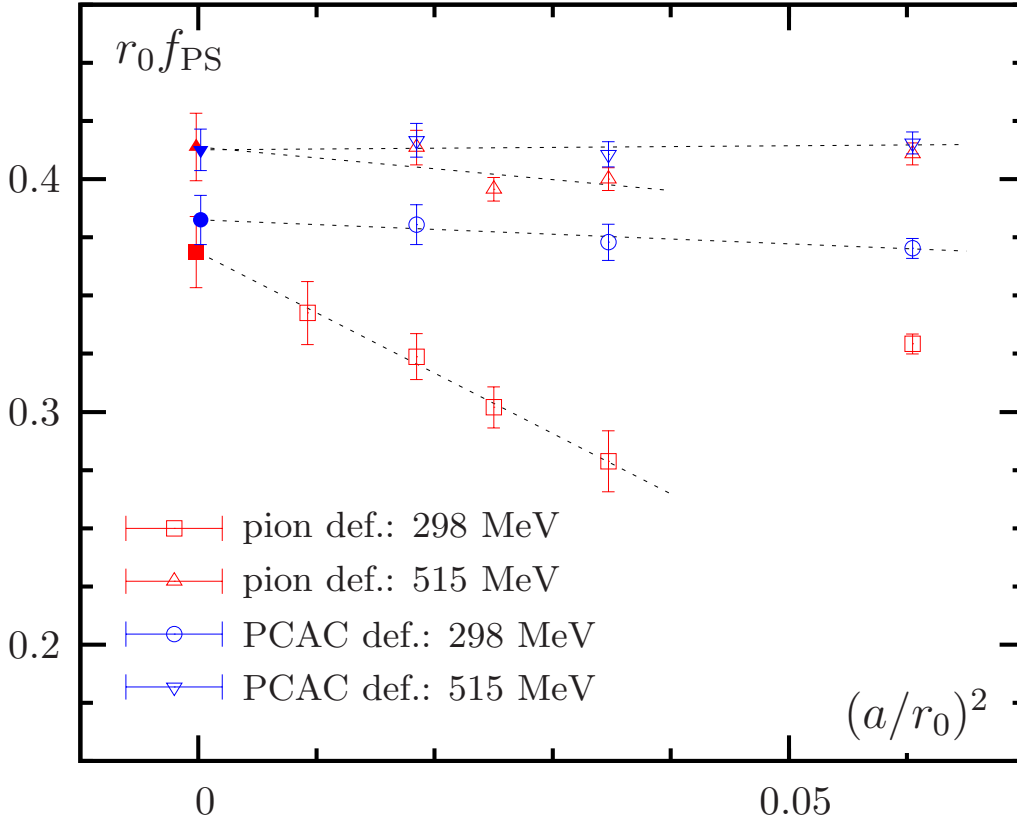


Figure 2.6: Scaling plot for $r_0 f_{PS}$ as a function of $(a/r_0)^2$ for $m_{PS} = 298$ MeV (open circles and open squares) and $m_{PS} = 515$ MeV (open triangles and open inverted triangles). The data points are represented by open symbols while the continuum extrapolations are plotted with filled symbols. Open squares and open triangles represent data obtained with the pion definition of κ_{crit} , while the open circles and the open reversed triangles represent data obtained with the PCAC definition. The PCAC definition continuum points are slightly displaced for better visibility.

results obtained with the PCAC definition of κ_{crit} .

The values for am_V for all our simulation points with the PCAC definition can be found in table C.5. In fig. 2.7 we show our results for the vector meson mass as a function of $(a/r_0)^2$ for values of $m_{PS} = 900$ MeV, 730 MeV and 270 MeV. As observed for f_{PS} , the continuum extrapolations for m_V are essentially flat down to pseudo scalar masses of 270 MeV for $(a/r_0)^2 \leq 0.06$, confirming again the expected $\mathcal{O}(a)$ improvement with mtmQCD.

2.3.5 Continuum extrapolation

The continuum extrapolated values of the pseudo scalar decay constant and the vector mass for nine values of m_{PS} are summarized in table C.6. The values quoted there have been extracted only from the data obtained with the PCAC definition, because

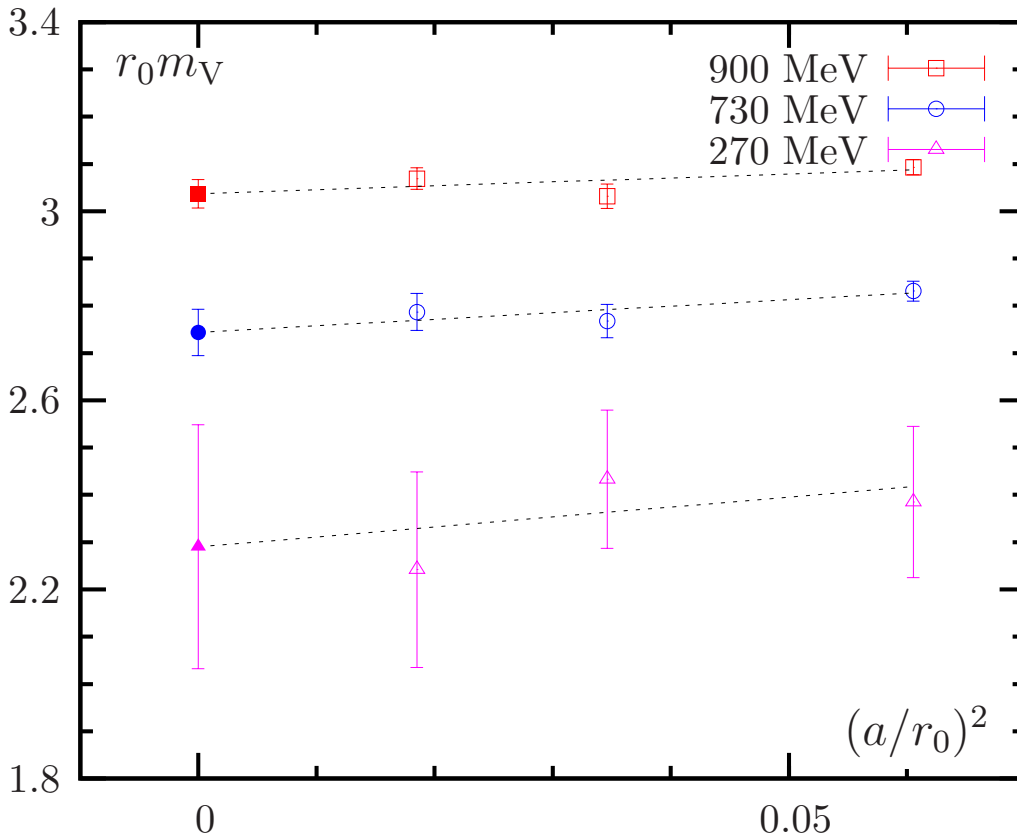


Figure 2.7: $r_0 m_V$ as a function of $(a/r_0)^2$ for three values of $r_0 m_{PS}$ with the PCAC definition of κ_{crit} .

they show much less residual lattice artifacts and the continuum extrapolations are thus more reliable.

In the figures 2.1(a) and 2.1(b), which can be found at the beginning of this chapter, we plot the continuum values of f_{PS} and m_V as functions of m_{PS}^2 in physical units. As a comparison we also plot continuum values that we extracted from the data presented in Ref. [79] where non-perturbatively $\mathcal{O}(a)$ improved Wilson fermions have been used. Both quantities show a linear behavior in the pseudo scalar mass squared without signs of artifacts as predicted from quenched chiral perturbation theory (proportional to $\log a$). Moreover, our results fully agree with the results extracted from Ref. [79], at least for the large values of m_{PS} where a comparison is possible. We extrapolated our results linearly to the chiral limit which is indicated by the dashed lines in the two panels of figure 2.1. The values for f_{PS} and m_V in the chiral limit are collected in table 2.4 together with the values of f_π , m_ρ , f_K and m_{K^*} (the last two in the $SU(3)$ symmetric limit). They were obtained either through the extrapolation to the chiral point (chiral limit, f_π , m_ρ), or by an interpolation (f_K , m_{K^*}).

The ratio $f_K/f_\pi = 1.11(5)$ is 10% smaller than the experimentally obtained

m_{PS} [GeV]	f_{PS} [GeV]	m_V [GeV]	
0.0	0.144(4)	0.873(32)	chiral limit
0.137	0.145(4)	0.880(32)	(f_π, m_ρ)
0.495	0.162(5)	0.971(34)	(f_K, m_{K^*})

Table 2.4: f_{PS} and m_V in the continuum (PCAC definition). The values are obtained from a linear fit on the smallest 4 masses (5 in the case of m_V) and correspond to: the values in the chiral limit (first row); f_π and m_ρ (second row); f_K and m_{K^*} in the $SU(3)$ symmetric limit (third row).

values. This deviation is, however, consistent with what was observed in previous quenched calculations [92, 80]. Moreover, the values of m_ρ and m_{K^*} turn out to be 10 – 15% larger than the experimental values again consistent with other quenched calculations [80].

2.3.6 Flavor breaking effects

As mentioned in section 1.2.4 the flavor chiral rotation to the twisted basis is in the continuum only a formal transformation that leaves the theory invariant. Therefore, even if flavor symmetry seems to be broken in the twisted basis, in the continuum it is only replaced by a modified, but equivalent flavor symmetry. This is not longer true at finite lattice spacing where the flavor symmetry is explicitly broken and only restored in the continuum limit. This manifests itself for instance as a non-vanishing difference of the charged m_{PS^+} and the neutral m_{PS^0} pseudo scalar masses, which is expected to vanish towards the continuum as a^2 . It is an important question whether this expectation proves true in practice.

If we consider the local bilinears $P^\pm = \bar{\chi}\gamma_5\frac{\tau^\pm}{2}\chi$ and $P^0 = \bar{\chi}\chi$ similar to what we explained in section 1.3, we can extract values for am_{PS^+} and am_{PS^0} from the following correlation functions:

$$\begin{aligned}
 C_{m_{\text{PS}^+}}(t) &= a^3 \sum_{\mathbf{x}} \langle [P^+(x)P^-(0)]_{\text{con}} \rangle , \\
 C_{m_{\text{PS}^0}}(t) &= a^3 \sum_{\mathbf{x}} \langle [P^0(x)P^0(0)]_{\text{con}} + [P^0(x)P^0(0)]_{\text{disc}} \rangle .
 \end{aligned}
 \tag{2-4}$$

Here we denote with $[\cdot]$ the fermionic contractions only and indicate with the subscripts “con” and “disc” the connected and the disconnected pieces of the correlation function, respectively. For the neutral pseudo scalar mass it is thus in general needed to evaluate the disconnected contribution

$$[P^0(x)P^0(0)]_{\text{disc}} = \text{Tr} \{ D_{\text{tm}}^{-1} \} (x) \text{Tr} \{ D_{\text{tm}}^{-1} \} (0) .
 \tag{2-5}$$

2.3. SCALING TEST OF MTMQCD

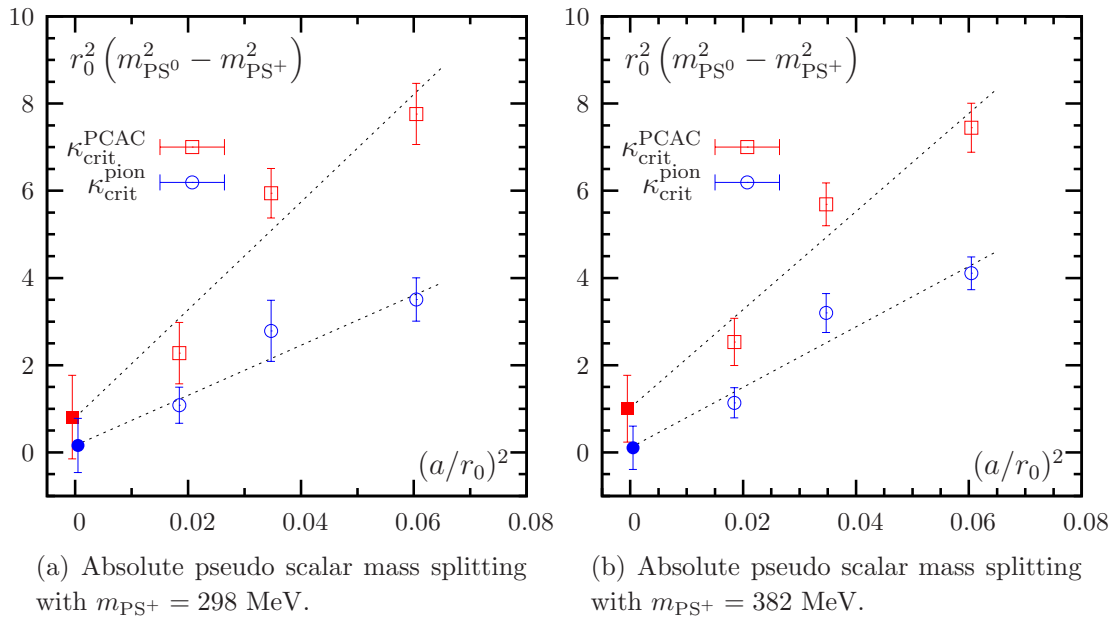


Figure 2.8: Absolute pseudo scalar mass difference as a function of $(a/r_0)^2$ for $m_{\text{PS}^+} = 298$ MeV and $m_{\text{PS}^+} = 382$ MeV employing the pion definition and the PCAC definition for the critical point.

with the vacuum contribution to $\text{Tr} \{D_{\text{tm}}^{-1}\}$ being subtracted and the trace is taken over color and Dirac indices. There are several techniques available to compute those contributions on the lattice, which are mostly based on stochastic estimators and usually rather computer time demanding (see for instance Refs. [93, 94]). However, since we work in the quenched approximation there is a possibility to investigate the aforementioned mass splitting from the connected piece in $C_{m_{\text{PS}^0}}$ only, by reinterpreting the connected piece with the help of the Osterwalder-Seiler (OS) action [95, 96].

The Osterwalder-Seiler action is identical to the twisted mass action, only the τ_3 matrix acting in flavor space is replaced with the unit matrix. Thus, $m_{\text{PS}^+} = m_{\text{PS}^0}$ because flavor symmetry is un-broken and

$$C_{m_{\text{PS}^+}}^{\text{OS}}(t) = C_{m_{\text{PS}^0}}^{\text{OS}}(t). \quad (2-6)$$

The key observation is now that in the quenched approximation, and only in the quenched approximation

$$C_{m_{\text{PS}^0}}^{\text{OS}}(t) = \left(C_{m_{\text{PS}^0}}^{\text{tm}}(t) \right)_{\text{con}}, \quad (2-7)$$

allowing us to interpret $(C_{m_{\text{PS}^0}}^{\text{tm}}(t))_{\text{con}}$ as the correlation function of a local operator. Hence, $(C_{m_{\text{PS}^0}}^{\text{tm}}(t))_{\text{con}}$ has a standard transfer matrix decomposition and the neutral pseudo scalar mass can be extracted from its exponential decay with good precision.

We have extracted values for am_{PS^0} for $\beta = 5.85$, $\beta = 6.0$ and $\beta = 6.2$ with the pion definition and the PCAC definition of κ_{crit} . They can be found in the tables of Ref. [96]. In figure 2.8 we show the mass difference $r_0^2(m_{\text{PS}^0}^2 - m_{\text{PS}^+}^2)$ as a function of $(a/r_0)^2$ for both definitions of κ_{crit} and for values of $m_{\text{PS}^+} = 298$ MeV and 387 MeV. As expected, the difference vanishes proportional to a^2 and the continuum extrapolated values are consistent with zero. However, the mass splitting is not small and in addition its size differs significantly for the two definition of κ_{crit} : the results obtained with the PCAC definition show larger flavor breaking effects. This is at first unexpected, because for all the other quantities used in this scaling test the $\mathcal{O}(a^2)$ artifacts are smaller for the PCAC definition when compared to the pion definition.

Our interpretation for this phenomenon is again based on the symmetries of the twisted mass formulation. With the PCAC definition parity is maximally restored at finite lattice spacing, but at the same time flavor symmetry is maximally broken, which is expressed in chiral perturbation theory in the fact that the mass splitting is proportional to $\sin(\omega)$ [97, 82]. We also remark that the mass splitting ($m_{\text{PS}^0} - m_{\text{PS}^+}$) comes out to be positive which is consistent with the realization of an Aoki phase scenario in the quenched approximation (see chapter 4 for more details).

2.4 Overlap versus twisted mass fermions

In the last section we have demonstrated that with mtmQCD $\mathcal{O}(a)$ improvement can be obtained without the need of any improvement coefficient, which is indeed a big advantage of this lattice QCD formulation. We mentioned in section 1.2.4 that tmQCD has a further advantage compared to the original Wilson formulation: the twisted mass parameter serves as an infra-red regulator for the low lying eigenvalues of the lattice Dirac operator [43]. Therefore, it is possible to perform simulations with pseudo scalar mass of about 270 MeV as we have also shown in the last section.

The tmQCD formulation shares these two properties with the overlap formulation, while the latter has the additional property of exact chiral symmetry on the lattice. All this makes a comparison between the two formulations rather interesting. Unfortunately, as also discussed in section 1.2.2, the overlap formulation is much more computer time demanding than the twisted mass formulation making it impossible for us to repeat the scaling test as presented in the last section with the overlap operator. We could only afford for simulations with the overlap formulation at one value of $\beta = 5.85$ on lattices of size $12^3 \times 24$ (for a first scaling study see [98]). We used the overlap operator as defined in Eq. (1-38) with the parameter $\rho = 1.6$ fixed and bare masses of $am_{\text{ov}} = 0.01, 0.02, 0.04, 0.06, 0.08, 0.10$ matching six values for $a\mu$ for the twisted mass simulations at $\beta = 5.85$. We approximated the inverse

2.4. OVERLAP VERSUS TWISTED MASS FERMIONS

am_{ov}	$am_{\text{PS}}^{\text{PP-SS}}$	$aLm_{\text{PS}}^{\text{PP-SS}}$	af_{PS}	am_{V}
0.01	0.140(20)	1.6	0.0934(90)	0.632(34)
0.02	0.196(14)	2.3	0.1012(53)	0.638(26)
0.04	0.280(10)	3.4	0.1060(34)	0.653(16)
0.06	0.346(8)	4.2	0.1106(25)	0.666(12)
0.08	0.401(7)	4.8	0.1157(22)	0.683(09)
0.10	0.451(6)	5.4	0.1209(21)	0.702(08)

Table 2.5: Values for the pseudo scalar mass with the overlap formulation at $\beta = 5.85$ on $12^3 \times 24$ lattices. In addition we provide values for $aLm_{\text{PS}}^{\text{PP-SS}}$, the pseudo scalar decay constant and vector mass for overlap at $\beta = 5.85$.

square root by means of Chebyshev polynomials to an absolute accuracy of 10^{-15} . In the construction of the polynomial we also project out the lowest 20 eigenvalues of $Q^2 = (\gamma_5 D_{\text{W}})^2$. For details on the numerical treatment of the overlap operator see Ref. [99]. With the overlap operator we performed measurements on 140 gauge configurations and extracted values for am_{PS} , af_{PS} and am_{V} .

The pseudo scalar mass has been extracted from C_{PP-SS}^{ov} (cf. Eq. 1-88). The scalar meson state did not affect the extraction, presumably because the quark masses are small enough to let the pseudo scalar state be dominant at sufficiently small values of t/a . The values for am_{PS} can be found in table 2.5. In addition to these values we provide values for $aLm_{\text{PS}}^{\text{PP-SS}}$ which we used to estimate the finite volume effects. Given these values of $aLm_{\text{PS}}^{\text{PP-SS}}$ and the experience from tmQCD, we expect very small finite volume effects for the five heaviest quark masses (again at a level of a few percent). For the lowest mass finite volume effects can be more relevant and therefore we usually do not include the corresponding data points in fits. However, the analysis of the quantities presented below suggests that also for this value of the quark mass $am_{\text{ov}} = 0.01$ finite volume effects are not larger than our statistical error.

In figure 2.9 we show the pseudo scalar mass squared as a function of the bare mass am_{bare} for the overlap and the twisted mass formulation at $\beta = 5.85$. For the overlap formulation the bare mass corresponds to am_{ov} while for twisted mass fermions to $a\mu$. For the twisted mass data we took solely the results obtained with the PCAC definition of κ_{crit} . In addition to our twisted mass and overlap data we show data points obtained with standard $\mathcal{O}(a)$ improved Wilson fermions [100, 101].

From figure 2.9 one can see that at equal value of the bare quark mass the value of am_{PS} is smaller for overlap fermions than for twisted mass fermions. This suggests that the renormalization factor Z_μ of the quark mass is larger for Wilson twisted mass fermions compared to the corresponding factor for overlap fermions. Nevertheless, it is evident that with both formulations values of about $m_{\text{PS}} = 270$ MeV can

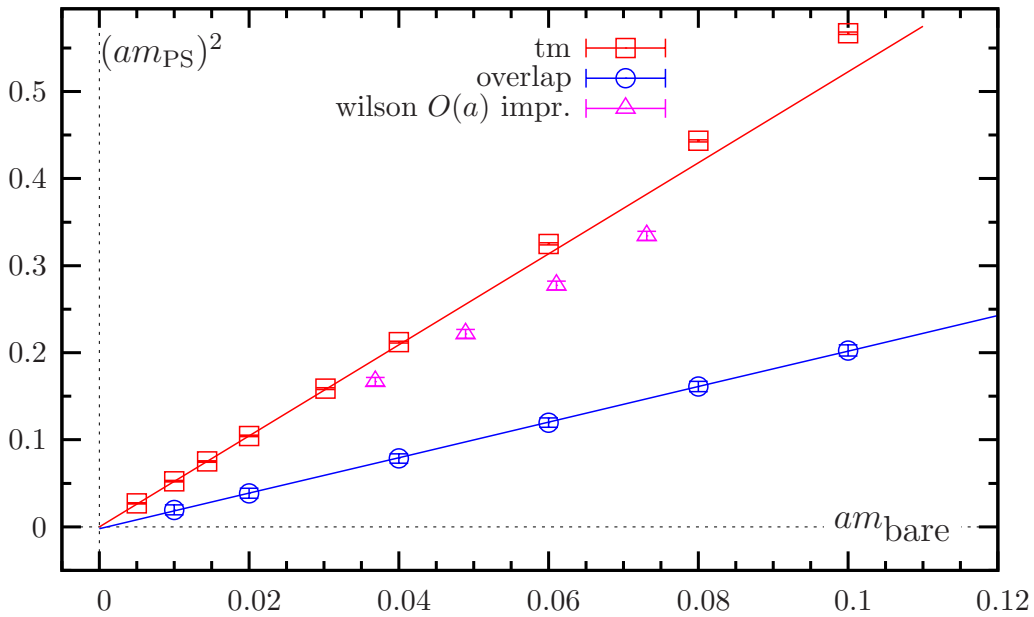


Figure 2.9: Comparison of am_{PS}^2 as a function of the bare quark mass between overlap and twisted mass formulation at $\beta = 5.85$. The overlap data are represented by open circles and the twisted mass data by open squares. In addition we plot data from simulations with $\mathcal{O}(a)$ improved Wilson fermions [100, 101]. m_{bare} labels the quark mass corresponding to μ for tmQCD and to m_{ov} the overlap. The solid lines are fits to the twisted mass and the overlap data. While for the fit we used in case of the overlap all the data points, for the twisted mass data we used only the four data points corresponding to the lowest pseudo scalar masses.

be reached without suffering from exceptional configurations. Moreover, for both formulations the squared pseudo scalar mass can be well approximated by a linear function of the bare quark mass down to the smallest mass values. In fact a linear extrapolation to the chiral limit gives in both cases a value for the intercept which is zero within the errors. For the twisted mass data we included only the lowest five masses in the fit, which was not necessary in case of overlap fermions.

In table 2.5 we have also collected the results for af_{PS} and am_{V} for the overlap operator at $\beta = 5.85$. The values for af_{PS} were determined using Eq. (1-93) and hence, do not require any renormalization constant. The vector mass m_{V} was extracted from $C_{\text{VV}}^{\text{ov}}$ (1-90).

Since we could perform the continuum extrapolation for these two quantities with the twisted mass formulation, we can compare the overlap results at $\beta = 5.85$ on the one hand to twisted mass results at the same β value and on the other hand to the continuum results. The result of this comparison can be found in the four panels of figure 2.10.

In this figure we plot r_0f_{PS} and r_0m_{V} as functions of $(r_0m_{\text{PS}})^2$. For the twisted mass results at $\beta = 5.85$ we use solely the data obtained with the PCAC definition

2.4. OVERLAP VERSUS TWISTED MASS FERMIONS

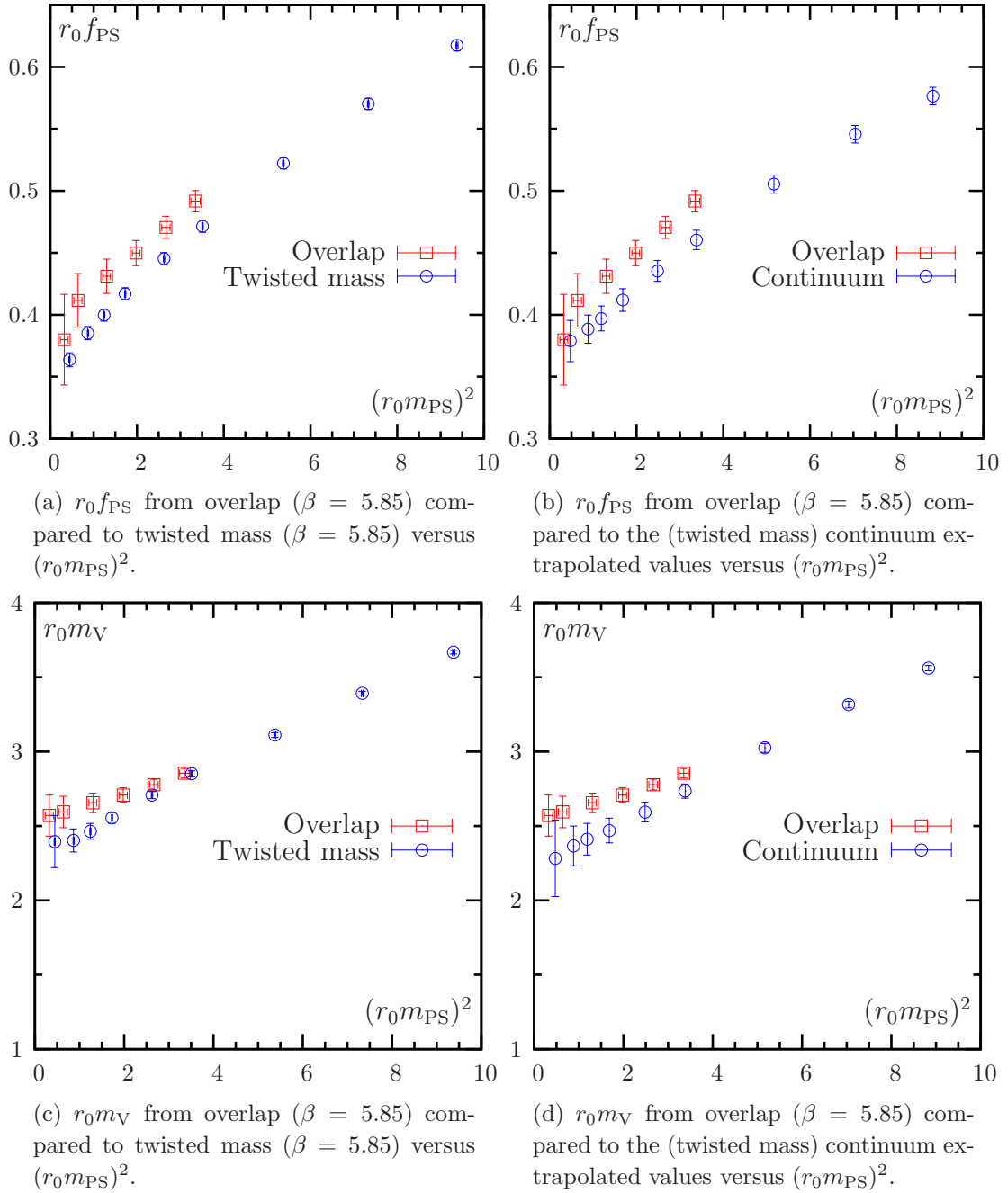


Figure 2.10: Comparison overlap and twisted mass formulation for f_{PS} and m_{V} .

of κ_{crit} . For both quantities a difference between the overlap results and the twisted mass results at finite values of the lattice spacing are visible, as well as between the overlap results and the continuum results. This suggests that the results obtained with the overlap operator are affected by small, but visible $\mathcal{O}(a^2)$ lattice artifacts.

Comparing 2.10(a) with 2.10(b) and 2.10(c) with 2.10(d) it is visible, that the data points for twisted mass at $\beta = 5.85$ and the continuum points are very close to

m_{PS} [MeV]	$a\mu$	am_{ov}
230	0.004	0.01
390	0.0125	0.03
555	0.025	0.06
720	0.042	0.10

Table 2.6: Bare quark masses for the cost comparison between overlap and twisted mass formulation at $\beta = 5.85$.

each other. This indicates again that the $\mathcal{O}(a^2)$ lattice artifacts in these quantities are very small.

2.4.1 Cost comparison

So far we have concentrated in this section on the comparison between the overlap and the twisted mass formulation of lattice QCD on the basis of physical results. We have seen that both the formulations are capable of simulations with pseudo scalar masses lower than 300 MeV. Despite the fact that the overlap formalism provides exact chiral symmetry on the lattice and is thus theoretically best founded for lattice simulations, it seems that twisted mass fermions can serve for many quantities as an equivalent alternative. This becomes even clearer by noting that with dynamical twisted mass fermions and a clever choice of the valence quark discretization mixing of operators with wrong chirality can often be avoided [46]. Thus we are eventually left to decide between the two formulations by means of a cost comparison.

In order to perform a cost comparison between twisted mass and overlap formulation we have chosen a set-up consisting of two quenched ensembles of 20 configurations with $L = T = 12$ and $L = T = 16$, respectively. Both were generated with the Wilson gauge action at $\beta = 5.85$ corresponding to a lattice spacing of $a = 0.12$ fm. We have tuned the bare twisted mass parameter $a\mu$ and the overlap bare quark mass am_{ov} such that the values of the pseudo scalar masses are matched. The actual values can be found in table 2.6.

We then invert the twisted mass and the overlap operator separately on two point-like sources η requiring a stopping criterion of $\|Ax - \eta\| < 10^{-14}$. We are working in the chiral basis (see appendix A.1) and have chosen the two sources to correspond to the two different chiral sectors. The inversions are usually performed with iterative solvers. In order to test the performance of different available solvers, we have implemented the minimal residual (MR), the conjugate gradient normal equation (CG(NE)), the conjugate gradient squared (CGS), the stabilized Bi-conjugate gradient (BiCGstab), the generalized minimal residual (GMRES) (see Ref. [102] for all of them) and the shifted minimal residual (SUMR) (cf. Ref. [103])

2.4. OVERLAP VERSUS TWISTED MASS FERMIONS

iterative solvers. The SUMR method is not applicable to the twisted mass operator, because it requires a shifted unitary operator. Moreover, some of the solvers fail to invert the twisted mass operator or converge only if D_{tm} is multiplied with γ_5 .

The computationally most expensive part in the inversion of the overlap operator is the approximation of $(A^\dagger A)^{-1/2}$ in Eq. (1-37), where we use the hermitian Wilson-Dirac operator as kernel $A \equiv Q = \gamma_5 D_W$. As mentioned before the square-root is approximated by means of Chebyshev polynomials, which have an degree of the order 200 – 300 in our particular set-up, if we project out the lowest 20 and 40 eigenvectors of Q on the 12^4 and 16^4 lattices, respectively. Hence, per application of the overlap operator the Wilson-Dirac operator must be applied order 400 – 600 times.

One strategy to reduce the number of Q application during the inversion of D_{ov} is to adapt the accuracy of the Chebyshev approximation. This can speed up the inversion by large factors since a reduction in the order of the polynomial enters multiplicatively in the total cost of the inversion. We denote the usage of adaptive precision by a subscript $_{\text{ap}}$ to the solver name. Depending on the solver the adaptive precision was applied in different ways. In case of the CG_{ap} we cut the polynomial as soon as the contribution to the resulting vector are smaller than the desired residuum by a factor 10^{-2} . This requires the full polynomial only at the beginning of the CG-search while towards the end polynomials of $\mathcal{O}(10)$ are sufficient. In the case of the MR_{ap} or the GMRES_{ap} on the other hand it is possible to start with a $\mathcal{O}(10)$ polynomial right at the beginning. From time to time the introduced error is corrected for by calculating the residuum to full precision. This corresponds to a restart of the solver, which is for these two particular solvers a natural procedure.

In the special case of the $\text{CG}(\text{NM})$ solver an additional factor of two can be saved when the overlap operator is inverted. Since in the $\text{CG}(\text{NM})$ the squared operator $\gamma_5 D_{\text{ov}} \gamma_5 D_{\text{ov}}$ - which is real and positive - is inverted, one can make use of the property of D_{ov} that $P_\pm D_{\text{ov}} P_\pm \propto P_\pm D_{\text{ov}}^\dagger D_{\text{ov}} P_\pm$, where $P_\pm = (1 \pm \gamma_5)/2$ denote the projectors on the positive and negative chiral sector. Therefore, if the sources are chiral half of the applications of D_{ov} can be saved. We denote this algorithm by CG_χ , which can also be combined with adaptive precision $\text{CG}_{\text{ap},\chi}$.

A further way to speed up the inversion of the overlap operator is the so called low mode preconditioning [104], which is supposed to help in the regime of quark masses lighter than what we have used. Therefore, in this chapter we did not include low mode preconditioning in the comparison. While for the overlap operator even/odd preconditioning cannot be applied due to the polynomial approximation of $(Q^\dagger Q)^{-1/2}$, it can be used for the twisted mass operator to reduce the inversion cost. See appendix B.1 for details on how to implement the inversion with even/odd preconditioning.

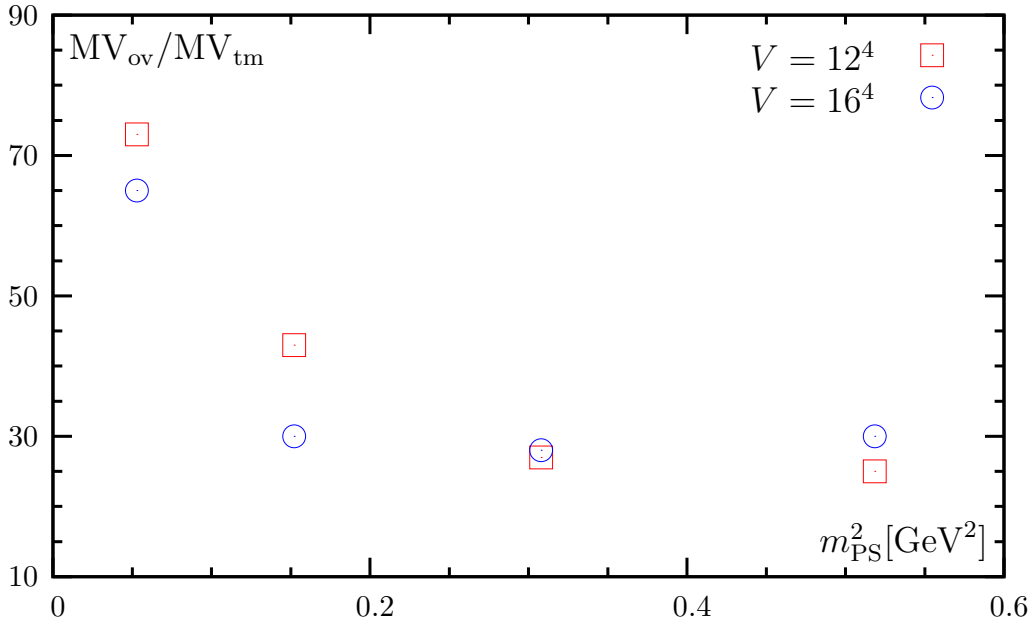


Figure 2.11: $MV_{\text{ov}}/MV_{\text{tm}}$ of the fastest available solver versus $m_{\text{PS}}^2 [\text{GeV}^2]$ for the two volumes $V = 12^4$ and $V = 16^4$.

We are left to define which quantity we use in the comparison of the different solvers and then also the two different operators. One application of D_{tm} is as expensive as one application of its even/odd preconditioned version as well as one application of the kernel of the overlap operator. Hence, it is natural to use the number of those applications for the comparison. Of course, the different iterative solvers have in general a different amount of additional linear algebra operations, which could be included in the comparison by measuring the wall clock time needed for the inversions. However, the latter measure is highly machine dependent and in addition our experience shows that the number of operator applications (which we will denote by matrix vector (MV) multiplications) is a sufficient criterion.

Our results clearly reveal that in case of the twisted mass operator the CG in combination with even/odd preconditioning is the best choice in the whole range of masses and for both the volumes we investigated here. For the overlap operator the GMRES_{ap} is the fastest algorithm of the iterative solvers we considered, apart from the simulation point with a 12^4 lattice volume and $m_{\text{PS}} = 230$ MeV where the $\text{CG}_{\chi,\text{ap}}$ is the fastest.

For the fastest available solvers we plot in figure 2.11 the ratio of overlap kernel applications MV_{ov} and applications the preconditioned twisted mass operator MV_{tm} versus m_{PS}^2 in physical units. Depending on the mass and the volume, the inversion of the twisted mass operator is a factor 20 to 70 faster than the inversion of the overlap operator at matched values of the pseudo scalar mass from 230 MeV to 720 MeV. When the two different volumes are compared the overlap operator performs slightly

better on the larger volume when compared to the twisted mass operator. But, of course, for a definite conclusion we would need at least one additional volume.

2.5 Conclusion

In this chapter we presented a detailed scaling test in the quenched approximation of the twisted mass lattice QCD formulation at maximal twist. To this end we have computed several physical quantities in a range of lattice spacings between 0.17 fm and 0.048 fm and pseudo scalar masses between 270 MeV and 1.2 GeV. In order to work at full twist ($\omega = \pi/2$) it is necessary to determine the critical value of the hopping parameter κ_{crit} . We have explored the pion definition and the PCAC definition for this parameter in order to investigate the influence of these particular choices on the lattice artifacts.

The results of our – quenched – study for the vector meson mass m_V and the pseudo scalar decay constant f_{PS} are very encouraging. Our data strongly suggest that in this setup the lattice spacing effects are substantially reduced with respect to standard Wilson fermions and consistent with vanishing $\mathcal{O}(a)$ discretization errors. This holds for both the pion definition and the PCAC definition of κ_{crit} . Using the PCAC definition of the critical mass the scaling region is found to start already at $(a/r_0)^2 \leq 0.06$ for the observables investigated here, while for the pion definition this region starts only at $(a/r_0)^2 \leq 0.04$. Moreover, the $\mathcal{O}(a^2)$ artifacts remain small for pseudo scalar masses down to 270 MeV when the PCAC definition is used, which is not the case for the pion definition. However, at the two smallest quark masses, we had to include for the pion definition a point at $\beta = 6.45$ in order to safely control the continuum limit extrapolation.

In the case of f_{PS} we have explicitly checked that both definitions of the critical mass lead *independently* to consistent values in the continuum limit. Nevertheless, for further simulations the PCAC definition of κ_c is clearly preferable as it leads to considerably smaller lattice artifacts at small quark masses, allowing at the same time for an enlargement of the scaling region.

We also investigated the flavor breaking effects in the twisted mass formulation. Flavor symmetry is explicitly broken by the twisted mass term at finite values of the lattice spacing. We have shown that the mass splitting between the neutral and the charged pseudo scalar state is not small. But, as expected, the splitting vanishes like a^2 in the continuum limit.

In addition we confronted at one value of the bare coupling constant $\beta = 5.85$ results obtained with the overlap formulation with corresponding results obtained with the twisted mass formulation of lattice QCD. We find that for the quantities investigated in this chapter the two formulations are compatible, in particular

pseudo scalar masses smaller than 300 MeV can be reached with both formulations. However, with a detailed test of various iterative solvers for both formulations we could show that (quenched) simulations with the twisted mass formulation are a factor of 20 to 70 faster than with the overlap formulation, depending on the mass under consideration.

Therefore, the results of this chapter clearly reveal that mtmQCD allows for reliable simulations at pseudo scalar meson masses of about 270 MeV without running into problems with exceptionally small eigenvalues. In addition lattice artifacts linear in a are absent and, when the PCAC definition of the critical mass is used, also the residual lattice artifacts are small in the whole range of masses investigated here. At the same time the costs are significantly less than what is needed for the overlap formulation. In view of future dynamical simulations with light quark masses this is, we think, a very important lesson.

Chapter 3

Accelerating the Hybrid Monte Carlo algorithm

In chapter 2 we have presented a scaling test of mtmQCD in the quenched approximation with values of the pseudo scalar mass lower than 300 MeV. Of course, it would be desirable to repeat such a study with two dynamical flavors of quarks and light values of their masses, at least for a first study. But, unfortunately, simulations with light dynamical quark flavors are by orders of magnitude more expensive in terms of computer time than the corresponding simulations in the quenched approximation. The reason for this is basically that the generation of one gauge field configuration is much more ($\mathcal{O}(100)$) expensive and the gauge configurations depend in full QCD on the values of the quark masses.

The Hybrid Monte Carlo (HMC) algorithm [75] is one widely used algorithm to perform dynamical simulations. It is an exact algorithm, which combines molecular dynamics evolution of the gauge fields with a Metropolis accept/reject step to correct for discretization errors in the numerical integration of the corresponding equations of motion. However, in its original form the HMC algorithm is even on computers available today not able to tackle simulations with light quarks on fine lattices. Due to increasing iteration numbers in the solvers and autocorrelation times the costs C are expected to increase as [105]

$$C = K \left(\frac{m_{\text{PS}}}{m_{\text{V}}} \right)^{-6} L^5 a^{-7},$$

where m_{PS} and m_{V} are the pseudo scalar and the vector mass, L is the spatial lattice extent and a the lattice spacing. The proportionality factor K was found to be too large to allow for simulations with realistic mass values on fine lattices [105]. Hence, during the last years a lot of effort has been invested to decrease K and to improve the cost scaling behavior of the HMC. The list of improvements that were found reaches, for instance, from even/odd preconditioning [106] over multiple time scale

integration [107] to mass preconditioning (Hasenbusch acceleration) [108, 109], to mention only those that are immediately relevant for the present work. It is worth noting that many of the known improvement tricks can be combined. In addition, alternative multi-boson methods [74] have been suggested, which, however, appear not to be superior to the HMC algorithm, although they have conceptual advantages compared to the HMC algorithm.

Recently in Ref. [110] a HMC variant as a combination of multiple time scale integration with domain decomposition as preconditioner with excellent scaling properties with the quark mass was presented. In addition, the value of K seems to be significantly lower than for other HMC variants. Thus this algorithm can be expected to be most promising when one wants to simulate small quark masses on fine lattices.

In this chapter we are going to present yet another variant of the HMC algorithm similar to the one of Refs. [111, 112] comprising multiple time scale integration with mass preconditioning. We test this algorithm for standard Wilson fermions at $\beta = 5.6$ and at pseudo scalar masses ranging from $m_{\text{PS}} = 380$ MeV to $m_{\text{PS}} = 670$ MeV, which are the simulation points of Ref. [110]. We show that in this situation the algorithm has similar scaling properties and performance as the method presented in Ref. [110]. From the performance data obtained with our HMC variant we find that K is reduced and the scaling of the cost with the quark mass is improved when compared to performance data available in the literature [105, 113, 114].

3.1 HMC algorithm

The variant of the HMC algorithm we will present here is applicable to a wide class of lattice Dirac operators, including twisted mass fermions, various improved versions, staggered fermions, and even the overlap operator. Nevertheless, in order to discuss a concrete example, we restrict ourselves in this chapter to the Wilson-Dirac operator (1-26) for $n_f = 2$ flavors of mass degenerate quarks with Wilson parameter r set to one. We do not expect that the algorithm properties depend significantly on the particular choice of the Dirac operator.

Since D_W (cf. Eq. (1-26)) fulfills the property $\gamma_5 D_W \gamma_5 = D_W^\dagger$ it is convenient to define the hermitian Wilson-Dirac operator

$$Q = \gamma_5 D_W. \quad (3-1)$$

After integrating out the fermion fields (cf. Eq. (1-15)) we have to deal with $\det(D_W)^2 = \det(Q^2)$ in the simulations. This is usually done by re-expressing the determinant with a Gaussian integral over bosonic fields ϕ, ϕ^\dagger :

$$\det(D_W)^2 = \det(Q)^2 \propto \int \mathcal{D}\phi^\dagger \mathcal{D}\phi \exp\left(-S_{\text{PF}}[U, \phi^\dagger, \phi]\right), \quad (3-2)$$

where $S_{\text{PF}}[U, \phi^\dagger, \phi] = |Q^{-1}\phi|^2$ is the so called pseudo fermion action. Formally the ϕ -fields are identical to the matter fields ψ since they have the same degrees of freedom, but follow the statistic of bosonic fields and are therefore called *pseudo fermion* fields. The partition function (1-21) is then given by

$$\mathcal{Z} = \int \mathcal{D}U \mathcal{D}\phi \mathcal{D}\phi^\dagger e^{-S_G[U] - S_{\text{PF}}[U, \phi, \phi^\dagger]} . \quad (3-3)$$

In order to perform Monte Carlo simulations for this partition function the integral over the pseudo fermion fields could be included in the important sampling process. This is, however, not needed, since for a given gauge field

$$e^{-S_{\text{PF}}[U, \phi, \phi^\dagger]} = e^{-R^\dagger R} \quad (3-4)$$

is a Gaussian distribution, which can be generated exactly. Therefore only the gauge fields have to be generated in a Markov chain, which is implemented by the Hybrid Monte Carlo algorithm. To set up the HMC algorithm [75] we introduce traceless Hermitian momenta $P_{x,\mu}$ as conjugate fields to the gauge fields $U_{x,\mu}$ and a Hamiltonian

$$H(P, U, \phi, \phi^\dagger) = \frac{1}{2} \sum_{x,\mu} P_{x,\mu}^2 + S_G[U] + S_{\text{PF}}[U, \phi, \phi^\dagger] . \quad (3-5)$$

The algorithm is then composed out of the following steps:

1. Global heat-bath for momenta and pseudo fermion fields.

The initial momenta are randomly produced according to a Gaussian distribution $\exp(-P^2/2)$. Moreover, random fields R are produced from a distribution like $\exp(-R^\dagger R)$ and the initial pseudo fermion fields are computed with $\phi = QR$.

2. Molecular dynamics evolution.

Production of a proposal gauge configuration U' and proposal momenta P' by *molecular dynamics* evolution (integrating Hamilton's equations of motion) of the gauge fields U and the momenta P at fixed pseudo fermion fields ϕ . If the integration of the corresponding equations of motion can be performed exactly, the Hamiltonian is conserved under this evolution.

3. Metropolis accept/reject step.

The proposals U' and P' are accepted with the probability $\min\{1, \exp(-\Delta H)\}$, where $\Delta H = H(P', U', \phi, \phi^\dagger) - H(P, U, \phi, \phi^\dagger)$.

This step is needed because the integration of the equations of motion can in practice be done only numerically and hence an acceptance step is needed to correct for discretization errors.

If the integration scheme is reversible and area preserving it is possible to prove that the HMC algorithm satisfies the *detailed balance* condition [75] and hence the HMC algorithm is exact.

Since the Hamiltonian is conserved up to discretization errors, the integration can be set up such that the acceptance rate is high while the gauge configurations are globally updated.

3.1.1 Molecular dynamics evolution

In the molecular dynamics part of the HMC algorithm the gauge fields U and the momenta P need to be evolved in a fictitious computer time t . With respect to t , Hamilton's equations of motion read

$$\frac{dU}{dt} = \frac{dH}{dP} = P, \quad \frac{dP}{dt} = -\frac{dH}{dU} = -\frac{dS}{dU}, \quad (3-6)$$

where we set $S = S_G + S_{PF}$ and d/dU , d/dP formally denote the derivative with respect to group elements. Since analytical integration of the former equations of motion is normally not possible, these equations must in general be integrated with a discretized integration scheme that is area preserving and reversible, such as the leap frog algorithm. The discrete update with integration step size $\Delta\tau$ of the gauge field and the momenta can be defined as

$$\begin{aligned} T_U(\Delta\tau) : \quad U &\rightarrow U' = \exp(i\Delta\tau P) U, \\ T_S(\Delta\tau) : \quad P &\rightarrow P' = P - i\Delta\tau \delta S, \end{aligned} \quad (3-7)$$

where δS is an element of the Lie algebra of $SU(3)$ and denotes the variation of S with respect to the gauge fields. The computation of δS is the most expensive part in the HMC algorithm since the variation of S_{PF} reads

$$\delta S_{PF} = -\phi^\dagger \frac{1}{Q^2} \delta(Q^2) \frac{1}{Q^2} \phi \quad (3-8)$$

and thus involves the inversion of the Wilson-Dirac operator. With (3-7) one basic time evolution step of the so called leap frog algorithm reads

$$T = T_S(\Delta\tau/2) T_U(\Delta\tau) T_S(\Delta\tau/2), \quad (3-9)$$

and a whole trajectory of length τ is achieved by $N_{MD} = \tau/\Delta\tau$ successive applications of the transformation T .

3.1.2 Integration with multiple time scales

Consider a Hamiltonian of the form

$$H = \frac{1}{2} \sum_{x,\mu} P_{x,\mu}^2 + \sum_{i=0}^k S_i[U], \quad (3-10)$$

with $k \geq 1$. For instance, with $k = 1$ S_0 might be identified with the gauge action and S_1 with the pseudo fermion action of Eq. (3-5).

Clearly, in order to keep the discretization errors in a leap frog like algorithm small, the time steps have to be small if the driving forces are large. Then, if the forces originating from the single parts in the Hamiltonian (3-10) differ significantly in their absolute values, it might be valuable to integrate the different S_i on time scales inverse proportionally deduced from the corresponding forces. This will maximally improve the algorithm performance if the most expensive part can be integrated with the largest molecular dynamics steps size.

The leap frog integration scheme can be generalized to multiple time scales as has been proposed in Ref. [107] without loss of reversibility and the area preserving property. The scheme with only one time scale can be recursively extended by starting with the definition

$$T_0 = T_{S_0}(\Delta\tau_0/2) T_U(\Delta\tau_0) T_{S_0}(\Delta\tau_0/2), \quad (3-11)$$

with T_U defined as in Eq. (3-7) and where $T_{S_i}(\Delta\tau)$ is given by

$$T_{S_i}(\Delta\tau) \quad : \quad P \quad \rightarrow \quad P - i\Delta\tau\delta S_i[U]. \quad (3-12)$$

As $\Delta\tau_0$ will be the smallest time scale, we can recursively define the basic update steps T_i , with time scales $\Delta\tau_i$ as

$$T_i = T_{S_i}(\Delta\tau_i/2) [T_{i-1}]^{N_{i-1}} T_{S_i}(\Delta\tau_i/2), \quad (3-13)$$

with integers N_i and $0 < i \leq k$. One full trajectory τ is then composed by $[T_k]^{N_k}$. The different time scales $\Delta\tau_i$ in Eq. (3-13) must be chosen such that the total number of steps on the i -th time scale N_{MD_i} times $\Delta\tau_i$ is equal to the trajectory length τ for all $0 \leq i \leq k$: $N_{MD_i}\Delta\tau_i = \tau$. This is obviously achieved by setting

$$\Delta\tau_i = \frac{\tau}{N_k \cdot N_{k-1} \cdot \dots \cdot N_i} = \frac{\tau}{N_{MD_i}}, \quad 0 \leq i \leq k, \quad (3-14)$$

where $N_{MD_i} = N_k \cdot N_{k-1} \cdot \dots \cdot N_i$.

In Ref. [107] also a partially improved integration scheme with multiple time scales was introduced, which reduces the size of the discretization errors. Again, we assume a Hamiltonian of the form (3-10) with now $k = 1$. By defining similar to T_0

$$T_{SW_0} = T_{S_0}(\Delta\tau_0/6) T_U(\Delta\tau_0/2) T_{S_0}(2\Delta\tau_0/3) T_U(\Delta\tau_0/2) T_{S_0}(\Delta\tau_0/6), \quad (3-15)$$

the basic update step of the improved scheme – usually referred to as the Sexton-Weingarten (SW) integration scheme – reads

$$\begin{aligned} T_{SW_1} = & T_{S_1}(\Delta\tau_1/6) \\ & [T_{SW_0}]^{N_0} T_{S_1}(2\Delta\tau_1/3) \\ & [T_{SW_0}]^{N_0} T_{S_1}(\Delta\tau_1/6), \end{aligned} \quad (3-16)$$

where $\Delta\tau_0 = \Delta\tau_1/(2N_0)$. This integration scheme not only reduces the size of the discretization errors, but also sets for S_0 a different time scale than for S_1 . Hence, it is one special example for an integration scheme with multiple time scales and can easily be extended to more than two time scales by recursively defining ($0 < i \leq k$):

$$\begin{aligned} T_{\text{SW}_i} &= T_{S_i}(\Delta\tau_i/6) \\ & [T_{\text{SW}_{i-1}}]^{N_{i-1}} T_{S_i}(2\Delta\tau_i/3) \\ & [T_{\text{SW}_{i-1}}]^{N_{i-1}} T_{S_i}(\Delta\tau_i/6). \end{aligned} \quad (3-17)$$

The different time scales for the SW integration scheme are defined by

$$\Delta\tau_i = \frac{\tau}{(2N_k) \cdot (2N_{k-1}) \cdot \dots \cdot (2N_i)} = \frac{\tau}{N_{\text{MD}_i}}, \quad i \leq k. \quad (3-18)$$

Note that the SW partially improved integration scheme was originally invented to make use of the fact that the computation of the variation of the gauge action is cheap as compared to the variation of the pseudo fermion action and in addition the time scales are chosen in order to cancel certain terms in the discretization error exactly [107]. It can be generalized to integrators of the form

$$T_{2\text{MN}_0} = T_{S_0}(\lambda\Delta\tau_0) T_U(\Delta\tau_0/2) T_{S_0}((1-2\lambda)\Delta\tau_0) T_U(\Delta\tau_0/2) T_{S_0}(\lambda\Delta\tau_0), \quad (3-19)$$

with a real parameter λ that needs to be tuned [115], which are called second order Minimal Norm (2MN) integration schemes. These schemes can be generalized to multiple time scales in exactly the same way as the SW integration scheme. However, we restrict ourselves in this work to the LF and the SW integration scheme only, also because from the results of Ref. [115] we do not expect a large improvement of 2MN schemes when compared to the SW scheme.

3.2 Mass Preconditioning

The arguments presented in this section are made for simplicity only for the not even/odd preconditioned Wilson-Dirac operator. The generalization to the even/odd preconditioned case is simple and can be found in Ref. [108] and the appendix B.1.

Mass preconditioning [108] – also known as Hasenbusch acceleration – relies on the observation that one can rewrite the fermion determinant as follows

$$\begin{aligned} \det(Q^2) &= \det(W^+W^-) \frac{\det(Q^2)}{\det(W^+W^-)} \\ &= \int \mathcal{D}\phi_1^\dagger \mathcal{D}\phi_1 \mathcal{D}\phi_2^\dagger \mathcal{D}\phi_2 e^{-\phi_1^\dagger \frac{1}{W^+W^-} \phi_1 - \phi_2^\dagger W^+ \frac{1}{Q^2} W^- \phi_2} \\ &= \int \mathcal{D}\phi_1^\dagger \mathcal{D}\phi_1 \mathcal{D}\phi_2^\dagger \mathcal{D}\phi_2 e^{-S_{\text{PF}_1} - S_{\text{PF}_2}}. \end{aligned} \quad (3-20)$$

The preconditioning operators W^\pm can in principle be freely chosen, but in order to let the preconditioning work W^+W^- should be a reasonable approximation of Q^2 , which is, however, cheaper to simulate. Moreover, to allow for Monte Carlo simulations, $\det(W^+W^-)$ must be positive. The generalized Hamiltonian (3-5) corresponding to Eq. (3-20) reads

$$H = \frac{1}{2} \sum_{x,\mu} P_{x,\mu}^2 + S_G[U] + S_{\text{PF}_1}[U, \phi_1, \phi_1^\dagger] + S_{\text{PF}_2}[U, \phi_2, \phi_2^\dagger], \quad (3-21)$$

and it can of course be extended to more than one additional field.

Note that a similar approach was presented in Ref. [116], in which the introduction of n pseudo fermion fields was coupled with the n -th root of the fermionic kernel.

One particular choice for W^\pm is to take for W^+ and W^- the one flavor components of a two flavor twisted mass operator

$$W^\pm = Q \pm i\mu, \quad (3-22)$$

which can be written as a two flavor operator as known from Eq. (1-47)

$$\begin{pmatrix} W^+ & \\ & W^- \end{pmatrix} = \gamma_5 \left[\begin{pmatrix} D_W & \\ & D_W \end{pmatrix} + i\mu\gamma_5\tau_3 \right] \quad (3-23)$$

One important property of this choice is that $W^+W^- = Q^2 + \mu^2$ and that $(W^+)^\dagger = W^-$. For small values of μ the product W^+W^- is certainly a reasonable approximation for Q^2 , but due to the mass shift μ^2 it is cheaper to invert. We remark that in general also Q itself can be a twisted mass operator.

In Ref. [109, 117] it was argued that the optimal choice for μ is given by $\mu^2 = \sqrt{\lambda_{\max}\lambda_{\min}}$. Here λ_{\max} (λ_{\min}) is the maximal (minimal) eigenvalue of Q^2 . The reason for the above quoted choice is as follows: the condition number of $Q^2 + \mu^2$ is approximately λ_{\max}/μ^2 and the one of $Q^2/(Q^2 + \mu^2)$ approximately μ^2/λ_{\min} . With $\mu^2 = \sqrt{\lambda_{\max}\lambda_{\min}}$ these two condition numbers are equal to $\sqrt{\lambda_{\max}/\lambda_{\min}}$, both of them being much smaller than the condition number of Q^2 which is $\lambda_{\max}/\lambda_{\min}$.

Since the force contribution in the molecular dynamics evolution is supposed to be proportional to some power of the condition number, the force contribution from the pseudo fermion part in the action is reduced and therefore the step size $\Delta\tau$ can be increased, in practice by about a factor of 2 [108, 109]. Therefore Q^2 must be inverted only about half as often as before and if the inversion of W^+W^- , which is needed to compute δS_{PF_1} , is cheap compared to the one of Q^2 the simulation speeds up by about a factor of two [108, 109].

One might wonder why the reduction of the condition number from K to \sqrt{K} gives rise to only a speedup factor of about 2. One reason for this is that one cannot

make use of the reduced condition number of $Q^2/(Q^2 + \mu^2)$ in the inversion of this operator, because in the actual simulation still the badly conditioned operator Q^2 must be inverted to compute the variation of $S_{\text{PF}_2} = \phi_2^\dagger \frac{W^+W^-}{Q^2} \phi_2$.

3.2.1 Mass preconditioning and multiple time scale integration

In the last subsection we have seen that mass preconditioning is indeed an effective tool to change the condition numbers of the single operators appearing in the factorization (3-20) compared to the original operator. But, this reduction of the condition numbers only influences the forces and *not* the number of iterations to invert the physical operator Q^2 .

Therefore, it might be advantageous to change the point of view: instead of tuning the condition numbers in a way à la Refs. [108, 109] we will exploit the possibility of arranging the forces by the help of mass preconditioning with the aim to arrange for a situation in which a multiple time scale integration scheme is favorable, as explained at the beginning of section 3.1.2.

The procedure can be summarized as follows: use mass preconditioning to split the Hamiltonian in different parts. The forces of the single parts should be adjusted by tuning the preconditioning mass parameter μ such that the more expensive the computation of δS_{PF_i} is, the less it contributes to the total force. This is possible because the variation of $(Q^2 + \mu^2)/Q^2$ is, for $|\mu| < 1$, (formally) reduced by a factor μ^2 compared to the variation of $1/Q^2$. In addition, $W^+W^- = Q^2 + \mu^2$ is significantly cheaper to invert than Q^2 . Then integrate the different parts on time scales chosen according to the magnitude of their force contribution.

The idea presented in this chapter is very similar to the idea of separating infrared and ultraviolet modes as proposed in Ref. [118]. This idea was applied to mass preconditioning by using only two time scales in Refs. [111, 112] in the context of clover improved Wilson fermions. However, a comparison of our results presented in the next section to the ones of Refs. [111, 112] is not possible, because volume, lattice spacing and masses are different.

3.3 Numerical results

3.3.1 Simulation points

In order to test the HMC variant introduced in the last sections, we decided to compare it with the algorithm proposed and tested in Ref. [110]. To this end we performed simulations with the same parameters as have been used in Ref. [110]:

	κ	m_q [MeV]	m_{PS} [MeV]	m_V [MeV]	r_0/a
<i>A</i>	0.1575	66(3)	665(17)	947(20)	6.04(10)
<i>B</i>	0.1580	34(1)	485(13)	836(24)	6.18(07)
<i>C</i>	0.15825	22(1)	380(17)	839(33)	6.40(15)

Table 3.1: The (un-renormalized) quark mass m_q , the pseudo scalar mass m_{PS} and the vector mass m_V are given in physical units at the three simulation points *A*, *B* and *C*. We use Wilson fermions at $\beta = 5.6$ on $24^3 \times 32$ lattices. The scale was set by the use of $r_0 = 0.5 \text{ fm}$ and we give the value of r_0/a at each simulation point. The values of all the quantities agree within the errors with the numbers quoted in Refs. [110, 68, 114], apart from the value for r_0/a at simulation point *B*, which disagrees by two sigmas to the value quoted in Ref. [68]. This is presumably due to the different methods to measure this quantity. For the measurements we used at each simulation point 100 thermalized configurations separated by 5 trajectories.

Wilson-Dirac operator with plaquette gauge action at $\beta = 5.6$ on $24^3 \times 32$ lattices. We have three simulation points *A*, *B* and *C* with values of the hopping parameter $\kappa = 0.1575$, $\kappa = 0.1580$ and $\kappa = 0.15825$, respectively. The trajectory length was set to $\tau = 0.5$. The details of the physical parameters corresponding to the different simulation points can be found in table 3.1. Additionally, this choice of simulation points allows at the two parameter sets *A* and *B* a comparison to results published in Ref. [114], where a HMC algorithm with a plain leap frog integration scheme was used.

In addition to the three simulation points *A*, *B* and *C* we have one additional point *D* with $\kappa = 0.15835$. According to Ref. [119] this value of κ corresponds to a pseudo scalar mass of about 294 MeV. Unfortunately, the history of our run is too short to be really conclusive, nevertheless we will use run *D* to get a preliminary idea of the performance of our algorithm towards even smaller quark and pseudo scalar masses.

3.3.2 Details of the implementation

We have implemented a HMC algorithm for two flavors of mass degenerate quarks with even/odd preconditioning and mass preconditioning with up to three pseudo fermion fields (cf. appendix B.1 on page 109). The boundary conditions are periodic in all directions apart from anti-periodic ones for the fermion fields in time direction. For the gauge action the usual Wilson plaquette gauge action (1-25) is used. The implementation is written in C and uses double precision throughout.

For the mass preconditioning we use

$$W_j^\pm = \gamma_5(D_W[U, m_0] \pm i\mu_j\gamma_5), \quad (3-24)$$

with $j = 1, 2$ for the factorization in Eq. (3-20), where the μ_j are the additional (unphysical) twisted mass parameters. Therefore, the pseudo fermion actions S_{PF_j} are given by

$$S_{\text{PF}_j}[U] = \begin{cases} \phi_1^\dagger \left(\frac{1}{W_1^+ W_1^-} \right) \phi_1 & j = 1, \\ \phi_j^\dagger \left(\frac{W_{j-1}^+ W_{j-1}^-}{Q^2} \right) \phi_j & j = N_{\text{PF}}, \\ \phi_j^\dagger \left(\frac{W_{j-1}^+ W_{j-1}^-}{W_j^+ W_j^-} \right) \phi_j & \text{otherwise,} \end{cases} \quad (3-25)$$

where we always chose $\mu_j > 0$ and $\mu_{j+1} > \mu_j$ for all values of j . N_{PF} denotes the actually used number of pseudo fermion fields.

We have implemented the leap frog (LF) and the Sexton-Weingarten (SW) integration schemes with multiple time scales each as described by Eq. (3-13) and Eq. (3-17), respectively, where k in both equations has to be identified with N_{PF} .

The time scales are defined as in Eq. (3-14) for the LF integration scheme and as in Eq. (3-18) for the SW scheme, with N_0 corresponding to the gauge action and N_j to S_{PF_j} for $N_{\text{PF}} \geq j > 0$. Note that for the LF integration scheme for one trajectory there are $N_{N_{\text{PF}}} \cdot \dots \cdot N_j + 1$ inversions of the corresponding operator needed, while for the SW integration scheme there are $2N_{N_{\text{PF}}} \cdot \dots \cdot 2N_j + 1$ inversions needed.

For the inversions we used the CG and the BiCGstab iterative solvers. As reported in section 2.4.1 the CG iterative solver is best suited for the even/odd preconditioned twisted mass operator. Thus we used for all inversions of mass preconditioning operators exclusively the CG iterative solver.

For the pure Wilson-Dirac operator D_W the BiCGstab iterative solver is known to perform best [120]. In case of dynamical simulations, however, usually the squared hermitian operator needs to be inverted and in this case the CG is comparable to the BiCGstab. Only in the acceptance step, where $\gamma_5 D_W$ (or rather the even/odd preconditioned version of it) needs to be inverted to a high precision, the usage of the CG would be wasteful. For this work we used the BiCGstab iterative solver for all inversions of either the pure Wilson-Dirac operator itself or $(\gamma_5 D_W)^2$.

The accuracy in the inversions was set during the computation of δS_{PF_j} to ϵ_j , which means that the inversions were stopped when the approximate solution ψ_j of $A_j \psi_j = \phi_j$ fulfills

$$\frac{\|\phi_j - A_j \psi_j\|}{\|\phi_j\|} \leq \epsilon_j,$$

where A_j denotes the operator corresponding to S_{PF_j} . During the inversions needed for the acceptance step the accuracy was set to $\tilde{\epsilon} = 10^{-10}$ for all pseudo fermion actions. The inversions in the acceptance step must be rather precise in order not to introduce systematic errors in the simulation, while for the force computation the precision can be relaxed as long as the reversibility violations are not too large. The

3.3. NUMERICAL RESULTS

	Int.	N_{PF}	N_{therm}	$N_{\{0,1,2,3\}}$	$\epsilon_1, \epsilon_2, \epsilon_3$	μ_1, μ_2	P_{acc}
<i>A</i>	SW	3	600	3, 2, 1, 3	$10^{-7}, 10^{-8}, 10^{-8}$	0.29, 0.057	0.86
<i>B</i>	SW	3	1000	3, 2, 1, 3	$10^{-8}, 10^{-8}, 10^{-8}$	0.25, 0.057	0.81
<i>C</i>	LF	2	1500	5, 6, 10, -	$10^{-8}, 10^{-8}, -$	0.054, -	0.80

Table 3.2: HMC algorithm parameters for the three simulation points. We give the integration scheme, the number of pseudo fermion fields N_{PF} , the number N_{therm} of trajectories of length 0.5 used to thermalize the systems, the number N_i of molecular dynamics steps for the multiple time scale integration scheme, the residues ϵ_i used in the solver for the force computation, the preconditioning mass parameter μ_i and the acceptance rate. We remind that N_0 corresponds to the gauge action.

values of ϵ_j and $\tilde{\epsilon}$ have been set such that the reversibility violations, which should be under control [121, 122, 123, 124], are on the same level as reported in Ref. [110], which means that the differences in the Hamiltonian are of the order* of 10^{-5} . The values for ϵ_j can be found in table 3.2.

The errors and autocorrelation times were computed with the so called Γ -method as explained in section 1.4.3, Eq. (1-103) on page 33 and in Ref. [78] (see also Ref. [77]).

3.3.3 Force contributions

The force contributions to the total force from the separate parts in the action we label by F_G for the gauge action and by F_j for the pseudo fermion action S_{PF_j} . Since the variation of the action with respect to the gauge fields is an element of the Lie algebra of $\text{SU}(3)$, we used $\|X\|^2 = -2 \text{Tr} X^2$ as the definition of the norm of such an element.

In order to better understand the influence of mass preconditioning on the HMC algorithm we computed the average and the maximal norm of the forces F_G, F_1, F_2 and F_3 on a given gauge field after all corresponding gauge field updates:

$$\begin{aligned} \|F\|_{\text{aver}} &= \frac{1}{4L^3T} \sum_{x,\mu} \|F(x, \mu)\|, \\ \|F\|_{\text{max}} &= \max_{x,\mu} \{\|F(x, \mu)\|\}, \end{aligned} \tag{3-26}$$

and averaged them over all measurements, which we indicate with $\langle \cdot \rangle$. Examples of force distributions for different runs can be found in figure 3.1. These investigations lead to the following observations generic to our simulation points:

*In case of 80% acceptance rate the average value of $\sqrt{\Delta H^2}$ is about 0.1. Therefore, a reversibility violation of the order 10^{-5} is supposed to be safe.

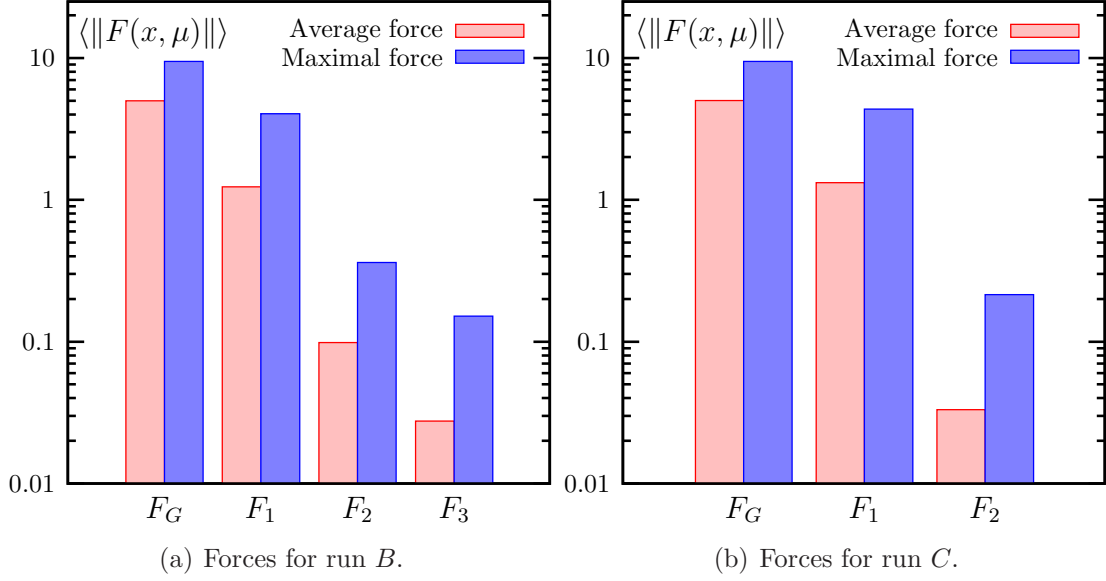


Figure 3.1: Average and maximal forces for simulation points *B* and *C*. The statistical errors are too small to be visible due to the large number of measurements.

- With the choice of parameters as given in table 3.2 the single force contributions are strictly hierarchically ordered with

$$\|F_G\|_{\text{aver,max}} > \|F_1\|_{\text{aver,max}} > \|F_2\|_{\text{aver,max}} > \|F_3\|_{\text{aver,max}}.$$
- The maximal force is up to one order of magnitude larger than the average force. This can only be explained by large local fluctuations in this quantity. These fluctuations become larger the smaller the mass is.

Moreover, the force ordering and sizes look very similar to the one reported in Ref. [110].

In a next step we performed some test trajectories without mass preconditioning in order to compare the fermionic forces with and without mass preconditioning. For the value of $\kappa = 0.15825$ (run *C*) the result can be found in figure 3.2. The bars labeled with *F* correspond to the fermion force without mass preconditioning. The labels *F*₁ and *F*₂ refer to the two fermionic forces for the run *C* with mass preconditioning. The following ratios are of interest:

$$\begin{aligned} \frac{\|F\|_{\text{aver}}}{\|F_1\|_{\text{aver}}} &\approx 1, & \frac{\|F\|_{\text{aver}}}{\|F_2\|_{\text{aver}}} &\approx 42, \\ \frac{\|F\|_{\text{max}}}{\|F_1\|_{\text{max}}} &\approx 1.3, & \frac{\|F\|_{\text{max}}}{\|F_2\|_{\text{max}}} &\approx 29. \end{aligned}$$

These ratios show that the average and maximal norm of *F*₂ is strongly reduced compared to the average and maximal norm of *F*. We observe that the maximal

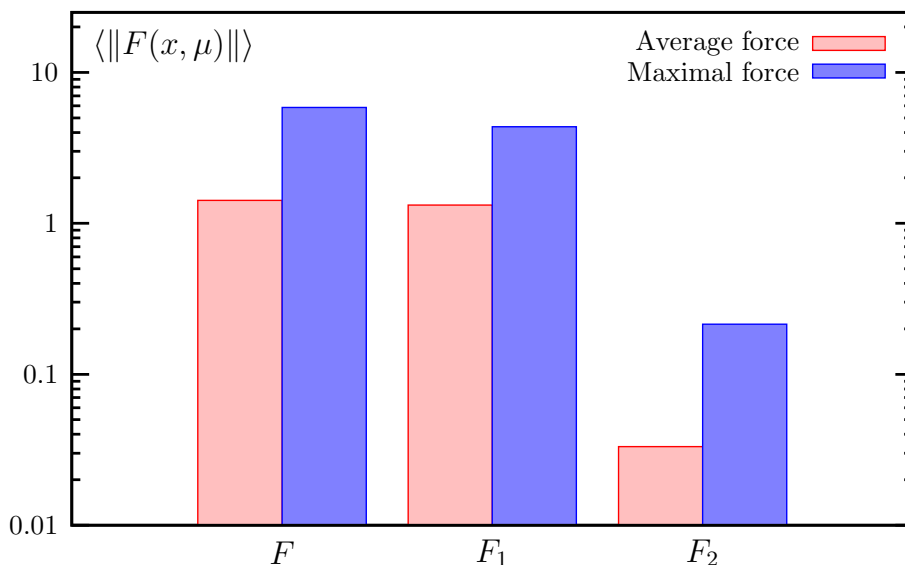


Figure 3.2: Comparison between the fermionic forces of run C (F_1 and F_2) and a run with $\kappa = 0.15825$ without mass preconditioning and multiple time scales (F). The statistical errors are too small to be visible.

norm is slightly less reduced than the average norm and, by varying μ_1 , we could confirm that the norm (average and maximal) of F_2 is roughly proportional to μ_1^2 .

As a further observation, one sees from figure 3.2 or from the ratios quoted above that the norm of F_1 is almost identical to the norm of F , which is the case for both the average and the maximal values.

From these investigations we think one can conclude the following: in the first place it is possible to tune the value of μ_1 (and possibly μ_2) such that the most expensive force contribution of F_2 (or F_3) to the total force becomes small. Secondly, since in the example above the force contributions for F and F_1 are almost identical – even though the masses are very different – we conclude that the norm of the forces does not explain the whole dynamics of the HMC algorithm. For this point see also the discussion in the next subsection.

3.3.4 Tuning the algorithm

As mentioned already in section 3.2.1 the tuning of the different mass parameters and time scales could become a delicate task. Therefore we decided to tune the parameters μ_1 and possibly μ_2 such that the molecular dynamics steps number $N_{N_{\text{PF}}}$ for the LF or $2N_{N_{\text{PF}}}$ for the SW integration scheme – the number of inversions of the original Wilson-Dirac operator in the course of one trajectory – is about the same as the corresponding values in Ref. [110]. The values we have chosen for the mass parameters μ_i and the step numbers N_i can be found in table 3.2 and one can

CHAPTER 3. ACCELERATING THE HMC ALGORITHM

	κ	N_{meas}	$\langle P \rangle$	$\tau_{\text{int}}(P)$
<i>A</i>	0.1575	740	0.57250(3)	6(2)
<i>B</i>	0.1580	1020	0.57339(3)	7(2)
<i>C</i>	0.15825	905	0.57384(4)	10(4)

Table 3.3: For the three runs this table contains the number of measurements for the plaquette N_{meas} , the mean plaquette expectation values and the corresponding autocorrelation times.

see by comparing to Ref. [110] that the step numbers N_i (or $2N_i$) are indeed quite similar.

The computation of the variation of S_G is, compared to the variations of the other action parts, almost negligible in terms of computer time. Therefore we set N_0 always large enough to ensure that the gauge part does not influence the acceptance rate negatively and we leave the gauge part out in the following discussion.

If one compares e.g. for simulation point *C* the average norm of the fermionic forces, then one finds that it increases like 1 : 40 ($\|F_2\| : \|F_1\|$). The maximal norm of the forces is accordingly strongly ordered, approximately like 1 : 20. The corresponding relations in the step numbers we had to choose (see the values in table 3.2) increase only like 1 : 6.

This indicates that the norm of the forces can indeed serve as a first criterion to tune the time scales and the preconditioning masses, by looking for a situation in which $\Delta\tau_i \|F_i\|_{\text{max}}$ is a constant independent of i . But, it cannot be the only criterion. Finally, the acceptance rate is determined by $\langle \exp(-\Delta H) \rangle$, which depends in a more complicated way on the forces, see e.g. Ref. [125].

It is well known that simulations with the HMC algorithm in particular for small quark masses become often unstable if the step sizes are too large. It is an important result that with the choice of parameters as can be found in table 3.2 our simulations appear to be very stable down to quark masses of the order of 20 MeV. We did encounter only few large, but not exceptional, fluctuations in ΔH during the runs. A typical history of ΔH and the average plaquette value can be found in figure 3.3 for run *C*. Note that even a pion mass of about 380 MeV might be still too large to observe the asymptotic behavior of the algorithm.

All our runs reproduce the average plaquette expectation values quoted in Ref. [110] and, where available, in Ref. [114] within the statistical errors. Our results together with the number of measurements N_{meas} and the integrated autocorrelation time can be found in table 3.3. We also measured the values of the pseudo scalar, the vector and the current quark mass and our numbers agree within errors with the

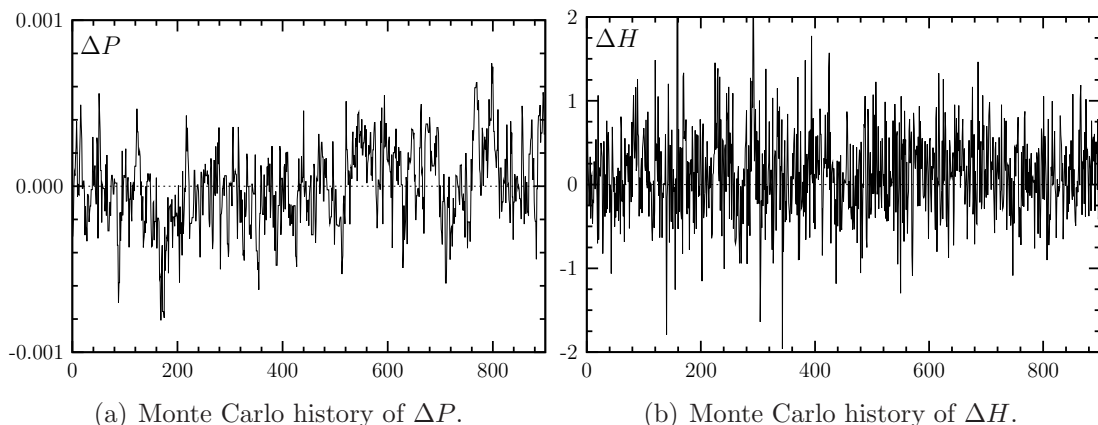


Figure 3.3: Monte Carlo histories of the deviation ΔP of the average plaquette from its mean value and of ΔH , both for simulation point C .

values quoted in Refs. [110, 114]. These measurements were done on 100 configurations separated by 5 trajectories at each simulation point and we computed the aforementioned quantities with the methods explained in section 1.3. In order to improve the signal we used Jacobi smearing and random sources. Our results in physical units can be found in table 3.1. Note that the value for m_V at simulation point C has to be taken with some caution, because the lattice time extend was a bit too small to be totally sure about the plateau.

In order to set the scale we determined the Sommer parameter r_0/a [59] as defined and explained in section 1.3.4 on page 28. For our calculation of r_0/a in this chapter we used the HYP static action[†] [64, 66], the tree-level improved force and potential [59] and we enhanced the overlap with the ground state of the potential using APE smeared [60] spatial gauge links. The results can be found in table 3.1. For run A and B our values for r_0/a agree very well within the errors with the value quoted in Ref. [68, 69]. One should keep in mind, however, that the values for r_0/a are computed on rather low statistics[‡].

3.3.5 Algorithm performance

Any statement about the algorithm performance has to include autocorrelation times. Since different observables can have in general rather different autocorrelation times, also the algorithm performance is observable dependent. However, in the following we will use the plaquette integrated autocorrelation time $\tau_{\text{int}}(P)$ to determine the performance.

[†]First results applying an improved static action in the computation of the static potential already appeared in [63, 69].

[‡]The computation of the values for r_0/a was performed by A. Shindler and U. Wenger.

	κ	ν	ν from [110]	ν from [114]
<i>A</i>	0.15750	0.09(3)	0.69(29)	1.8(8)
<i>B</i>	0.15800	0.11(3)	0.50(17)	5.1(5)
<i>C</i>	0.15825	0.23(9)	0.28(9)	-

Table 3.4: Values of the cost figure ν compared to the corresponding values of Refs. [110] and [114], where available.

The values we measured for $\tau_{\text{int}}(P)$ can be found in table 3.3. It is interesting to observe that for runs *A* and *B* the values for $\tau_{\text{int}}(P)$ are smaller than the one found for the domain decomposition method. An explanation for this may be that in the algorithm of Ref. [110] a subset of all link variables is kept fixed during the molecular dynamics evolution, while in our HMC variant all link variables are updated.

Our value for $\tau_{\text{int}}(P)$ for run *A* is almost identical to the corresponding one found in Ref. [114]. In contrast, for simulation point *B* our value is a factor of three smaller, which is – we think – partly due to the significantly smaller acceptance rate of about 60% quoted in Ref. [114] for this point and partly due to the algorithmic improvements presented in this chapter.

A measure for the performance of the pure algorithm, which is implementation and machine independent, but incorporating the autocorrelation times is provided by the cost figure

$$\nu = 10^{-3}(2n + 3)\tau_{\text{int}}(P) \quad (3-27)$$

that has been introduced in Ref. [110]. n in Eq. (3-27) stands for either N_{NPF} in case a LF integration scheme is used or $2N_{\text{NPF}}$ in case a SW integration scheme is used. ν represents the average number of inversions of the Wilson-Dirac operator with the physical mass in units of thousands as needed to generate a statistically independent value of the average plaquette. Hence, in giving values for ν , we neglect the overhead coming from the remaining parts of the Hamiltonian.

Our values for ν together with the corresponding numbers from Ref. [110] and Ref. [114] are given in table 3.4. Compared to Ref. [110] our values for ν are smaller for simulation points *A* and *B* and comparable for run *C*. In contrast, the cost figure for the HMC algorithm with plain leap frog integration scheme is at least a factor 10 larger than the values found for our HMC algorithm variant. This gain is, of course, what we aimed for by combining multiple time scale integration with mass preconditioning and hence confirms our expectation. Unfortunately, due to the large statistical uncertainties of the ν values it is not possible to give a scaling of the cost figure with the mass. This holds for our values of ν as well as the ones of Ref. [110].

	N_{MV}			$\tau_{\text{int}}(P) \cdot \sum N_{\text{MV}}$	
	S_{PF_1}	S_{PF_2}	S_{PF_3}	this work	Ref. [114]
<i>A</i>	3800	4600	6600	90000	190750
<i>B</i>	6000	6900	11900	173600	1280000
<i>C</i>	31000	25500	-	565000	-

Table 3.5: Rounded number of matrix vector multiplications needed during one trajectory of length 0.5 for the different pseudo fermion actions without the usage of a chronological solver guess. We give also the sum of our numbers multiplied by the plaquette autocorrelation time and as a comparison the corresponding number from Ref. [114], where available.

3.3.6 Simulation cost

Although the value of ν is a sensible performance measure for the algorithm itself, since it is independent of the machine, the actual implementation and the solver, it cannot serve to estimate the actual computer resources (costs) needed to generate one independent configuration. Assuming that the dominant contribution to the total cost stems from the matrix vector (MV) multiplications, we give in table 3.5 the average number of MV multiplications N_{MV} needed for the different pseudo fermion actions to evolve the system for one trajectory of length $\tau = 0.5$. In addition we give the sum of these MV multiplications multiplied with the plaquette autocorrelation time together with the corresponding number from Ref. [114].

In order to compare to the numbers of Ref. [114] we remark that the lattice time extent is $T = 40$ in Ref. [114] compared to $T = 32$ in our case, but we do not expect a large influence on the MV multiplications coming from this small difference. Large influence on the MV multiplications, however, we expect from ll-SSOR preconditioning [126] that was used in Ref. [114] in combination with a chronological solver guess (CSG) [127].

Initially, when one compares the values of the cost figure for our HMC algorithm with the one of the plain leap frog algorithm as used in Ref. [114], one might expect that the number of MV multiplications shows a similar behavior as a function of the quark mass. However, inspecting table 3.5, we see that in terms of MV multiplications at simulation point *A* the HMC algorithm of Ref. [114] is only a factor of 2 slower than the variant presented in this chapter, while the values of ν are by a factor of about 20 different. The reason for this is two-fold: On the one hand ll-SSOR preconditioning together with a CSG method is expected to perform better than only even/odd preconditioning. On the other hand we think that the quark mass at this simulation point is still not small enough to gain significantly from multiple time scale integration. This illustrates that indeed the value of ν is not

immediately conclusive for the actual cost of the algorithm.

At simulation point B the relative factor between the MV multiplications needed by the two algorithms is already about 7. And finally, it is remarkable that for simulation point C the costs with our HMC variant are still a factor of 2 smaller than the costs for simulation point B with the algorithm used in Ref. [114], even though the masses are very different.

From this comparison we conclude that especially in the regime of small quark masses the HMC algorithm presented in this work is significantly faster than a HMC algorithm with single time scale leap frog integration scheme.

By looking at table 3.5 one notices that especially for simulation point C the number of MV multiplications needed for preconditioning is larger than the one needed for the physical operator. This comes from the fact that with the choice of algorithm parameters we have used the number of molecular dynamics steps for the mass preconditioned operator is large. This possibly indicates potential to further improve the performance by tuning the preconditioning masses and time scales.

We stress here again that the number of matrix vector operations is highly solver dependent, and therefore, every improvement to reduce the solver iterations will decrease the cost for one trajectory. Promising improvements are for instance the use of a chronological inversion method [127] (or similar methods [128]) or the use of a solver based on domain decomposition as adopted for QCD in Ref. [129]. We tested the chronological inversion method and found in total not more than 20% gain in matrix vector operations.

Finally, it is interesting to compare the number of matrix vector multiplications reported in table 3.5 with a HMC algorithm where mass preconditioning and multiple time scale improvements are switched off and CSG is not used. For instance for a simulation with a Sexton-Weingarten improved integration scheme at $\kappa = 0.15825$ there are 120 molecular dynamics steps needed to get acceptance. This corresponds to 240 inversions of Q^2 , which amounts to about 720000 matrix vector multiplications. Compared to run C this is at least a factor 10 more. We did only a few trajectories to get an estimate for this number, so we cannot say anything about autocorrelation time.

Of course it would be interesting to compare also to a HMC algorithm with mass preconditioning but without multiple time scale integration. This, however, needs again a tuning of the mass parameters and would therefore be quite costly and we did not attempt to test this situation here.

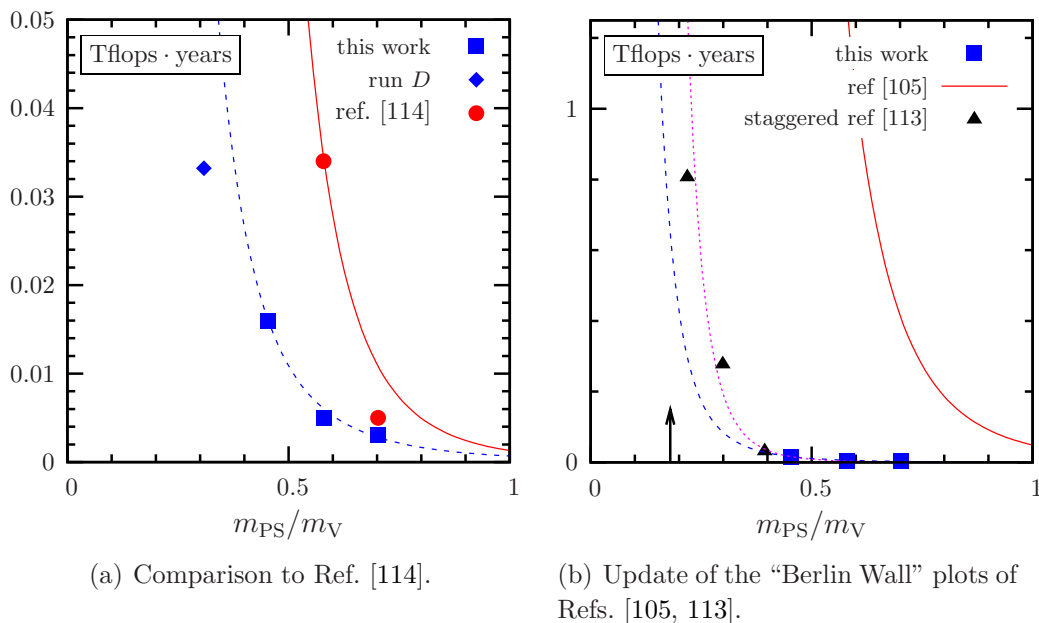


Figure 3.4: Computer resources needed to generate 1000 independent configurations of size $24^3 \times 40$ at a lattice spacing of about 0.08 fm in units of Tflops · years as a function of m_{PS}/m_V . In (a) we compare our results represented by squares to the results of Ref. [114] represented by circles. The lines are functions proportional to $(m_{PS}/m_V)^{-4}$ (dashed) and $(m_{PS}/m_V)^{-6}$ (solid) with a coefficient such that they cross the data points corresponding to the lightest pseudo scalar mass. The diamond represents the preliminary result of run D (see text). In (b) we compare to the formula of Eq. 3-28 [105] (solid line) by extrapolating our data with $(m_{PS}/m_V)^{-4}$ (dashed) and with $(m_{PS}/m_V)^{-6}$ (dotted), respectively. The arrow indicates the physical pion to rho meson mass ratio. Additionally, we add points from staggered simulations as were used for the corresponding plot in Ref. [113]. Note that all the cost data were scaled to match a lattice time extend of $T = 40$.

3.3.7 Scaling with the mass

An important property of an algorithm for lattice QCD is the scaling of the costs with the simulated quark mass. The naive expectation is that the number of solver iterations grows like m_q^{-1} and also the number of molecular dynamics steps is proportional to m_q^{-1} , see for instance Ref. [130] or Ref. [105]. Since also the integrated autocorrelation time is assumed to grow like m_q^{-1} , it is expected that the HMC algorithm costs scale with the quark mass as m_q^{-3} or equivalently as m_{PS}^{-6} . In contrast, for our HMC algorithm variant we expect a much weaker scaling of $\Delta\tau$ and also of the number of solver iterations. Indeed, we see that the costs for our HMC algorithm variant is consistent with a m_q^{-2} or m_{PS}^{-4} behavior when the autocorrelation time is taken into account.

We have translated the number of matrix vector multiplications from table 3.5 into costs in units Tflops · years and plotted the computer resources needed to generate 1000 independent configurations of size $24^3 \times 40$ at a lattice spacing of ~ 0.08 fm

as a function of $m_{\text{PS}}/m_{\text{V}}$ in figure 3.4(a) together with the results of Ref. [114]. Note that we have scaled our costs like $(40/32)^{1.25}$ corresponding to the expected volume dependence (cf. [105]) to match the different time extents and, moreover, we used the plaquette autocorrelation time as an estimate for the autocorrelation time.

The solid (dashed) line is not a fit to the data, but a function proportional to $(m_{\text{PS}}/m_{\text{V}})^{-4}$ ($(m_{\text{PS}}/m_{\text{V}})^{-6}$) with a coefficient that is fixed by the data point corresponding to the lightest pseudo scalar mass. These functional dependencies on $(m_{\text{PS}}/m_{\text{V}})$ describe the data reasonably well. However, from our few data points it is not possible to decide on the value of the exponent in the quark mass dependence of the costs. But, it is clear from the figure that with multiple time scale integration and mass preconditioning the “wall” – which renders simulations at some point infeasible – is moved towards smaller values of the quark mass.

An additional indication for the scaling properties of the algorithm towards smaller masses is given by the preliminary result of run D . It is represented in figure 3.4(a) by the single diamond. For this point we used our current number of MV multiplications as measured for run D and the value for m_{PS} as given in Ref. [119]. Moreover, we extrapolated the value am_{V} in κ and the value of τ_{int} in $1/m_{\text{PS}}^2$. The result as we plot it in the figure thus has certainly a significant error. Nevertheless, even if the “true” result will be a factor of two larger, the point is still in excellent agreement with the anticipated scaling proportional to $(m_{\text{PS}}/m_{\text{V}})^{-4}$.

On a larger scale we can compare the extrapolations of our cost data to the formula given in Ref. [105]

$$C = K \left(\frac{m_{\text{PS}}}{m_{\text{V}}} \right)^{-z_{\pi}} L^{z_L} a^{-z_a}, \quad (3-28)$$

where the constant K can be found in Ref. [105] and $z_{\pi} = 6$, $z_L = 5$ and $z_a = 7$. The result of this comparison is plotted in figure 3.4(b), which is an update of the “Berlin Wall” figure that can be found in Ref. [113]. We plot the simulation costs in units of Tflops · years versus $m_{\text{PS}}/m_{\text{V}}$, where we again scaled the numbers in order to match a lattice time extend of $T = 40$. The dashed and the dotted lines are extrapolations from our data proportional to $(m_{\text{PS}}/m_{\text{V}})^{-4}$ and $(m_{\text{PS}}/m_{\text{V}})^{-6}$, respectively, again matching the data point corresponding to the lightest pseudo scalar mass. The solid line corresponds to Eq. (3-28) with K taken from Ref. [105]. In addition we plot data from staggered simulations as were used for the plot in Ref. [113]. That the corresponding points lie nearly on top of the dotted line is accidental.

Conservatively one can conclude from figure 3.4(b) that with the HMC algorithm described in this chapter at least simulations with $m_{\text{PS}}/m_{\text{V}} \approx 0.3$ are feasible, even though $L = 1.93$ fm is too small for such values of the masses. Taking the more optimistic point of view by assuming that the costs scale with $z_{\pi} = 4$, even simulation

with the physical $m_{\text{PS}}/m_{\text{V}}$ ratio and a lattice spacing of 0.08 fm become accessible, with again the caveat that L/a needs to be increased.

Independent of the value for z_{π} , figure 3.4(b) reveals that the costs for simulations with staggered fermions and with Wilson fermions in a comparable physical situation are of the same order of magnitude, if for the simulations with Wilson fermions an algorithm like the one presented in this work is used. It would be interesting to see whether the techniques applied in this work perform similarly well for staggered fermions.

We would like to point out that we did not try to tune the parameters to their optimal values. The aim of this work was to give a first comparison of mass preconditioned HMC algorithm with multiple time scale integration to existing performance data, i.e. data for a HMC algorithm preconditioned by domain decomposition [110] and data for the HMC algorithm variant of Ref. [114]. We are confident that there are still improvements possible by further tuning of the parameters in our variant of the HMC algorithm.

3.4 Conclusion

In this chapter we have presented and tested a variant of the HMC algorithm combining multiple time scale integration with mass preconditioning (Hasenbusch acceleration). It is based on the idea to arrange mass preconditioning such that the force contributions from the different parts in the Hamiltonian are strictly ordered with respect to the absolute value of the force and that the most expensive part has the smallest contribution to the total force. Then the most expensive part can be integrated on the largest time scale.

Our aim was to perform a first investigation of the performance properties of this HMC algorithm by comparing it to other state of the art HMC algorithm variants in the same physical situation, i.e. for pseudo scalar masses in the range of 380 to 670 MeV, a lattice spacing of about 0.08 fm and a lattice size of $L \approx 2$ fm with two flavors of mass degenerate Wilson fermions. We verified our implementation by comparing results for the plaquette and for the pseudo scalar, the vector and the current quark mass to results available in the literature finding full agreement.

We have shown that indeed the aforementioned idea can be realized by tuning the additional (unphysical) mass parameters introduced for mass preconditioning. In this set-up the performance of our variant in terms of the cost figure in Eq. (3-27) is compatible to the one observed for the HMC algorithm with multiple time scales and domain decomposition as preconditioner introduced in Ref. [110] and clearly superior to the one for the HMC algorithm with a simple leap frog integration scheme as used in Ref. [114].

While the cost figure provides a clean algorithm performance measure we also compare the simulation costs in units of Tflops · years to existing data. This comparison is summarized in an update of the “Berlin wall” plot of Ref. [113], which can be found in figure 3.4. We could show that with the HMC algorithm presented in this chapter the wall is moved towards smaller values of the quark mass and that simulations with a ratio of $m_{\text{PS}}/m_{\text{V}} \approx 0.3$ become feasible at a lattice spacing of around 0.08 fm and $L \approx 2$ fm. We have preliminary results for a simulation point with a pseudo scalar mass of around 300 MeV and $m_{\text{PS}}/m_{\text{V}} \approx 0.3$, which is in excellent agreement with all the results mentioned above. In particular this simulation point seems to confirm that the algorithm costs scale proportional to $(m_{\text{PS}}/m_{\text{V}})^{-4}$.

The HMC variant presented here has the advantage of being applicable to a wide variety of Dirac operators, including in principle also the overlap operator. In addition its implementation is straightforward, in particular in an already existing HMC code. We remark that the parallelization properties of our HMC variant and the one of the algorithm presented in [110] can be very different depending on whether a fine- or a coarse-grained massively parallel computer architecture is used.

From a stability point of view our results reveal that even for Wilson fermions it is very well possible to simulate quark masses of the order of 20 MeV when using the algorithmic ideas presented in this work. The presently ongoing simulation with even smaller quark mass is also running without any practical problems, but the statistics is not yet adequate to say something definite. However, it is a remarkable result by itself that there are now at least two algorithms available allowing for *stable* simulations with Wilson fermions at low values of the quark masses. Remarkable, because only short time ago this was thought to be hardly possible and it immediately raises the question for an explanation: one can speculate that the observed stability is mainly due to noise reduction provided by the additional pseudo fermion fields and former simulations yielded problems, because the stochastic approximation for the determinant was not sufficient.

The results presented in this chapter are mostly based on empirical observations and on simulations for only one value of the coupling constant $\beta = 5.6$. It remains to be seen how our HMC variant behaves for larger values of β , which, as well as smaller quark masses and theoretical considerations about the scaling properties with the quark mass needs further investigations. Moreover, a more systematic study of the interplay between integration schemes, step sizes, (preconditioning and physical) masses and the simulation costs is needed. Those investigations will hopefully also provide a better understanding of the algorithm itself and its dynamics. Of course, the algorithm should also be tested for tmQCD, even though we do not expect a large difference to the pure Wilson case.

Finally, we think that there are further improvements possible by the usage of a Polynomial HMC (PHMC) algorithm [131, 132, 133, 134]. With such an algorithm

one could treat the lowest eigenvalues of the Dirac operator exactly and/or by reweighting. In this set-up the large fluctuations in the force might be significantly reduced, if the lowest eigenvalues are responsible for those. Then it might be possible to further reduce the number of inversions of the badly conditioned physical operator needed to evolve the system. In addition, a PHMC algorithm would immediately allow for simulations with three or more flavors of quarks.

Chapter 4

Phase structure of lattice QCD

Understanding the phase structure of lattice QCD is an important pre-requisite before starting large scale simulations [113]. As we will see below, the main result of this chapter is that Wilson type fermions generically show a first order phase transition around the chiral point for lattice spacings below a certain threshold. The existence of this first order phase transition, which is due to lattice artifacts, has important consequences for simulations with this kind of lattice fermions.

First, approaching the “physical point”, at which the pion mass assumes its value as measured in experiment, the algorithms used in lattice simulations suffer from a substantial slowing down, as explained in the previous chapter. Since simulations are also restricted to finite lattice spacings, lattice QCD in general involves extrapolations to the physical point and to the continuum. In order to be able to control those extrapolations it is necessary to know whether there is indeed a smooth extrapolation possible.

Second, numerical simulations in the vicinity of first or second order phase transitions are problematic. It is not obvious whether the algorithms correctly sample configuration space in such a situation. And, assuming they do so, usually one observes practical problems such as long thermalization and autocorrelation times. Therefore, as – unlike the finite temperature case – the phase transition itself is not the topic under investigation, one would like to avoid to simulate close to a phase transition.

It is important to realize that the above mentioned first order phase transition for Wilson type lattice fermions imposes a lower bound on the value of the pseudo scalar mass that can be simulated at a given lattice spacing. The actual value of the lower bound to the pseudo scalar mass, however, turns out to depend on the particular gauge action discretization providing an opportunity to circumvent or at least reduce the aforementioned complications.

We remark that strictly speaking in a finite volume with a finite number of lattice

points there cannot exist a phase transition. Nevertheless, already in a finite, but large enough volume the effects of a phase transition in infinite volume can be visible and rather strong.

For QCD it is widely believed that chiral symmetry is spontaneously broken by the ground state and one expects a first order phase transition where the scalar condensate jumps from negative to positive value when the mass is changed from small negative to small positive values, or vice versa. In this chapter we are going to discuss in lattice chiral perturbation theory how the explicit chiral symmetry violations of Wilson type fermions at non-vanishing lattice spacing will affect the continuum picture, which will allow an interpretation of our numerical results. Thereafter, we will present our numerical results on the phase structure of lattice QCD in a regime of lattice spacings where so far a systematic study was missing. These results will provide evidence for a first order phase transition, as mentioned before. Finally, we will outline how to reduce the effects of this phase transition by choosing a different discretization of the gauge part in the action.

4.1 Effective potential model

As we have discussed in section 1.2.2, the massless Lagrangian of QCD for two flavors of quarks is symmetric under the chiral group $SU_V(2) \times SU_A(2)$, which is spontaneously broken down to $SU_V(2)$ by the ground state of the theory. To describe reality the Lagrangian contains a quark mass term explicitly breaking the aforementioned symmetry. Therefore the vector and axial currents are not exactly conserved.

However, since the masses of up- and down-quark are small, also the divergence of the currents vanishes approximately and the masses might be treated as a small perturbation to the massless theory as is done in the chiral perturbation theory (χ PT) [135, 136, 137]. From a principle point of view lattice calculations include all the low energy structure of QCD with the quark masses being free parameters. Therefore, there is a priori no need for a χ PT if expectation values can be computed on the lattice on a non-perturbative level at realistic values of the quark mass. However, nowadays lattice calculations are not yet able to provide reliable results obtained with values of the quark masses as small as estimated in experiment.

Hence, χ PT might serve as an useful tool to connect the results for physical quantities obtained from lattice simulations performed at un-physically large quark masses with those at the physical point. However, χ PT is valid only for small quark masses below a certain upper bound or in different words it has a finite convergence radius, which allows to make contact to lattice calculations if and only if lattice simulations with masses below this bound are possible. The actual value of this

bound is unknown, but lattice simulations with pion masses below 300 MeV will most likely be needed.

As originally χ PT is an effective low energy theory of continuum QCD, it is not valid at finite values of the lattice spacing a . However, one can also formulate an effective theory including the lattice spacing, which is then treated as a small, additional parameter.

4.1.1 Chiral perturbation theory on the lattice

In section 1.2.3 we have introduced the Symanzik effective theory, which is expected to describe the lattice theory close to the continuum by an effective continuum Lagrangian. The usual terms in the continuum Lagrangian are supplemented by contributions proportional to powers of the lattice spacing. We have also discussed in section 1.2.3 that the effective fermionic Lagrangian is of the form (see Eq. (1-44))

$$\mathcal{L}_{\text{eff}} \sim \bar{\psi}(\gamma_\mu D_\mu + m)\psi + c_{\text{sw}} a \bar{\psi} i \sigma_{\mu\nu} F_{\mu\nu} \psi + \mathcal{O}(a^2). \quad (4-1)$$

This amounts to QCD with a Pauli term. And since the Pauli term transforms under chiral rotations exactly like a mass term [41], the corresponding chiral Lagrangian is already known.

We will not discuss lattice χ PT (L χ PT) in detail, as it can for instance be found for Wilson lattice QCD in [138], but we rather follow the qualitative discussion of reference [41] to immediately access the phase structure of lattice QCD at small quark masses.

In terms of a SU(2) matrix-valued field Σ , transforming under independent SU(2) rotation U_L and U_R as $\Sigma \rightarrow U_L \Sigma U_R^\dagger$ the kinetic part (which is the chiral Lagrangian in absence of mass and Pauli term) can be written as

$$\mathcal{L}_\chi = \frac{f_\pi^2}{4} \text{Tr} \left(\partial^\mu \Sigma^\dagger \partial_\mu \Sigma \right), \quad (4-2)$$

and the potential energy is given by

$$\mathcal{V}_\chi = -\frac{c_1}{4} \text{Tr} \left(\Sigma + \Sigma^\dagger \right) + \frac{c_2}{16} \left\{ \text{Tr} \left(\Sigma + \Sigma^\dagger \right) \right\}^2 \quad (4-3)$$

with coefficients $c_1 \sim m \Lambda_{\text{QCD}}^3 + a \Lambda_{\text{QCD}}^5$ and $c_2 \sim m^2 \Lambda_{\text{QCD}}^2 + m a \Lambda_{\text{QCD}}^4 + a^2 \Lambda_{\text{QCD}}^6$. The factors or Λ_{QCD} are required by dimensional analysis and dimensionless coefficients of order unity are dropped [41]. In the following we will be particularly interested in the case where $m' = m - a \Lambda_{\text{QCD}}^2 \sim a^2 \Lambda_{\text{QCD}}^3$. In this regime the coefficients can be simplified to

$$c_1 \sim m' \Lambda_{\text{QCD}}^3, \quad c_2 \sim a^2 \Lambda_{\text{QCD}}^6. \quad (4-4)$$

In this situation the two terms in the potential energy become comparable, since $c_1 \sim c_2$, and it will lead to the prediction of a non-trivial phase structure of Wilson lattice QCD for small quark masses. Moreover, the coefficient c_1 parameterizes the bare quark mass and c_2 is proportional to a^2 .

By denoting with A the flavor singlet component of Σ , the latter can be expressed like

$$\Sigma = A + i \sum_{r=1}^3 B_r \tau_r, \quad (4-5)$$

with $1 = A^2 + \sum B_r B_r$ and τ_r the three Pauli matrices. This constrains A , which corresponds to the scalar condensate, to lie between -1 and 1 inclusive. The potential then reads

$$\mathcal{V}_\chi = -c_1 A + c_2 A^2 = A(c_2 A - c_1). \quad (4-6)$$

Note that B_3 corresponds to $\langle \bar{\psi} \gamma_5 \tau_3 \psi \rangle$.

In order to find the ground state Σ_0 the effective potential has to be minimized. Σ can then be expanded around Σ_0 and observables like pion masses can be extracted in terms of the quark mass, c_2 and the pion decay constant f_π . And it turns out that the results depend strongly on the sign of the coefficient c_2 , which parameterizes the lattice artifacts.

In case of *positive* c_2 there exists a phase in the region of bare quark masses defined by $-2c_2 < c_1 < 2c_2$ where the flavor symmetry is broken. This expresses itself in massless charged pions, because they are Goldstone bosons of spontaneous flavor symmetry breaking, but a massive uncharged pion.

At the boundaries $c_1 = \pm 2c_2$ all three pions are massless and the system undergoes a second order phase transition, while outside the phase with ($|c_1| > 2c_2$) the flavor symmetry is preserved by the ground state and all three pions are massive. This scenario is called the *Aoki scenario*, because S. Aoki first predicted the existence of such a flavor symmetry breaking phase [139, 140, 141].

The alternative is that c_2 is *negative*. In this case the flavor symmetry is preserved in the whole range of values for c_1 , but it does not exist a value of c_1 where the pions are massless, since the pion masses are given by [41]

$$m_\pi^2 = \frac{1}{f_\pi^2} (|c_1| + 2|c_2|). \quad (4-7)$$

At $c_1 = 0$ the vacuum expectation value of Σ jumps from $A = -1$ to $A = +1$. Since this jump appears at non-zero pion mass the thermo-dynamical description of the behavior near $c_1 = 0$ corresponds to a first order phase transition.

This situation is called *normal scenario* due to its similarity to the continuum first order chiral phase transition around zero quark mass. In the continuum, of course, the pions as Goldstone bosons become massless in the chiral point in contrast

4.1. EFFECTIVE POTENTIAL MODEL

to a non-zero minimal value of the pion mass characterizing the normal scenario on the lattice.

The minimal value of the squared pion mass is in $L\chi$ PT for vanishing value of c_1 proportional to c_2 , which is of $\mathcal{O}(a^2)$. This means, if a negative value of c_2 is realized in Wilson lattice QCD the minimal value of m_π^2 will vanish like a^2 when the continuum limit is performed. On the other hand, if the Aoki scenario is realized, the width of the Aoki phase will vanish like $\Delta m \sim a^3$ [41]. Unfortunately $L\chi$ PT by itself is not able to make any prediction about the sign of c_2 .

Although the just discussed phase structure for Wilson lattice QCD from $L\chi$ PT is known already quite some time [41], the corresponding investigation for the Wilson twisted mass formulation of lattice QCD was missing. It was published in Refs. [142, 97, 143] only after we published parts of the results that will be presented in this chapter. The introduction of a twisted mass term in the low energy effective Lagrangian turns out to be straightforward. The only difference is one additional term in the effective potential energy

$$\mathcal{V}_\chi = -c_1 A - c_3 B_3 + c_2 A^2. \quad (4-8)$$

The new coefficient $c_3 \sim \mu$ is parameterizing the twisted mass in the effective theory while the other coefficients are given as in the former discussion. If we now consider values of $m' \sim \mu \sim a^2 \Lambda_{\text{QCD}}^3$ we again have to distinguish between negative and positive values of c_2 . The sign of c_3 depends on the value of μ , which we choose to be positive, since it does not influence the qualitative picture of the phase structure.

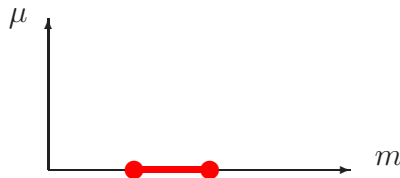


Figure 4.1: Phase diagram in the $m - \mu$ plane for the Aoki scenario.

For $c_2 > 0$ and $\mu \neq 0$ fixed the Aoki scenario changes as follows: charged and uncharged pions are massive for all values of c_1 , even though the flavor $SU(2)$ symmetry is explicitly broken. The value A changes continuously as a function of c_1 and there is no phase transition. The situation for $c_2 > 0$ is summarized schematically in figure 4.1. For $\mu = 0$ the two second order phase transition points are indicated by the filled circles. On the line between these two circles the charged pions are massless and the uncharged pion is massive. Note that this line corresponds to a first order transition line where B_3 jumps when the values of μ sweep across zero.

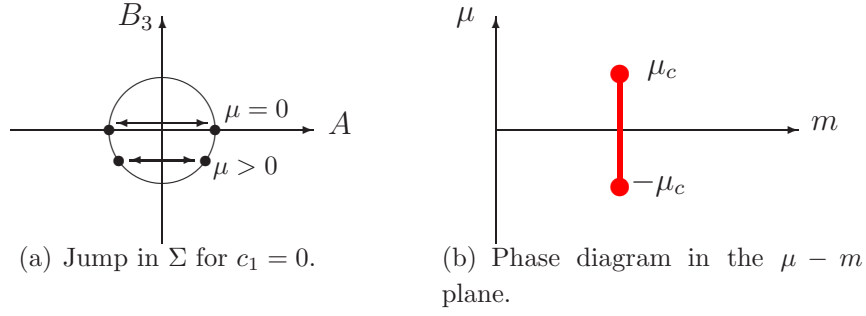


Figure 4.2: Phase diagram and jump in Σ for the normal scenario ($c_2 < 0$).

In the case of $c_2 < 0$ the influence of a non-zero value for the twisted mass parameter μ is different. Here μ plays the rôle of a “magnetic field” and shifts away the minimum of the potential \mathcal{V}_χ from the values $A = \pm 1$. As indicated in figure 4.2(a) B_3 adopts a non-zero value and $|A| < 1$. With increasing modulus of μ the two minima of the potential approach each other and the transition line has an endpoint for

$$\mu_c \sim a^2. \quad (4-9)$$

For $\mu \neq 0$ flavor and parity symmetry are explicitly broken and at $\mu = \mu_c$ the uncharged pion becomes massless while the charged pions remain massive. At $c_1 = 0$ and $\mu < \mu_c$ a first order phase transition takes place. The corresponding phase diagram is schematically represented in figure 4.2(b), where we indicate the endpoints of the first order transition line at μ_c by filled circles.

Before turning to a numerical check of L χ PT it is useful to summarize what was known in the literature so far about the phase structure of Wilson lattice QCD. In a recent re-investigation [144, 145] of the Aoki phase with the Wilson gauge action for values of the coupling $\beta < 5$ the authors found in agreement with previous publications [146, 141, 147, 148, 149] evidence for an Aoki phase only for values of $\beta < 4.6$. On the other hand, for values of $\beta = 4.6$ and $\beta = 5$ they found no evidence for the realization of an Aoki scenario.

For values of $\beta > 5$ (Wilson plaquette gauge action), however, a systematic investigation of the phase structure was missing. Even though there exist several indications in the literature (see for instance [150]) for the realization of the normal scenario in this region of β values the connection to the aforementioned results from L χ PT was never explicitly mentioned.

4.2 Numerical results

4.2.1 Simulation points

We have chosen three values of the bare coupling constant β to study the phase structure of lattice QCD with Wilson plaquette gauge action and Wilson twisted mass fermions. The simulation parameters can be found in table 4.1. The values of $a\mu$ were chosen such that $r_0\mu$ is roughly constant for all values of β . Moreover, the lattice sizes were taken to have a physical volume of at least 2 fm in order to allow for save extraction of meson physics.

The values of the lattice spacing quoted in table 4.1 are estimates to give an orientation. At $\beta = 5.2$ we have in addition to the lattice size quoted in table 4.1 results for $16^3 \times 32$ lattices to check for finite volume effects.

β	$L^3 \times T$	$a\mu$	a [fm]
5.1	$12^3 \times 24$	0.013	0.20
5.2	$12^3 \times 24$	0.010	0.16
5.3	$16^3 \times 32$	0.008	0.14

Table 4.1: Simulation points for Wilson plaquette gauge action.

For the next subsections we will mainly concentrate on the results obtained at $\beta = 5.2$, because the results are qualitatively the same for the three values of β . After this description of the first order phase transition phenomenon we will in a next step analyze also the scaling with the lattice spacing. For all these investigations we used the hopping parameter representation of the Wilson twisted mass lattice action in the twisted basis as given by Eq. (1-46).

4.2.2 Thermal cycles

We started our investigation of the phase diagram of zero temperature lattice QCD by performing thermal cycles in κ while keeping fixed $\beta = 5.2$ and the value of the twisted mass parameter $a\mu$. These cycles are performed such that a starting value of κ_{start} is chosen and then κ is incremented, without performing further intermediate thermalization sweeps, until a final value of κ_{final} is reached. At this point the procedure is reversed and κ is decremented until the starting value κ_{start} is obtained back. At each value of κ 150 configurations are produced and averaged over.

In fig. 4.3 we show three such thermal cycles, performed at $a\mu = 0$, $a\mu = 0.01$ and $a\mu = 0.1$ from bottom to top. In the cycles signs of hysteresis effects can be seen for $a\mu = 0$ and $a\mu = 0.01$ while for the largest value of $a\mu = 0.1$ such

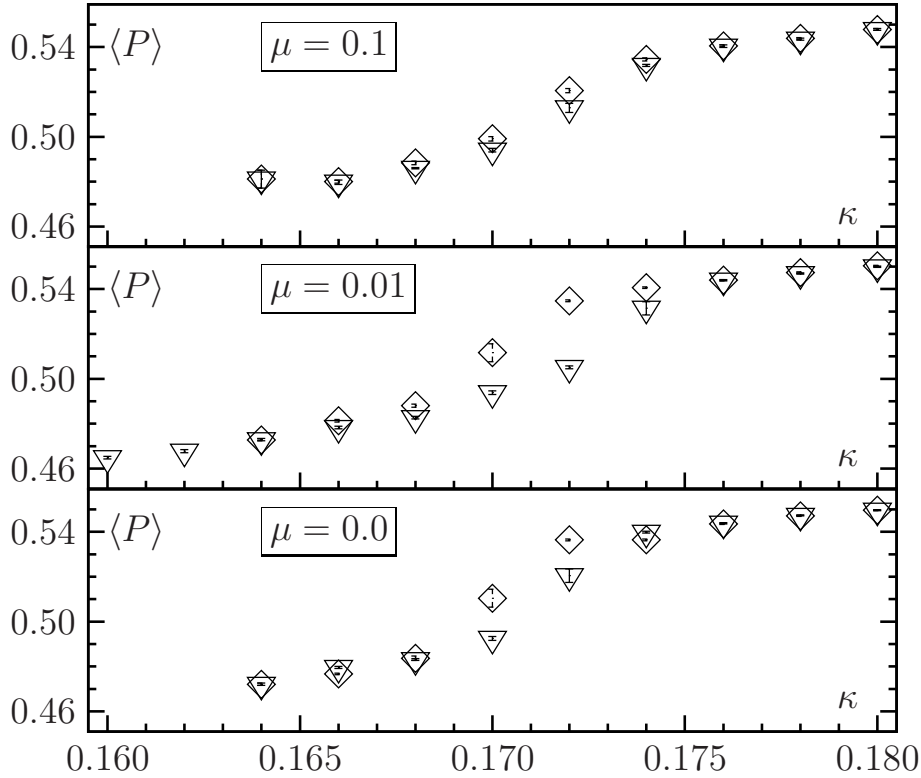


Figure 4.3: Thermal cycles in κ on $8^3 \times 16$ lattices at $\beta = 5.2$. The plaquette expectation value is shown for $a\mu = 0.1$, $a\mu = 0.01$ and $a\mu = 0$ (top to bottom). The triangles refer to increasing κ -values, the diamonds to decreasing ones. The errors are the naive statistical errors, without taking the autocorrelation time into account.

effects are hardly visible. Hysteresis effects in thermal cycles *may be* signs of the existence of a first order phase transition. However, they should only be taken as first indications. Nevertheless, they provide most useful hints for further studies to search for meta-stable states.

4.2.3 Meta-stabilities

Guided by the results from the thermal cycles, we next performed simulations at fixed values of $a\mu$ and κ , starting with ordered and disordered configurations, staying again at $\beta = 5.2$. In fig. 4.4 we show the Monte Carlo time evolution of the plaquette expectation value, in most cases on a $12^3 \times 24$ lattice. For several values of κ we find coexisting branches with different average values of the plaquette. The gap (the “latent heat”) appears to be rather large. At $\kappa = 0.1717$ we show the history of the plaquette expectation value also on a larger ($16^3 \times 32$) lattice. It seems that the gap in the plaquette expectation value does not depend much on the lattice size, suggesting that the meta-stability we observe here is not a finite volume effect. In most cases the twisted mass is $a\mu = 0.01$, except for the picture right in the upper

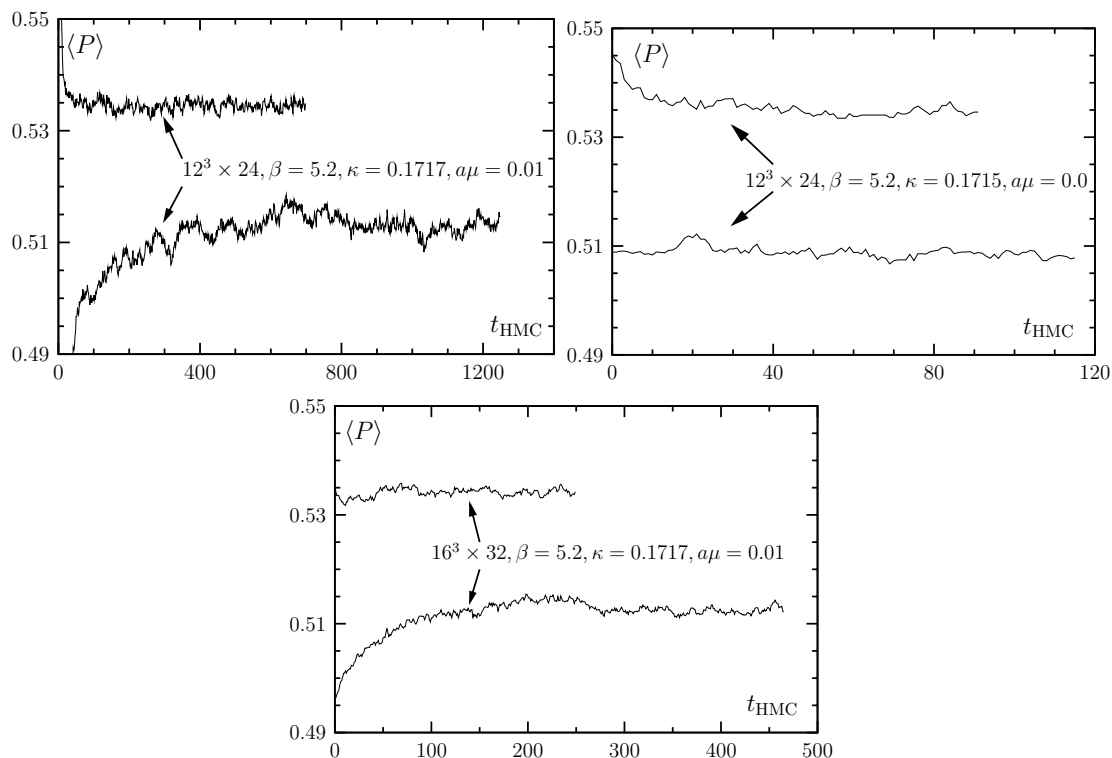


Figure 4.4: Meta-stable states at $\beta = 5.2$. The plaquette value is plotted as a function of the HMC time, i.e. the number of trajectories. The lattice size is $12^3 \times 24$ except for the bottom figure where it is $16^3 \times 32$. The value for the twisted mass parameter is $a\mu = 0.01$ except for the rightmost figure where it is $a\mu = 0$.

line where it is $a\mu = 0$. Since the meta-stabilities are also observed with $a\mu = 0$ it is already excluded that the phenomenon is only due to the twisted mass term.

The lifetime of a meta-stable state, i.e. the time before a tunneling to the stable branch occurs, depends on the algorithm used. In fact, one may wonder, whether the appearance of the meta-stable states seen in fig. 4.4 may not be purely an artefact of our algorithms. We cannot completely exclude this possibility but we believe it is very unlikely: we employed two very different kinds of algorithms in our simulations as explained in subsection 1.4.2. We observe the meta-stable states with both, the HMC and the TSMB algorithm. We also inter-changed configurations between the two algorithms: a configuration generated with algorithm A was iterated further with algorithm B and vice versa. We find that in such situations the plaquette expectation value remains in the state where it has been before the interchange of configurations took place. In addition, as we shall see below, the two states can be characterized by well defined and markedly different values of basic physical quantities. We therefore conclude that the meta-stable states are a generic phenomenon of lattice QCD in the Wilson or Wilson twisted mass formulation. This conclusion is strongly supported by the fact that it is consistent with the picture provided by

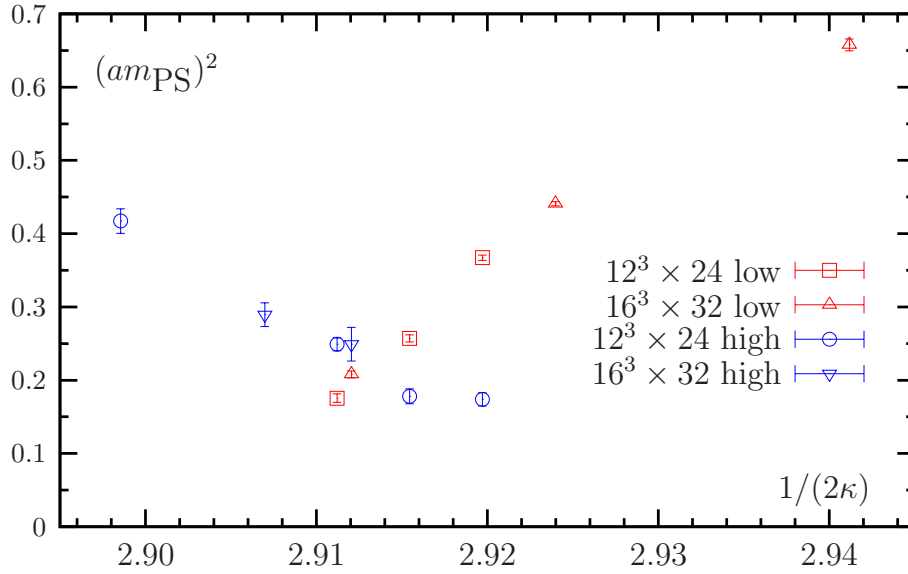


Figure 4.5: The pseudo scalar mass squared in lattice units as a function of $1/(2\kappa)$ on two lattice sizes measured separately on configurations in the two (meta)stable states. These runs were performed at $\beta = 5.2$ and $a\mu = 0.01$.

L χ PT, which we discussed in section 4.1.1.

4.2.4 Pseudo scalar and quark masses

While in the preceding paragraph we looked at the plaquette expectation value indicating a first order phase transition we will discuss this phenomenon in the following paragraph in terms of pseudo scalar and quark masses. To this end we selected separately configurations with high and with low plaquette expectation value and measured the pseudo scalar mass m_{PS} and the untwisted PCAC quark mass m_χ^{PCAC} , which we obtained as explained in section 1.3.

In fig. 4.5 we show the pseudo scalar mass squared in lattice units as a function of $1/(2\kappa)$. We observe that the pseudo scalar mass is rather large and the most striking effect in the graph is that it can have two different values at the same κ value. Moreover, the minimal value of the pseudo scalar mass is not zero, but assumes a rather large value.

If we consider the quark mass m_χ^{PCAC} in fig. 4.6, we see that in the states with a low plaquette expectation value the mass is positive while for high values of the plaquette expectation it is negative. These quark masses with opposite sign coexist for some values of κ .

Figs. 4.4-4.6 clearly reveal that for small enough values of μ meta-stabilities show up in the quantities we have investigated, such as m_{PS} , m_χ^{PCAC} and the average plaquette, if m_0 is close to its critical value. What “small enough μ ” means is

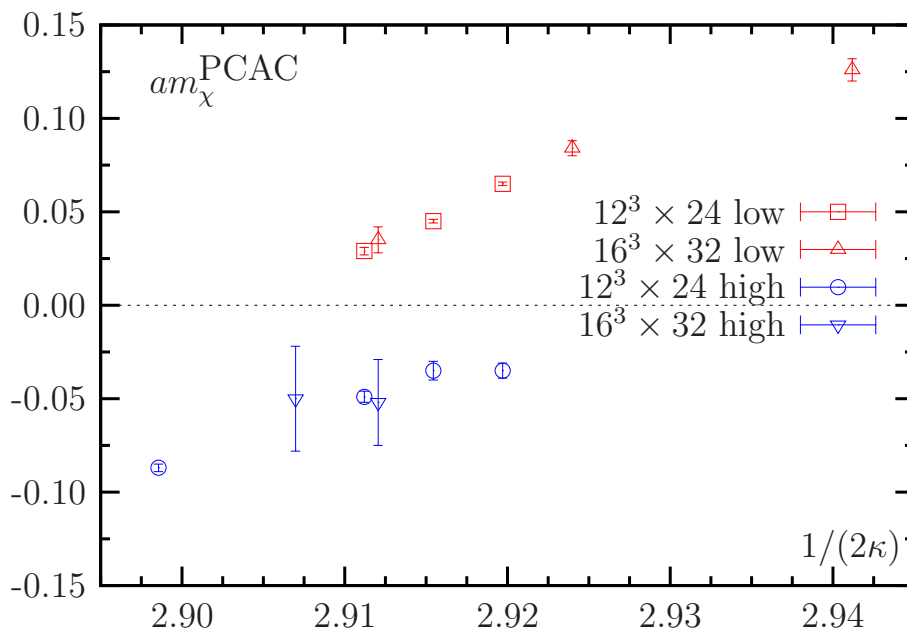


Figure 4.6: The PCAC quark mass m_χ^{PCAC} in lattice units as defined in Eq. (1-80) as a function of $1/(2\kappa)$ on two lattice sizes measured separately on configurations in the two (meta)stable states. The values of $\beta = 5.2$ and $a\mu = 0.01$ are fixed.

likely to change with β , see below. As a matter of fact, when m_0 is significantly larger (smaller) than m_{crit} we find m_χ^{PCAC} to be positive (negative) and no signal of meta-stabilities.

Of course, one might ask the question how we can know that the values of m_0 are close to its critical value, since we do not observe any value for m_0 where $m_\chi^{\text{PCAC}} = 0$. The answer to this question is that – as argued already above – in a finite volume there exist no phase transition and therefore, m_χ^{PCAC} must be an analytical function of m_0 . Only due to the disability of the algorithm to correctly sample configuration space in this region of m_0 values, we get the impression (on the remnant) of a first order phase transition. Analyticity in turn implies that in the region of meta-stabilities an optimal algorithm would find a value of m_0 where $m_\chi^{\text{PCAC}} = 0$, which is, however, a finite volume effect: in infinite volume physical observables such as m_χ^{PCAC} jump at the phase transition point and hence, m_χ^{PCAC} does not become zero. The chiral point is then defined at the phase transition point.

The remark that meta-stabilities take place for m_0 close to its critical value is important for the interpretation of the observed phenomenon. As we explained in section 4.1, $L\chi\text{PT}$ predicts two different scenarios for the phase structure at small quark masses. The so called normal scenario is characterized as follows: firstly, a first order phase transition appears when the untwisted quark mass sweeps across zero as long as the twisted mass parameter is smaller than a critical value μ_c . Secondly, the pseudo scalar mass has a minimal value that is larger than zero.

As meta-stabilities are expected to take place in the vicinity of a first order phase transition, we therefore conclude that at this value of β the normal scenario is realized.

Even though we discussed in this subsection only one value of β , the same phenomenon appears for $\beta = 5.1$ and for $\beta = 5.3$: we observe meta-stabilities and a non-zero value of the minimal pseudo scalar mass. In order to simplify the language, we denote simulation points in the phase with high plaquette expectation value with “high”, and correspondingly the simulation points with low plaquette expectation value with “low”.

The dependence of the phase transition on the lattice spacing will be the topic of the following subsection.

4.2.5 The phase transition as a function of the lattice spacing

As mentioned already before, we have apart from $\beta = 5.2$ also simulation points with $\beta = 5.1$ and $\beta = 5.3$. The details for all our simulation points can be found in tables 4, 5, and 6 of Ref. [151]. The values of $a\mu$ were fixed for each β value such that $r_0\mu \approx 0.03$ for all values of β . Note that the value of r_0/a depends on the value of the quark mass and therefore, we had to choose a reference value for r_0/a . We have chosen this reference point to have $(r_0m_{\text{PS}})^2 = 1.5$ and interpolated our data for $\beta = 5.1$ and $\beta = 5.3$ to this point, while for $\beta = 5.2$ a short extrapolation was necessary. We again used the ROOT and MINUIT packages from CERN to perform the corresponding fits. The parameters are summarized in table 4.1.

In figure 4.7 we have plotted the plaquette expectation value $\langle P \rangle$ as a function of $1/(2\kappa)$ for the three values of β . The β -dependence shows that the gap in the plaquette expectation value ΔP decreases substantially when moving from $\beta = 5.1$ ($a \approx 0.20$ fm) to $\beta = 5.3$ ($a \approx 0.14$ fm). One possible definition for the quantity ΔP is the difference between low and high phase plaquette expectation value at the smallest value of κ where a meta-stability occurs. Moreover, one can see in figure 4.7 that the meta-stability region in $1/(2\kappa)$ gets narrower with increasing values of β . Other quantities than ΔP show a similar behavior. Also the gap in m_χ^{PCAC} between positive and negative values shrinks significantly with increasing values of β .

We remark that the first order phase transition exists also in the continuum limit where it occurs as the jump of the scalar condensate as a consequence of spontaneous chiral symmetry breaking. Of course, in the continuum limit, the phase transition occurs only for $\mu = 0$ and the jump in $\langle P \rangle$ will disappear.

An interesting practical question is, at what value of the lattice spacing a the minimal pseudo scalar mass $m_{\text{PS}}^{\text{min}}$ that can be simulated without meta-stability as-

β	r_0/a	$r_0\mu$	ΔP	$m_{\text{PS}}^{\text{min}} [\text{MeV}]$
5.1	2.497(29)	0.0327	0.0399(1)	$\gtrsim 600$
5.2	3.124(85)	0.0312	0.0261(1)	$\gtrsim 630$
5.3	3.628(60)	0.0290	0.0077(4)	$\gtrsim 470$

Table 4.2: Reference values for r_0/a together with $r_0\mu$, the plaquette gap ΔP and $m_{\text{PS}}^{\text{min}}$ for the three β values.

sumes a value of, say, 300 MeV. At such a value contact to χ PT could be established. Unfortunately, the precise determination of the meta-stability region in κ and of a minimal pseudo scalar mass is very difficult. We can estimate the meta-stability region in κ from our data by including all κ values where meta-stabilities occur. A lower bound for the minimal pseudo scalar mass in the low plaquette phase (high plaquette phase) is then represented by the value of m_{PS} at the lower (higher) end of this κ interval. In the following we will mainly focus on the low plaquette phase since this is the natural choice for studying lattice QCD.

We give in table 4.2 estimates for the minimal pseudo scalar masses in the low plaquette phase in physical units. In addition, we provide estimates for ΔP . In principle, it would be the natural next step to extrapolate the minimal pseudo scalar mass and ΔP as a function of the lattice spacing. However, our present data do not allow for a reliable and safe extrapolation. First of all, the determination of the minimal pseudo scalar mass has a large ambiguity in itself since we do not know for sure, which simulation point is stable or meta-stable. Second, the only three values of β we have used give a too short lever arm to perform a trustworthy extrapolation. And, last, the values of r_0/a are very different in the two phases, which makes it particularly difficult to follow ΔP as a function of the lattice spacing, since ΔP contains information from both phases.

Nevertheless, an estimate on a more qualitative level yields a value of the lattice spacing of $a \sim 0.07 \text{ fm} - 0.1 \text{ fm}$ where simulations with pseudo scalar masses of about 300 MeV can be performed without being affected by the first order phase transition*.

At this point we can complete the picture of the phase structure of lattice QCD with Wilson type quarks. It is schematically plotted in figure 4.8 in the β - μ - κ -space on the basis of the predictions of $L\chi$ PT and the numerical findings as presented in this section and in the literature. For values of β smaller than about five it was found that the Aoki scenario is realized [144, 145]. The phase is located around the critical value of κ , its width in κ diminishes with increasing values of β and, as

*We remark that we have indications for meta-stabilities even at the parameters of run D from the last chapter, i.e. $a \sim 0.08 \text{ fm}$, $\mu = 0$ and $m_{\text{PS}} \sim 300 \text{ MeV}$. However, this might turn out to be a thermalization phenomenon.

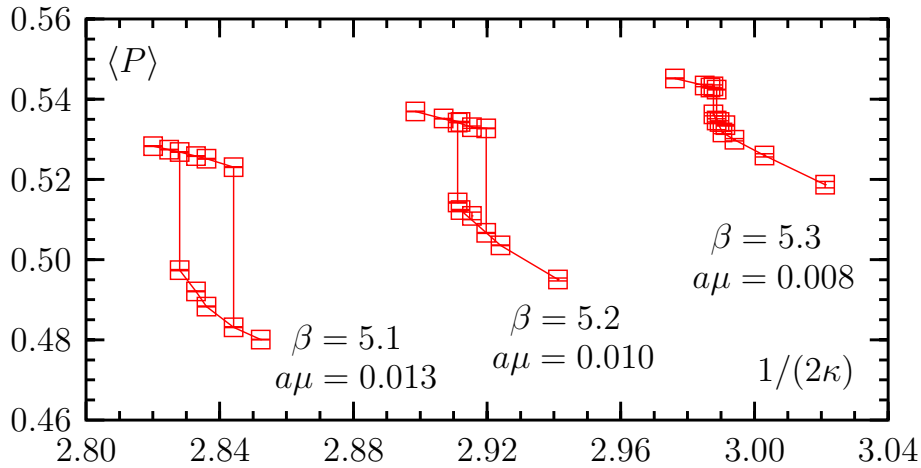


Figure 4.7: Evolution of the gap in the plaquette expectation value for the three values of β .

discussed in section 4.1, the Aoki phase disappears for non-vanishing values of μ . The resulting area is plotted in dark-gray in figure 4.8. On the boundaries of this area a second order phase transition takes place.

Since the first order phase transition in the normal scenario takes place for κ equal to its critical value and values of $|\mu|$ smaller than the critical value μ_c , the corresponding area of first order phase transition points is orthogonal to the Aoki phase area. It is plotted in light-gray in figure 4.8. As discussed in section 4.1 the value of μ_c goes to zero as a^2 when the continuum limit is approached. Therefore the width of the first order phase transition area gets correspondingly smaller. On the boundary of this area are second order endpoints located.

In the region of β values between the two scenarios there are no data available describing how the crossover exactly looks like. The design of this crossover as it is plotted in figure 4.8 is a guess under the assumption that the two scenarios do not exist in parallel for the same set of parameters and that the crossover is smooth in β . However, this brings the discussion to the question how reliable predictions of $L\chi$ PT are, when the lattice spacing is larger than 0.15 fm. If $L\chi$ PT to order a^2 would explain the phase structure for the whole above mentioned range of lattice spacings, the sign of c_2 must change as a function of β .

We cannot answer this question, even though we think that it is very likely that higher order lattice artefacts contribute significantly to the phase structure if the lattice spacing is large. These higher order effects could then also avoid the necessity of a sign change in c_2 , because the crossover can then be explained by higher order lattice artifacts.

Finally, let us discuss the implications of the observed phase structure on simulations in lattice QCD with Wilson like fermions. First of all the understanding of

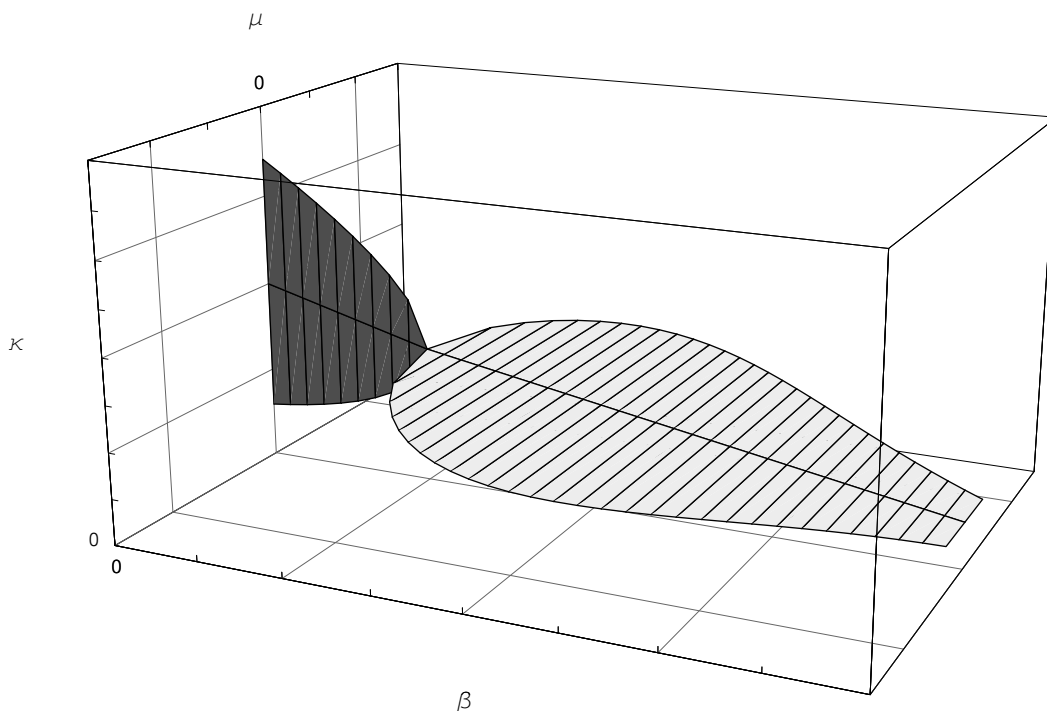


Figure 4.8: Schematic phase-diagram of Wilson twisted mass lattice QCD. The light-gray surface is the area where a first order phase transition takes place ($c_2 < 0$) when κ is crossing its critical value. At the boundary of this area a second order endpoint is located. The dark-gray area is the parameter region where the Aoki scenario is realized ($c_2 > 0$). At the border line of this area a second order phase transition takes place. At the value of β where the two areas touch each other, c_2 is supposed to be identically zero.

the phase structure is (or should be) an important pre-requisite for any large scale simulation in lattice QCD, a piece of information that was missing so far.

Unfortunately, from an only practical point of view the actual phase structure makes simulations with Wilson gauge action and Wilson like fermions difficult, if not unfeasible. The reason is the following: due to the meta-stability phenomenon simulations with pseudo scalar masses of the order of 300 MeV must be performed with lattice spacings of $a \lesssim 0.1$ fm and correspondingly large volumes in lattice units. To perform then a reliable continuum extrapolation simulations with at least two even smaller lattice spacings with the same physical volumes are needed. All together large scale simulations in this setup become rather demanding, and are therefore not realistic.

Note that in the two dimensional Gross-Neveu model [152] it is possible to compute the phase structure analytically. Depending on the parameters the outcome is very similar to the phase structure as observed for Wilson lattice QCD [153, 154, 155, 156].

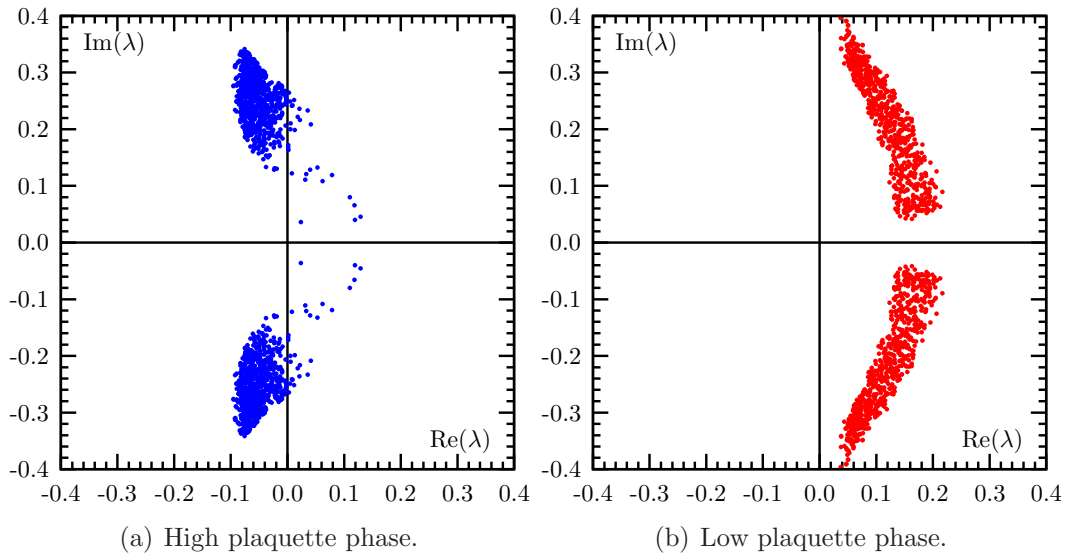


Figure 4.9: Eigenvalues λ of the Wilson twisted mass fermion matrix with small absolute value in case of the Wilson plaquette action at $\beta = 5.2$, $\mu = 0.01$, $\kappa = 0.1715$ on a $12^3 \times 24$ lattice. Both high plaquette and low plaquette spectra are shown.

4.2.6 Changing the gauge action

As explained above we think that simulations with Wilson plaquette gauge action and Wilson like fermions are not feasible due to the existence of the first order phase transition. This is of course a rather un-satisfactory result and a cure is needed if one wants to stick to Wilson twisted mass fermions in order to use the automatic $\mathcal{O}(a)$ improvement property of mtmQCD.

If one compares the infra-red end of the eigenvalue spectrum of the twisted mass operator at identical parameters, but on gauge configurations from the high and the low plaquette phase separately as we show it in figure 4.9[†], one can see that the first order phase transition is visible as a strong difference in infra-red spectrum. Therefore, the phase structure might be influenced by changing the infra-red eigenvalue spectrum, which is well known to be possible by changing the discretization of the gauge action (see for instance [157, 158, 159, 160]). In fact, the JLQCD collaboration reported in Refs. [150, 161] for lattice QCD $N_f = 3$ non perturbatively improved flavors of quarks that the meta-stability phenomenon disappears when the Iwasaki gauge action [162, 163] or the tadpole improved Symanzik gauge action [13] is used instead of the Wilson plaquette gauge action.

In order to check the effect of changing the gauge action on the phase structure our collaboration performed simulations with two additional gauge actions: the DBW2 gauge action [164, 165] and the tree level Symanzik (tlSym) improved gauge

[†]In the figure one can nicely see the eigenvalue free strip around the real axis, which is due to the twisted mass term serving as an infra-red cut-off to the spectrum.

action [166, 167]. Both of these belong to a one-parameter family of gauge actions and include, besides the usual (1×1) Wilson plaquette term, planar rectangular (1×2) Wilson loops:

$$S_g = \sum_x \left(c_0 \sum_{1 \leq \mu < \nu; \mu, \nu=1}^4 \frac{1}{3} \{1 - \text{Re Tr}(U_\square)\} + c_1 \sum_{\mu \neq \nu; \mu, \nu=1}^4 \frac{1}{3} \{1 - \text{Re Tr}(U_{x, \mu, \nu}^{1 \times 2})\} \right), \quad (4-10)$$

with the normalization condition $c_0 = 1 - 8c_1$. (The notation c_0, c_1 is conventional. c_1 should not be confused with the parameter c_1 in the effective potential model). The coefficient c_1 in Eq. (4-10) takes different values for the two choices of gauge actions mentioned above:

$$c_1 = \begin{cases} -1.4088 & \text{DBW2 gauge action} \\ -1/12 & \text{tlSym gauge action .} \end{cases} \quad (4-11)$$

Clearly, $c_1 = 0$ corresponds to the original Wilson plaquette gauge action. Note that the value of $c_1 = -0.331$ corresponds to the Iwasaki gauge action [162, 163].

The c_1 value for the DBW2 gauge action was determined by renormalization group considerations. The value of $c_1 = -1/12$ corresponds to the value computed at tree level in order to improve the gauge action à la Symanzik. Thus, the tlSym gauge action is in between the DBW2 and the Wilson plaquette gauge action, even though the numerical value of c_1 suggests that it is closer to the Wilson gauge action.

Our collaboration obtained very promising results for the DBW2 gauge action, which are published in Ref. [168]. If we tune the parameters with the DBW2 gauge action such that the lattice spacings are comparable to the one measured with Wilson plaquette gauge action at $\beta = 5.2$, we find that the strength of the phase transition is significantly reduced. Moreover, for even smaller lattice spacings, in a situation now comparable to the one with Wilson plaquette gauge action at $\beta = 5.3$, we do not find evidence for a first order phase transition. For the tlSym gauge action the investigations are still ongoing, but the preliminary results are similar to the one observed with the DBW2 gauge action. We summarize the results in figure 4.10, where we show the plaquette expectation value as a function of κ for the DBW2, the tlSym and the Wilson plaquette gauge action at approximately the same value for $a = 0.2$ fm. For all of the three actions meta-stabilities are visible. The value of $a\mu$ was only for the Wilson plaquette gauge action different from zero, which should, however, decrease the effect for the plaquette gauge action.

Passing in figure 4.10 from the Wilson over the tlSym to the DBW2 gauge action the jump in the plaquette expectation value becomes clearly smaller. In addition,

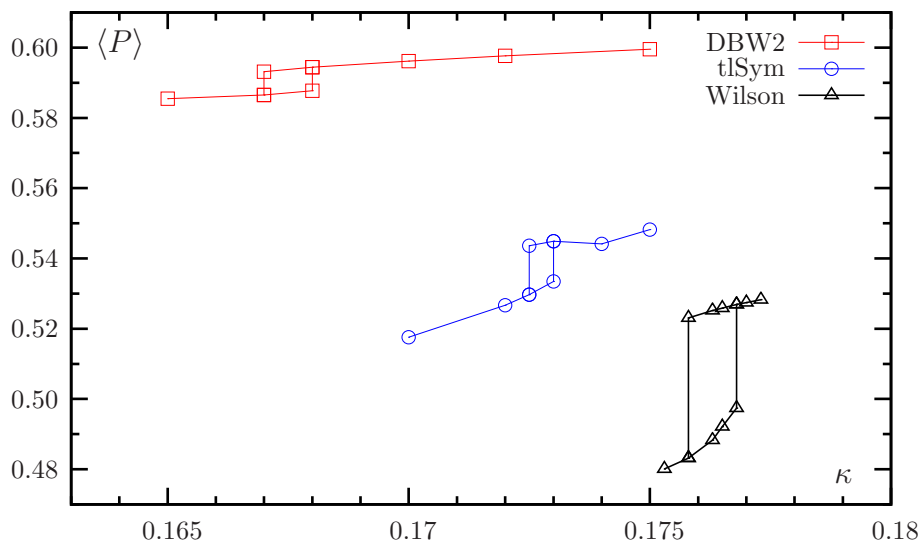


Figure 4.10: Plaquette expectation values as a function of κ for the three different gauge actions at approximately identical value of the lattice spacing $a = 0.2$ fm. For the DBW2 and the tISym gauge action the value of μ was identical zero, whereas for the Wilson gauge action we used $a\mu = 0.013$. The continuous lines only connect the points and are meant to guide the eyes. The meta-stable region is indicated by connecting high and low phase at the first and the last meta-stable point.

the difference between DBW2 and tISym is much smaller than the difference between tISym and Wilson plaquette. This outcome, if confirmed, is surprising, since one could expect the effect to “scale” with c_1 .

Nevertheless, these results make us confident that changing the gauge action is indeed a tool to weaken the effects of the first order phase transition, in agreement with earlier findings [150, 161]. With a different gauge action than the Wilson plaquette gauge action dynamical simulation with twisted mass fermions at maximal twist might then become possible with lattice spacings equal or lower than 0.16 fm and pseudo scalar masses small enough.

4.3 Conclusion

The main result of this chapter is that close enough to the continuum the phase structure in lattice theories with Wilson or Wilson twisted mass fermions is the expected continuum phase structure of QCD distorted by lattice artifacts.

In detail, we have explored the phase structure of lattice QCD with Wilson twisted mass fermions and the Wilson plaquette gauge action. We have investigated three values of the bare coupling $\beta = 5.1, 5.2, 5.3$ each with fixed value of $a\mu$. By changing the hopping parameter κ we encountered strong meta-stabilities for all three values of β , visible in long living meta-stable states with either a low or a high

plaquette expectation value. The PCAC quark mass m_χ^{PCAC} in the different meta-stable branches is positive for the branch with low plaquette expectation value and negative for the branch with high plaquette expectation value. At the same time, the pseudo scalar mass does not vanish at the chiral point but has a minimum at a rather large value, which is at $\beta = 5.3$ still about 500 MeV. We stress here that the aforementioned lower bound for the pseudo scalar mass does not originate from algorithmic or technical problems, but it is a physical property of the lattice theory. In fact, it would be interesting to investigate the phase structure of lattice QCD with different formulations, such as the staggered or even the overlap formulation.

This phenomenon finds a natural interpretation in the effective potential model from lattice chiral perturbation theory, both for pure Wilson fermions with $\mu = 0$ [41] and for Wilson twisted mass fermions with $\mu \neq 0$ [142, 97, 143]: the so-called normal scenario with a first order phase transition and a non-vanishing pseudo scalar mass at the chiral point is realized for the β value that we investigated.

We clearly observe that this first order phase transition weakens substantially, when β is increased. Unfortunately, we cannot quantitatively locate the value of the lattice spacing, where the effects of the first order phase transition becomes negligible and where a minimal pseudo scalar mass of, say, 300 MeV can be reached. As an estimate of such a value of the lattice spacing we give $a \approx 0.1$ fm. Of course, this would mean that a continuum extrapolation of physical results obtained on lattices with linear extent of at least $L = 2$ fm would be very demanding, since the starting point for such simulations would already require large lattices. It is therefore very important to find alternative actions such that the value of the lattice spacing can be lowered without running into problems with the first order phase transition.

With our results together with results available in the literature we were able to draw a schematic picture of the phase structure of lattice QCD with Wilson like fermions. While for values of β smaller than 5.0 there is evidence for the Aoki scenario, in the range of β values between 5.1 and 5.3 we find evidence for the normal scenario. The phase diagram is summarized in fig. 4.8.

Our collaboration also investigated the change of the observed phase structure with two different gauge actions to replace the Wilson plaquette gauge action. These are the DBW2 and the tree level Symanzik improved gauge actions, both of which belong to a one-parameter family of gauge actions. The results we obtain are very promising in a sense that it seems to become possible to reduce significantly the effects of the first order phase transition at lattice spacings comparable to the one used in this chapter (see Ref. [168] for details). This makes us confident that indeed with a different gauge action than the Wilson plaquette gauge action dynamical simulations with automatic $\mathcal{O}(a)$ improved Wilson twisted mass fermions are feasible.

Finally, we remark that in the twisted mass formulation also non-degenerate

quark flavors can be simulated without losing the property of automatic $\mathcal{O}(a)$ improvement [169, 46]. Together with the twisted mass parameter serving as an infra-red cut-off for the eigenvalue spectrum, the twisted mass formulation then becomes a promising candidate for large scale simulations with dynamical up-, down- and strange quark.

Summary and Outlook

In the year 2003 an intriguing paper was published: Frezzotti and Rossi realized that with the Wilson twisted mass formulation of lattice QCD at maximal twist $\mathcal{O}(a)$ improvement can be obtained without the need of additional improvement coefficients [42], which need to be computed non-perturbatively in the standard Symanzik improvement programme. Of course, such a theoretical proposition needs a check in practice and hence we performed a scaling test of maximally twisted mass QCD (mtmQCD), in the quenched approximation as a start. The results can be found in chapter 2 and the most important findings are the following:

- thanks to the twisted mass term playing the rôle of an infrared cut-off for the Dirac operator eigenvalue spectrum, it is possible to simulate pseudo scalar masses as low as 270 MeV without having problems with exceptional configurations.
- physical observables determined with mtmQCD show no $\mathcal{O}(a)$ lattice artifacts. With the PCAC definition of κ_{crit} also higher order lattice artifacts are small, even with scalar mass values as small as 270 MeV. (cf. figure 2.3 on page 43, figure 2.6 on page 46 and figure 2.7 on page 47)
- the effects of the explicit flavor symmetry breaking in mtmQCD are sizable when the charged/neutral pseudo scalar mass splitting is considered. Nevertheless, the splitting is a lattice artifact and we could show that it vanishes proportional to a^2 , see figure 2.8 on page 49.
- while simulations with pseudo scalar mass values below 300 MeV are also possible with overlap fermions, a comparison of computational costs revealed that simulations with twisted mass fermions are a factor of 20 to 70 faster than simulations with overlap fermions.

Aiming at large scale simulations in full lattice QCD we developed – in addition to the investigation of mtmQCD as the potential formulation – a new variant of the Hybrid Monte Carlo algorithm in order to make full QCD simulations with light quark masses affordable. The new variant, a HMC with a combination of mass

preconditioning and multiple time scale integration as presented in chapter 3, is based on the idea to precondition the fermion determinant in such a way that the most expensive part contributes the least to the total molecular dynamics force and can be therefore integrated on the largest time scale. It is applicable to a wide variety of lattice Dirac operators and moreover straightforward to implement.

Our simulations clearly show that with this new HMC variant (full dynamical) simulations with Wilson type fermions and realistic quark masses are possible with reasonable computational effort. This we could illustrate in an update of the so called “Berlin Wall” figure [105, 113]. With our HMC variant the “Wall” is shifted definitely towards smaller values of m_{PS}/m_V , see figure 3.4 on page 77.

Taking the results of these two chapters together, we have now a sound basis for performing large scale simulations with light quark masses. However, our first investigations in full lattice QCD revealed a surprising result: we have shown in chapter 4 that close enough to the continuum the phase structure of lattice QCD with Wilson type fermions and the Wilson plaquette gauge action exhibits the expected continuum phase structure distorted by lattice artifacts. Our investigation yielded that in a range of lattice spacings between 0.2 and 0.13 fm there exists a first order phase transition at the chiral point. This phase transition is characterized by a jump in the plaquette expectation value, the existence of meta-stabilities and most importantly by the fact that the minimal value of the charged pseudo scalar mass is significantly larger than zero. For the investigated lattice spacing the latter value lies well above 450 MeV.

This phenomenon finds its natural interpretation in terms of an effective potential model depicted in lattice chiral perturbation theory, where a first order phase transition is predicted as one of two possible scenarios emerging due to $\mathcal{O}(a^2)$ lattice artifacts. We stress here that the first order phase transition is a generic property of Wilson type fermions and it is not restricted to only the twisted mass formulation, which was mainly used for the simulations. The phase structure is summarized schematically in figure 4.8 on page 97.

This result on the one hand makes clear that indeed the knowledge of the phase structure is an essential pre-requisite before starting large scale simulations that was missing so far. On the other hand, the rather large minimal value of the charged pseudo scalar mass does not allow for simulations with realistic quark masses at affordable lattice spacings with mtmQCD and the Wilson plaquette gauge action. However, the size of the lattice artifacts and hence the size of for instance the minimal value of the pseudo scalar mass is certainly depending on the discretization of the gauge action. There were already hints for this in the literature and we could confirm that by adding a rectangular part to the Wilson plaquette gauge action the effects of the phase transition can be reduced when compared to the Wilson plaquette gauge action at equal lattice spacing. This means that for example with the DBW2 or the

tree level improved Symanzik gauge action the minimal pseudo scalar mass value is significantly reduced and we are optimistic that at lattice spacings of about 0.15 fm dynamical simulations can be performed without being affected by the first order phase transition.

Recapitulating, we think the results of this work exhibit a sound basis for future large scale dynamical simulations with realistic masses and small enough lattice spacing. Of course, the next natural step is to repeat a scaling test for twisted mass QCD with $n_f = 2$ dynamical flavors of quarks, as we presented it in chapter 2 for the quenched approximation. In such a scaling test already some interesting continuum extrapolations should become possible. Then, in case of a positive outcome, we will proceed to $n_f = 2 + 1 + 1$ dynamical quark flavors, with up- and down-quark mass degenerate and strange- and charm-quark with different, non-degenerate masses. At the latest in these simulations such interesting items like ρ -decay, string breaking, π phase shift, or η and η' meson masses will be addressed.

SUMMARY AND OUTLOOK

Appendix A

Conventions

A.1 Dirac matrices

The following convention for the the Euclidean γ matrices is used:

$$\begin{aligned} \gamma_0 &= \begin{pmatrix} 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 \\ +1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \end{pmatrix}, & \gamma_1 &= \begin{pmatrix} 0 & 0 & 0 & +i \\ 0 & 0 & +i & 0 \\ 0 & -i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \\ \gamma_2 &= \begin{pmatrix} 0 & 0 & 0 & +1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ +1 & 0 & 0 & 0 \end{pmatrix}, & \gamma_3 &= \begin{pmatrix} 0 & 0 & +i & 0 \\ 0 & 0 & 0 & -i \\ -i & 0 & 0 & 0 \\ 0 & +i & 0 & 0 \end{pmatrix}. \end{aligned} \tag{A-1}$$

They are hermitian and satisfy the anti-commutation relation

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}. \tag{A-2}$$

With the above choice for γ_0 we have chosen the chiral representation where $\gamma_5 = \gamma_1\gamma_2\gamma_3\gamma_0$ is diagonal:

$$\gamma_5 = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{A-3}$$

The projection operators on left- and right-handed chirality then read

$$P_- = \frac{1}{2}(1 - \gamma_5), \quad P_+ = \frac{1}{2}(1 + \gamma_5), \tag{A-4}$$

respectively.

Appendix B

Algorithmic details

B.1 Even/odd preconditioning

In this appendix we describe how even/odd [106, 170] preconditioning can be used in the HMC algorithm in presence of a twisted mass term. By setting the twisted mass parameter to zero, even/odd preconditioning for the Wilson-Dirac operator can easily be recovered from the formulae presented in the following.

We start with the lattice fermion action in the hopping parameter representation in the χ -basis written as

$$\begin{aligned} S[\chi, \bar{\chi}, U] &= \sum_x \left\{ \bar{\chi}(x) [1 + 2i\kappa\mu\gamma_5\tau^3] \chi(x) \right. \\ &\quad \left. - \kappa\bar{\chi}(x) \sum_{\mu=1}^4 \left[U(x, \mu) (r + \gamma_\mu) \chi(x + a\hat{\mu}) \right. \right. \\ &\quad \left. \left. + U^\dagger(x - a\hat{\mu}, \mu) (r - \gamma_\mu) \chi(x - a\hat{\mu}) \right] \right\} \\ &\equiv \sum_{x,y} \bar{\chi}(x) M_{xy} \chi(y) . \end{aligned} \tag{B-1}$$

similar to Eq. (1-46) in section 1.2.4. For convenience we define $\tilde{\mu} = 2\kappa\mu$. Using the matrix M one can define the hermitian (two flavor) operator.

$$Q \equiv \gamma_5 M = \begin{pmatrix} Q^+ & \\ & Q^- \end{pmatrix} \tag{B-2}$$

Appendix B Algorithmic details

where the sub-matrices Q^\pm can be factorized as follows:

$$\begin{aligned} Q^\pm &= \gamma_5 \begin{pmatrix} 1 \pm i\tilde{\mu}\gamma_5 & M_{eo} \\ M_{oe} & 1 \pm i\tilde{\mu}\gamma_5 \end{pmatrix} = \gamma_5 \begin{pmatrix} M_{ee}^\pm & M_{eo} \\ M_{oe} & M_{oo}^\pm \end{pmatrix} \\ &= \begin{pmatrix} \gamma_5 M_{ee}^\pm & 0 \\ \gamma_5 M_{oe} & 1 \end{pmatrix} \begin{pmatrix} 1 & (M_{ee}^\pm)^{-1} M_{eo} \\ 0 & \gamma_5 (M_{oo}^\pm - M_{oe} (M_{ee}^\pm)^{-1} M_{eo}) \end{pmatrix}. \end{aligned} \quad (\text{B-3})$$

Note that $(M_{ee}^\pm)^{-1}$ can be computed to be

$$(1 \pm i\tilde{\mu}\gamma_5)^{-1} = \frac{1 \mp i\tilde{\mu}\gamma_5}{1 + \tilde{\mu}^2}. \quad (\text{B-4})$$

Using $\det(Q) = \det(Q^+) \det(Q^-)$ the following relation can be derived

$$\begin{aligned} \det(Q^\pm) &\propto \det(\hat{Q}^\pm) \\ \hat{Q}^\pm &= \gamma_5 (M_{oo}^\pm - M_{oe} (M_{ee}^\pm)^{-1} M_{eo}), \end{aligned} \quad (\text{B-5})$$

where \hat{Q}^\pm is only defined on the odd sites of the lattice. In the HMC algorithm the determinant is stochastically estimated using pseudo fermion field ϕ_o : Now we write the determinant with pseudo fermion fields:

$$\begin{aligned} \det(\hat{Q}^+ \hat{Q}^-) &= \int \mathcal{D}\phi_o \mathcal{D}\phi_o^\dagger \exp(-S_{\text{PF}}) \\ S_{\text{PF}} &\equiv \phi_o^\dagger \left(\hat{Q}^+ \hat{Q}^- \right)^{-1} \phi_o, \end{aligned} \quad (\text{B-6})$$

where the fields ϕ_o are defined only on the odd sites of the lattice. In order to compute the force corresponding to the effective action S_{PF} we need the variation of S_{PF} with respect to the gauge fields (using $\delta(A^{-1}) = -A^{-1}\delta A A^{-1}$):

$$\begin{aligned} \delta S_{\text{PF}} &= -[\phi_o^\dagger (\hat{Q}^+ \hat{Q}^-)^{-1} \delta \hat{Q}^+ (\hat{Q}^+)^{-1} \phi_o + \phi_o^\dagger (\hat{Q}^-)^{-1} \delta \hat{Q}^- (\hat{Q}^+ \hat{Q}^-)^{-1} \phi_o] \\ &= -[X_o^\dagger \delta \hat{Q}^+ Y_o + Y_o^\dagger \delta \hat{Q}^- X_o] \end{aligned} \quad (\text{B-7})$$

with X_o and Y_o defined on the odd sides as

$$X_o = (\hat{Q}^+ \hat{Q}^-)^{-1} \phi_o, \quad Y_o = (\hat{Q}^+)^{-1} \phi_o = \hat{Q}^- X_o, \quad (\text{B-8})$$

where $(\hat{Q}^\pm)^\dagger = \hat{Q}^\mp$ has been used. The variation of \hat{Q}^\pm reads

$$\delta \hat{Q}^\pm = \gamma_5 \left(-\delta M_{oe} (M_{ee}^\pm)^{-1} M_{eo} - M_{oe} (M_{ee}^\pm)^{-1} \delta M_{eo} \right), \quad (\text{B-9})$$

and one finds

$$\begin{aligned} \delta S_{\text{PF}} &= -(X^\dagger \delta Q^+ Y + Y^\dagger \delta Q^- X) \\ &= -(X^\dagger \delta Q^+ Y + (X^\dagger \delta Q^+ Y)^\dagger) \end{aligned} \quad (\text{B-10})$$

where X and Y are now defined over the full lattice as

$$X = \begin{pmatrix} -(M_{ee}^-)^{-1}M_{eo}X_o \\ X_o \end{pmatrix}, \quad Y = \begin{pmatrix} -(M_{ee}^+)^{-1}M_{eo}Y_o \\ Y_o \end{pmatrix}. \quad (\text{B-11})$$

In addition $\delta Q^+ = \delta Q^-$, $M_{eo}^\dagger = \gamma_5 M_{oe} \gamma_5$ and $M_{oe}^\dagger = \gamma_5 M_{eo} \gamma_5$ has been used. Since the bosonic part is quadratic in the ϕ_o fields, the ϕ_o are generated at the beginning of each molecular dynamics trajectory with

$$\phi_o = \hat{Q}^+ R, \quad (\text{B-12})$$

where R is a random spinor field taken from a Gaussian distribution with norm one.

Inversion

In addition to even/odd preconditioning in the HMC algorithm as described above, it can also be used to speed up the inversion of the fermion matrix.

Due to the factorization (B-3) the full fermion matrix can be inverted by inverting the two matrices appearing in the factorization

$$\begin{pmatrix} M_{ee}^\pm & M_{eo} \\ M_{oe} & M_{oo}^\pm \end{pmatrix}^{-1} = \begin{pmatrix} 1 & (M_{ee}^\pm)^{-1}M_{eo} \\ 0 & (M_{oo}^\pm - M_{oe}(M_{ee}^\pm)^{-1}M_{eo}) \end{pmatrix}^{-1} \begin{pmatrix} M_{ee}^\pm & 0 \\ M_{oe} & 1 \end{pmatrix}^{-1}.$$

The two factors can be simplified as follows:

$$\begin{pmatrix} M_{ee}^\pm & 0 \\ M_{oe} & 1 \end{pmatrix}^{-1} = \begin{pmatrix} (M_{ee}^\pm)^{-1} & 0 \\ -M_{oe}(M_{ee}^\pm)^{-1} & 1 \end{pmatrix}$$

and

$$\begin{aligned} & \begin{pmatrix} 1 & (M_{ee}^\pm)^{-1}M_{eo} \\ 0 & (M_{oo}^\pm - M_{oe}(M_{ee}^\pm)^{-1}M_{eo}) \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & -(M_{ee}^\pm)^{-1}M_{eo}(M_{oo}^\pm - M_{oe}(M_{ee}^\pm)^{-1}M_{eo})^{-1} \\ 0 & (M_{oo}^\pm - M_{oe}(M_{ee}^\pm)^{-1}M_{eo})^{-1} \end{pmatrix}. \end{aligned}$$

The complete inversion is now performed in two separate steps: First we compute for a given source field $\phi = (\phi_e, \phi_o)$ an intermediate result $\varphi = (\varphi_e, \varphi_o)$ by:

$$\begin{pmatrix} \varphi_e \\ \varphi_o \end{pmatrix} = \begin{pmatrix} M_{ee}^\pm & 0 \\ M_{oe} & 1 \end{pmatrix}^{-1} \begin{pmatrix} \phi_e \\ \phi_o \end{pmatrix} = \begin{pmatrix} (M_{ee}^\pm)^{-1}\phi_e \\ -M_{oe}(M_{ee}^\pm)^{-1}\phi_e + \phi_o \end{pmatrix}.$$

This step requires only the application of M_{oe} and $(M_{ee}^\pm)^{-1}$, the latter of which is given by Eq (B-4). The final solution $\psi = (\psi_e, \psi_o)$ can then be computed with

$$\begin{pmatrix} \psi_e \\ \psi_o \end{pmatrix} = \begin{pmatrix} 1 & (M_{ee}^\pm)^{-1}M_{eo} \\ 0 & (M_{oo}^\pm - M_{oe}(M_{ee}^\pm)^{-1}M_{eo}) \end{pmatrix}^{-1} \begin{pmatrix} \varphi_e \\ \varphi_o \end{pmatrix} = \begin{pmatrix} \varphi_e - (M_{ee}^\pm)^{-1}M_{eo}\psi_o \\ \psi_o \end{pmatrix},$$

where we defined

$$\psi_o = (M_{oo}^\pm - M_{oe}(M_{ee}^\pm)^{-1}M_{eo})^{-1}\varphi_o.$$

Therefore the only inversion that has to be performed numerically is the one to generate ψ_o from φ_o and this inversion involves only an operator that is better conditioned than the original fermion operator.

B.2 Multiple mass solver for twisted mass fermions

In this appendix we show that within the Wilson twisted mass fermion formulation it is possible to apply the multi mass solver (MMS) [171, 172, 173] method to the conjugate gradient (CG) algorithm. We will call this algorithm CG-M and give here the details of the implementation.

The advantage of the MMS is that it allows the computation of the solution of the following linear system

$$(A + \sigma)x - b = 0 \tag{B-13}$$

for several values of σ simultaneously, using only as many matrix-vector operations as the solution of a single value of σ requires.

We want to invert the Wilson twisted mass operator at a certain value of the twisted mass μ_0 obtaining automatically all the solutions for other values μ_k (with $|\mu_k| \geq |\mu_0|$). We use the twisted mass operator D_{tm} as defined in Eq. (1-47) and denote the number of additional twisted mass values with N_m . The operator can be split up as

$$D_{\text{tm}} = D_{\text{tm}}^{(0)} + i(\mu_k - \mu_0)\gamma_5\tau^3, \quad D_{\text{tm}}^{(0)} = D_W + m_0 + i\mu_0\gamma_5\tau^3. \tag{B-14}$$

The trivial observation is that

$$D_{\text{tm}}D_{\text{tm}}^\dagger = D_{\text{tm}}^{(0)}D_{\text{tm}}^{(0)\dagger} + \mu_k^2 - \mu_0^2, \tag{B-15}$$

where we have used $\gamma_5 D_W \gamma_5 = D_W^\dagger$. Now clearly we have a shifted linear system $(A + \sigma_k)x - b = 0$ with $A = D_{\text{tm}}^{(0)}D_{\text{tm}}^{(0)\dagger}$ and $\sigma_k = \mu_k^2 - \mu_0^2$. In the following we describe the CG-M algorithm in order to solve the problem $(A + \sigma_k)x - b = 0$. The lower index indicates the iteration steps of the solver, while the upper index k refers to the shifted problem with σ_k .

B.2. MULTIPLE MASS SOLVER FOR TWISTED MASS FERMIONS

CG – M Algorithm

$$x_0^k = 0, r_0 = p_0^k = b, \alpha_{-1} = \zeta_{-1}^k = \zeta_0^k = 1, \beta_0^k = \beta_0 = 0$$

for $i = 0, 1, 2, \dots$

$$\alpha_n = \frac{(r_n, r_n)}{(p_n, Ap_n)}$$

$$\zeta_{n+1}^k = \frac{\zeta_n^k \alpha_{n-1}}{\alpha_n \beta_n (1 - \frac{\zeta_n^k}{\zeta_{n-1}^k}) + \alpha_{n-1} (1 - \sigma_k \alpha_n)}$$

$$\alpha_n^k = \alpha_n \frac{\zeta_{n+1}^k}{\zeta_n^k}$$

$$x_{n+1}^k = x_n^k + \alpha_n^k p_n^k$$

$$x_{n+1} = x_n + \alpha_n p_n$$

$$r_{n+1} = r_n - \alpha_n A p_n$$

convergence check

$$\beta_{n+1} = \frac{(r_{n+1}, r_{n+1})}{(r_n, r_n)}$$

$$p_{n+1} = r_{n+1} + \beta_{n+1} p_n$$

$$\beta_{n+1}^k = \beta_{n+1} \frac{\zeta_{n+1}^k \alpha_n^k}{\zeta_n^k \alpha_n}$$

$$p_{n+1}^k = \zeta_{n+1}^k r_{n+1} + \beta_{n+1}^k p_n^k$$

end for

We give here the algorithm explicitly again, since it has a different definition of ζ_{n+1}^k compared to the one of Ref. [173]. This version allows to avoid roundoff errors when $\sigma_k = \mu_k^2 - \mu_0^2$ becomes too large.

We remind that when using a MMS the eventual preconditioning has to retain the shifted structure of the linear system. This means for example that it is not compatible with even/odd preconditioning.

Appendix C

Tables

β	5.7	5.85	6.0	6.1	6.2	6.45
$a\mu_1$	0.2455(23)	0.1682(26)	0.1385(66)	0.1129(41)	0.1004(27)	0.0720(28)
$a\mu_2$	0.3237(16)	0.2256(22)	0.1764(42)	0.1482(27)	0.1298(23)	0.0914(27)
$a\mu_3$	0.4434(11)	0.3122(19)	0.2373(32)	0.2030(21)	0.1768(17)	—
$a\mu_4$	0.6272(9)	0.4452(14)	0.3335(22)	0.2865(15)	0.2463(15)	—
$a\mu_5$	0.7767(9)	0.5535(12)	0.4134(17)	0.3534(13)	0.3037(13)	—
$a\mu_6$	0.9074(9)	0.6488(13)	0.4839(16)	0.4130(13)	0.3546(12)	—
$a\mu_7$	1.0255(8)	0.7358(12)	0.5491(14)	0.4676(12)	0.4021(11)	—

Table C.1: Values of am_{PS} with pion definition of κ_{crit} .

β	5.7	5.85	6.0	6.1	6.2	6.45
$a\mu_1$	0.0986(10)	0.0782(13)	0.0516(17)	0.0466(14)	0.0437(13)	0.0329(13)
$a\mu_2$	0.1195(10)	0.0890(12)	0.0632(11)	0.0546(09)	0.0500(11)	0.0361(11)
$a\mu_3$	0.1418(11)	0.1003(12)	0.0740(09)	0.0623(08)	0.0562(10)	—
$a\mu_4$	0.1685(11)	0.1149(12)	0.0859(09)	0.0716(08)	0.0637(10)	—
$a\mu_5$	0.1902(11)	0.1273(13)	0.0949(09)	0.0790(08)	0.0698(09)	—
$a\mu_6$	0.2112(12)	0.1390(14)	0.1029(10)	0.0858(08)	0.0754(09)	—
$a\mu_7$	0.2320(13)	0.1501(14)	0.1104(10)	0.0919(09)	0.0806(09)	—

Table C.2: Values of af_{PS} with pion definition of κ_{crit} .

β	5.7	5.85	6.0	6.2
$a\mu_1$	0.2323(18)	0.1640(23)	0.1217(66)	0.0934(24)
$a\mu_2$	0.3245(15)	0.2289(17)	0.1708(50)	0.1276(21)
$a\mu_3$	0.4598(12)	0.3232(13)	0.2396(33)	0.1779(18)
$a\mu_4$	0.6564(10)	0.4606(11)	0.3403(22)	0.2492(13)
$a\mu_5$	0.8114(10)	0.5701(10)	0.4214(17)	0.3071(12)
$a\mu_6$	0.9451(10)	0.6658(09)	0.4925(14)	0.3588(10)
$a\mu_7$	1.0647(09)	0.7530(09)	0.5579(14)	0.4062(09)
$a\mu_8$	0.3892(14)	0.2741(15)	0.2038(40)	0.1519(20)
$a\mu_9$	0.5678(11)	0.3984(12)	0.2948(26)	0.2160(16)

Table C.3: Values of am_{PS} with PCAC definition of κ_{crit} .

β	5.7	5.85	6.0	6.2
$a\mu_1$	0.1267(14)	0.0894(14)	0.0689(27)	0.0512(16)
$a\mu_2$	0.1345(13)	0.0947(13)	0.0711(13)	0.0532(13)
$a\mu_3$	0.1472(12)	0.1025(12)	0.0763(10)	0.0567(10)
$a\mu_4$	0.1697(12)	0.1159(12)	0.0858(10)	0.0633(08)
$a\mu_5$	0.1914(13)	0.1284(11)	0.0944(10)	0.0694(08)
$a\mu_6$	0.2134(14)	0.1402(11)	0.1025(10)	0.0751(08)
$a\mu_7$	0.2358(15)	0.1518(11)	0.1100(10)	0.0803(08)
$a\mu_8$	0.1403(13)	0.0983(12)	0.0734(11)	0.0548(11)
$a\mu_9$	0.1589(12)	0.1095(12)	0.0813(10)	0.0601(09)

Table C.4: Values of af_{PS} with PCAC definition $\kappa_{\text{crit}}^{\text{PCAC}}$.

β	5.7	5.85	6.0	6.2
$a\mu_1$	0.716(67)	0.589(43)	0.458(30)	0.306(26)
$a\mu_2$	0.773(36)	0.591(19)	0.451(20)	0.317(20)
$a\mu_3$	0.854(20)	0.628(09)	0.467(12)	0.339(11)
$a\mu_4$	0.973(15)	0.701(05)	0.517(07)	0.378(05)
$a\mu_5$	1.076(11)	0.765(04)	0.560(06)	0.415(03)
$a\mu_6$	1.178(09)	0.834(03)	0.614(04)	0.452(02)
$a\mu_7$	1.277(08)	0.902(03)	0.666(03)	0.488(02)
$a\mu_8$	0.812(26)	0.606(13)	0.464(14)	0.327(16)
$a\mu_9$	0.919(14)	0.666(06)	0.494(08)	0.359(07)

Table C.5: Vector meson masses am_V for all simulation points with the PCAC definition $\kappa_{\text{crit}}^{\text{PCAC}}$.

m_{PS} [GeV]	f_{PS} [GeV]	m_{V} [GeV]
0.272	0.1500(66)	0.904(102)
0.372	0.1538(45)	0.937(053)
0.432	0.1572(40)	0.955(042)
0.514	0.1631(36)	0.978(033)
0.624	0.1724(33)	1.027(026)
0.728	0.1823(31)	1.083(019)
0.900	0.2002(29)	1.198(012)
1.051	0.2161(28)	1.313(009)
1.177	0.2283(28)	1.410(007)

Table C.6: f_{PS} and m_{V} in the continuum (only data from PCAC definition $\kappa_{\text{crit}}^{\text{PCAC}}$).

Bibliography

- [1] H. Weyl. *Electron and gravitation*. Z. Phys. 56, 330–352 (1929).
- [2] F. Jegerlehner. *Theoretical precision in estimates of the hadronic contributions to $(g-2)_\mu$ and $\alpha(\text{QED})(M(Z))$* . Nucl. Phys. Proc. Suppl. 126, 325–334 (2004), hep-ph/0310234.
- [3] D. J. Gross and F. Wilczek. *Ultraviolet behavior of non-Abelian gauge theories*. Phys. Rev. Lett. 30, 1343–1346 (1973).
- [4] S. Weinberg. *Nonabelian gauge theories of the strong interactions*. Phys. Rev. Lett. 31, 494–497 (1973).
- [5] H. Fritzsch, M. Gell-Mann, and H. Leutwyler. *Advantages of the color octet gluon picture*. Phys. Lett. B47, 365–368 (1973).
- [6] D. J. Gross and F. Wilczek. *Asymptotically free gauge theories. 1*. Phys. Rev. D8, 3633–3652 (1973).
- [7] H. D. Politzer. *Reliable perturbative results for strong interactions?* Phys. Rev. Lett. 30, 1346–1349 (1973).
- [8] H. D. Politzer. *Asymptotic freedom: an approach to strong interactions*. Phys. Rept. 14, 129–180 (1974).
- [9] G. 't Hooft. Unpublished remarks at the 1972 Marseille Conference on Yang-Mills Fields.
- [10] K. G. Wilson. *Confinement of quarks*. Phys. Rev. D10, 2445–2459 (1974).
- [11] K. G. Wilson. *The renormalization group: Critical phenomena and the kondo problem*. Rev. Mod. Phys. 47, 773 (1975).
- [12] K. Symanzik. Some topics in quantum field theory. In R. Schrader et al., editor, *Mathematical problems in theoretical physics*, volume 153, pages 47–58, 1981. Presented at 6th Int. Conf. on Mathematical Physics, Berlin, West Germany.

BIBLIOGRAPHY

- [13] K. Symanzik. *Continuum limit and improved action in lattice theories. 1. principles and ϕ^4 theory*. Nucl. Phys. B226, 187 (1983).
- [14] K. Symanzik. *Continuum limit and improved action in lattice theories. 2. $O(N)$ nonlinear sigma model in perturbation theory*. Nucl. Phys. B226, 205 (1983).
- [15] M. E. Peskin and D. V. Schroeder. *An Introduction to quantum field theory*. Westview Press, 1995.
- [16] S. Weinberg. *The Quantum theory of fields. Vol. 1: Foundations*. Cambridge University Press, 1995.
- [17] T. P. Cheng and L. F. Li. *Gauge theory of elementary particle physics: Problems and solutions*. Oxford, UK: Clarendon, 2000.
- [18] I. Montvay and G. Münster. *Quantum fields on a lattice*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1994.
- [19] H.J. Rothe. *Lattice gauge theories*. World Scientific, Singapore, 1992.
- [20] R. Gupta. Introduction to lattice QCD. Lectures given at Les Houches Summer School in Theoretical Physics, Session 68, 1997.
- [21] R. P. Feynman. *Space-time approach to non-relativistic quantum mechanics*. Rev. Mod. Phys. 20, 367–387 (1948).
- [22] K. Osterwalder and R. Schrader. *Axioms for euclidean Green's functions*. Commun. Math. Phys. 31, 83–112 (1973).
- [23] S. Eidelman et al. *Review of Particle Physics*. Physics Letters B 592, 1+ (2004).
- [24] H. B. Nielsen and M. Ninomiya. *No go theorem for regularizing chiral fermions*. Phys. Lett. B105, 219 (1981).
- [25] H. B. Nielsen and M. Ninomiya. *Absence of neutrinos on a lattice. 1. proof by homotopy theory*. Nucl. Phys. B185, 20 (1981).
- [26] H. B. Nielsen and M. Ninomiya. *Absence of neutrinos on a lattice. 2. intuitive topological proof*. Nucl. Phys. B193, 173 (1981).
- [27] L. H. Karsten and J. Smit. *Lattice fermions: species doubling, chiral invariance, and the triangle anomaly*. Nucl. Phys. B183, 103 (1981).
- [28] P. H. Ginsparg and K. G. Wilson. *A remnant of chiral symmetry on the lattice*. Phys. Rev. D25, 2649 (1982).

-
- [29] M. Lüscher. *Exact chiral symmetry on the lattice and the Ginsparg- Wilson relation*. Phys. Lett. B428, 342–345 (1998), [hep-lat/9802011](#).
- [30] M. F. Atiyah and I. M. Singer. *The Index of elliptic operators. 5*. Annals Math. 93, 139–149 (1971).
- [31] P. Hasenfratz, V. Laliena, and F. Niedermayer. *The index theorem in QCD with a finite cut-off*. Phys. Lett. B427, 125–131 (1998), [hep-lat/9801021](#).
- [32] H. Neuberger. *Exactly massless quarks on the lattice*. Phys. Lett. B417, 141–144 (1998), [hep-lat/9707022](#).
- [33] H. Neuberger. *More about exactly massless quarks on the lattice*. Phys. Lett. B427, 353–355 (1998), [hep-lat/9801031](#).
- [34] F. Niedermayer. *Exact chiral symmetry, topological charge and related topics*. Nucl. Phys. Proc. Suppl. 73, 105–119 (1999), [hep-lat/9810026](#).
- [35] P. Hernandez, K. Jansen, and M. Lüscher. *Locality properties of Neuberger’s lattice Dirac operator*. Nucl. Phys. B552, 363–378 (1999), [hep-lat/9808010](#).
- [36] D. B. Kaplan. *A Method for simulating chiral fermions on the lattice*. Phys. Lett. B288, 342–347 (1992), [hep-lat/9206013](#).
- [37] V. Furman and Y. Shamir. *Axial symmetries in lattice QCD with Kaplan fermions*. Nucl. Phys. B439, 54–78 (1995), [hep-lat/9405004](#).
- [38] K. Jansen. *Domain wall fermions and chiral gauge theories*. Phys. Rept. 273, 1–54 (1996), [hep-lat/9410018](#).
- [39] M. Lüscher, S. Sint, R. Sommer, and P. Weisz. *Chiral symmetry and $O(a)$ improvement in lattice QCD*. Nucl. Phys. B478, 365–400 (1996), [hep-lat/9605038](#).
- [40] B. Sheikholeslami and R. Wohlert. *Improved continuum limit lattice action for qcd with Wilson fermions*. Nucl. Phys. B259, 572 (1985).
- [41] S. R. Sharpe and Jr. Singleton, R. *Spontaneous flavor and parity breaking with Wilson fermions*. Phys. Rev. D58, 074501 (1998), [hep-lat/9804028](#).
- [42] R. Frezzotti and G. C. Rossi. *Chirally improving Wilson fermions. I: $O(a)$ improvement*. JHEP 08, 007 (2004), [hep-lat/0306014](#).
- [43] R. Frezzotti, P. A. Grassi, S. Sint, and P. Weisz. *A local formulation of lattice QCD without unphysical fermion zero modes*. Nucl. Phys. Proc. Suppl. 83, 941–946 (2000), [hep-lat/9909003](#).

BIBLIOGRAPHY

- [44] R. Frezzotti, P. A. Grassi, S. Sint, and P. Weisz. *Lattice QCD with a chirally twisted mass term*. JHEP 08, 058 (2001), [hep-lat/0101001](#).
- [45] R. Frezzotti, G. Martinelli, M. Papinutto, and G. C. Rossi. *Reducing cutoff effects in maximally twisted lattice QCD close to the chiral limit*. (2005), [hep-lat/0503034](#).
- [46] R. Frezzotti and G. C. Rossi. *Chirally improving Wilson fermions. II: Four-quark operators*. JHEP 10, 070 (2004), [hep-lat/0407002](#).
- [47] C. Michael. *Particle decay in lattice gauge theory*. Nucl. Phys. B327, 515 (1989).
- [48] T. A. DeGrand. *Resonance masses from Monte Carlo simulations (with emphasis on the rho meson)*. Phys. Rev. D43, 2296–2300 (1991).
- [49] R. M. Baxter et al. *Quenched heavy light decay constants*. Phys. Rev. D49, 1594–1605 (1994), [hep-lat/9308020](#).
- [50] R. Frezzotti and S. Sint. *Some remarks on $O(a)$ improved twisted mass QCD*. Nucl. Phys. Proc. Suppl. 106, 814–816 (2002), [hep-lat/0110140](#).
- [51] M. Della Morte, R. Frezzotti, and J. Heitger. *Quenched twisted mass QCD at small quark masses and in large volume*. Nucl. Phys. Proc. Suppl. 106, 260–262 (2002), [hep-lat/0110166](#).
- [52] K. Jansen, A. Shindler, C. Urbach, and I. Wetzorke. *Scaling test for Wilson twisted mass QCD*. Phys. Lett. B586, 432–438 (2004), [hep-lat/0312013](#).
- [53] T. Blum et al. *Quenched lattice QCD with domain wall fermions and the chiral limit*. Phys. Rev. D69, 074502 (2004), [hep-lat/0007038](#).
- [54] S. J. Dong et al. *Chiral properties of pseudoscalar mesons on a quenched $20^* \times 4$ lattice with overlap fermions*. Phys. Rev. D65, 054507 (2002), [hep-lat/0108020](#).
- [55] C. Gattringer et al. *Quenched spectroscopy with fixed-point and chirally improved fermions*. Nucl. Phys. B677, 3–51 (2004), [hep-lat/0307013](#).
- [56] L. Giusti, C. Hoelbling, and C. Rebbi. *Light quark masses with overlap fermions in quenched QCD*. Phys. Rev. D64, 114508 (2001), [hep-lat/0108007](#). Erratum-ibid.D65:079903,2002.
- [57] N. Garron, L. Giusti, C. Hoelbling, L. Lellouch, and C. Rebbi. *$B(K)$ from quenched QCD with exact chiral symmetry*. Phys. Rev. Lett. 92, 042001 (2004), [hep-ph/0306295](#).

- [58] Y. Chen et al. *Chiral logarithms in quenched QCD*. Phys. Rev. D70, 034502 (2004), [hep-lat/0304005](#).
- [59] R. Sommer. *A New way to set the energy scale in lattice gauge theories and its applications to the static force and alpha-s in SU(2) Yang-Mills theory*. Nucl. Phys. B411, 839–854 (1994), [hep-lat/9310022](#).
- [60] M. Albanese et al. *Glueball masses and string tension in lattice QCD*. Phys. Lett. B192, 163 (1987).
- [61] M. Lüscher and U. Wolff. *How to calculate the elastic scattering matrix in two- dimensional quantum field theories by numerical simulation*. Nucl. Phys. B339, 222–252 (1990).
- [62] E. Eichten and B. Hill. *An effective field theory for the calculation of matrix elements involving heavy quarks*. Phys. Lett. B234, 511 (1990).
- [63] A. Hasenfratz, R. Hoffmann, and F. Knechtli. *The static potential with hypercubic blocking*. Nucl. Phys. Proc. Suppl. 106, 418–420 (2002), [hep-lat/0110168](#).
- [64] M. Della Morte et al. *Lattice HQET with exponentially improved statistical precision*. Phys. Lett. B581, 93–98 (2004), [hep-lat/0307021](#).
- [65] M. Della Morte et al. *Static quarks with improved statistical precision*. Nucl. Phys. Proc. Suppl. 129, 346–348 (2004), [hep-lat/0309080](#).
- [66] M. Della Morte, A. Shindler, and R. Sommer. *On lattice actions for static quarks*. (2005), [hep-lat/0506008](#).
- [67] A. Hasenfratz and F. Knechtli. *Flavor symmetry and the static potential with hypercubic blocking*. Phys. Rev. D64, 034504 (2001), [hep-lat/0103029](#).
- [68] G. S. Bali et al. *Static potentials and glueball masses from QCD simulations with Wilson sea quarks*. Phys. Rev. D62, 054503 (2000), [hep-lat/0003012](#).
- [69] G. S. Bali, H. Neff, T. Duessel, T. Lippert, and K. Schilling. *Observation of string breaking in QCD*. Phys. Rev. D71, 114513 (2005), [hep-lat/0505012](#).
- [70] G. S. Bali et al. *String breaking with dynamical Wilson fermions*. Nucl. Phys. Proc. Supl. 140, 609–611 (2004), [hep-lat/0409137](#).
- [71] C.-N. Yang and R. L. Mills. *Conservation of isotopic spin and isotopic gauge invariance*. Phys. Rev. 96, 191–195 (1954).

BIBLIOGRAPHY

- [72] M. Guagnelli, R. Sommer, and H. Wittig. *Precision computation of a low-energy reference scale in quenched lattice QCD*. Nucl. Phys. B535, 389–402 (1998), [hep-lat/9806005](#).
- [73] S. Necco and R. Sommer. *The $N(f) = 0$ heavy quark potential from short to intermediate distances*. Nucl. Phys. B622, 328–346 (2002), [hep-lat/0108008](#).
- [74] M. Lüscher. *A New approach to the problem of dynamical quarks in numerical simulations of lattice QCD*. Nucl. Phys. B418, 637–648 (1994), [hep-lat/9311007](#).
- [75] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth. *Hybrid monte carlo*. Phys. Lett. B195, 216–222 (1987).
- [76] I. Montvay. *An Algorithm for Gluinos on the Lattice*. Nucl. Phys. B466, 259–284 (1996), [hep-lat/9510042](#).
- [77] N. Madras and A. D. Sokal. *The Pivot algorithm: a highly efficient Monte Carlo method for selfavoiding walk*. J. Statist. Phys. 50, 109–186 (1988).
- [78] U. Wolff. *Monte Carlo errors with less errors*. Comput. Phys. Commun. 156, 143–153 (2004), [hep-lat/0306017](#).
- [79] J. Garden, J. Heitger, R. Sommer, and Wittig H. *Precision computation of the strange quark’s mass in quenched QCD*. Nucl. Phys. B571, 237–256 (2000), [hep-lat/9906013](#).
- [80] D. Pleiter. XXX. PhD thesis, Freie Universitt Berlin, 2001.
- [81] S. Aoki and O. Bär. *Twisted-mass QCD, $O(a)$ improvement and Wilson chiral perturbation theory*. Phys. Rev. D70, 116011 (2004), [hep-lat/0409006](#).
- [82] S. R. Sharpe and Jackson M. S. Wu. *Twisted mass chiral perturbation theory at next-to-leading order*. Phys. Rev. D71, 074501 (2005), [hep-lat/0411021](#).
- [83] S. R. Sharpe. *Observations on discretization errors in twisted-mass lattice QCD*. (2005), [hep-lat/0509009](#).
- [84] S. Aoki and O. Bär. *Determining the low energy parameters of Wilson chiral perturbation theory*. (2005), [hep-lat/0509002](#).
- [85] *The ROOT system home page*. [root.cern.ch/](#).
- [86] *MINUIT home page*. [seal.web.cern.ch/seal/snapshot/work-packages/mathlibs/minuit/home.html](#).

- [87] C. R. Allton et al. *Gauge invariant smearing and matrix correlators using Wilson fermions at Beta = 6.2*. Phys. Rev. D47, 5128–5137 (1993), [hep-lat/9303009](#).
- [88] M. Guagnelli et al. *Finite size effects of a pion matrix element*. Phys. Lett. B597, 216–221 (2004), [hep-lat/0403009](#).
- [89] M. Gell-Mann, R. J. Oakes, and B. Renner. *Behavior of current divergences under SU(3) x SU(3)*. Phys. Rev. 175, 2195–2199 (1968).
- [90] L. Giusti, F. Rapuano, M. Talevi, and A. Vladikas. *The QCD chiral condensate from the lattice*. Nucl. Phys. B538, 249–277 (1999), [hep-lat/9807014](#).
- [91] K. Jansen et al. in preparation, 2005.
- [92] J. Heitger, R. Sommer, and H. Wittig. *Effective chiral Lagrangians and lattice QCD*. Nucl. Phys. B588, 377–399 (2000), [hep-lat/0006026](#). and references therein.
- [93] M. Foster and C. Michael. *Quark mass dependence of hadron masses from lattice QCD*. Phys. Rev. D59, 074503 (1999), [hep-lat/9810021](#).
- [94] C. McNeile and C. Michael. *The eta and eta' mesons in QCD*. Phys. Lett. B491, 123–129 (2000), [hep-lat/0006020](#).
- [95] K. Osterwalder and E. Seiler. *Gauge field theories on the lattice*. Ann. Phys. 110, 440 (1978).
- [96] K. Jansen et al. *Flavour breaking effects of Wilson twisted mass fermions*. (2005), [hep-lat/0507032](#). Accepted for publication in Phys. Lett. B.
- [97] L. Scorzato. *Pion mass splitting and phase structure in twisted mass QCD*. Eur. Phys. J. C37, 445–455 (2004), [hep-lat/0407023](#).
- [98] J. Wennekers and H. Wittig. *On the renormalized scalar density in quenched QCD*. (2005), [hep-lat/0507026](#).
- [99] P. Hernandez, K. Jansen, and L. Lellouch. *A numerical treatment of Neuberger's lattice Dirac operator*. (2000), [hep-lat/0001008](#).
- [100] P. Hernandez, K. Jansen, L. Lellouch, and H. Wittig. *Non-perturbative renormalization of the quark condensate in Ginsparg-Wilson regularizations*. JHEP 07, 018 (2001), [hep-lat/0106011](#).
- [101] P. Hernandez, K. Jansen, L. Lellouch, and H. Wittig. *Scalar condensate and light quark masses from overlap fermions*. Nucl. Phys. Proc. Suppl. 106, 766–771 (2002), [hep-lat/0110199](#).

BIBLIOGRAPHY

- [102] Y. Saad. *Iterative Methods for sparse linear systems*. SIAM, 2nd edition, 2003.
- [103] C. F. Jagels and L. Reichel. *A Fast Minimal Residual Algorithm for Shifted Unitary Matrices*. *Numerical Linear Algebra with Applications* 1(6), 555–570 (1994).
- [104] L. Giusti, C. Hoelbling, M. Lüscher, and H. Wittig. *Numerical techniques for lattice QCD in the epsilon- regime*. *Comput. Phys. Commun.* 153, 31–51 (2003), [hep-lat/0212012](#).
- [105] A. Ukawa. *Computational cost of full QCD simulations experienced by CP-PACS and JLQCD Collaborations*. *Nucl. Phys. Proc. Suppl.* 106, 195–196 (2002).
- [106] T. A. DeGrand and P. Rossi. *Conditioning techniques for dynamical fermions*. *Comput. Phys. Commun.* 60, 211–214 (1990).
- [107] J. C. Sexton and D. H. Weingarten. *Hamiltonian evolution for the hybrid monte carlo algorithm*. *Nucl. Phys.* B380, 665–678 (1992).
- [108] M. Hasenbusch. *Speeding up the Hybrid-Monte-Carlo algorithm for dynamical fermions*. *Phys. Lett.* B519, 177–182 (2001), [hep-lat/0107019](#).
- [109] M. Hasenbusch and K. Jansen. *Speeding up lattice QCD simulations with clover-improved Wilson fermions*. *Nucl. Phys.* B659, 299–320 (2003), [hep-lat/0211042](#).
- [110] M. Lüscher. *Schwarz-preconditioned HMC algorithm for two-flavour lattice QCD*. *Comput. Phys. Commun.* 165, 199 (2005), [hep-lat/0409106](#).
- [111] A. Ali Khan et al. *Accelerating Hasenbusch’s acceleration of hybrid Monte Carlo*. *Nucl. Phys. Proc. Suppl.* 129, 853–855 (2004), [hep-lat/0309078](#).
- [112] A. Ali Khan et al. *Accelerating the hybrid Monte Carlo algorithm*. *Phys. Lett.* B564, 235–240 (2003), [hep-lat/0303026](#).
- [113] K. Jansen. *Actions for dynamical fermion simulations: Are we ready to go?* *Nucl. Phys. Proc. Suppl.* 129, 3–16 (2004), [hep-lat/0311039](#).
- [114] B. Orth, T. Lippert, and K. Schilling. *Finite-size effects in lattice QCD with dynamical Wilson fermions*. *Phys. Rev.* D72, 014503 (2005), [hep-lat/0503016](#).
- [115] T. Takaishi and P. de Forcrand. *Testing and tuning new symplectic integrators for hybrid Monte Carlo algorithm in lattice QCD*. (2005), [hep-lat/0505020](#).

- [116] M. A. Clark and A. D. Kennedy. *Accelerating fermionic molecular dynamics*. (2004), [hep-lat/0409134](#).
- [117] M. Della Morte et al. *Simulating the Schroedinger functional with two pseudo-fermions*. *Comput. Phys. Commun.* 156, 62–72 (2003), [hep-lat/0307008](#).
- [118] M. J. Peardon and J. Sexton. *Multiple molecular dynamics time-scales in hybrid Monte Carlo fermion simulations*. *Nucl. Phys. Proc. Suppl.* 119, 985–987 (2003), [hep-lat/0209037](#).
- [119] Martin Lüscher. *Lattice QCD with light Wilson quarks*. *PoS(LAT2005)002* (2005), [hep-lat/0509152](#).
- [120] A. Frommer, V. Hannemann, B. Nockel, T. Lippert, and K. Schilling. *Accelerating Wilson fermion matrix inversions by means of the stabilized bi-conjugate gradient algorithm*. *Int. J. Mod. Phys. C5*, 1073–1088 (1994), [hep-lat/9404013](#).
- [121] K. Jansen and C. Liu. *Study of Liapunov exponents and the reversibility of molecular dynamics algorithms*. *Nucl. Phys. Proc. Suppl.* 53, 974–976 (1997), [hep-lat/9607057](#).
- [122] C. Liu, A. Jaster, and K. Jansen. *Liapunov exponents and the reversibility of molecular dynamics algorithms*. *Nucl. Phys. B524*, 603–617 (1998), [hep-lat/9708017](#).
- [123] R. G. Edwards, I. Horvath, and A. D. Kennedy. *Instabilities and non-reversibility of molecular dynamics trajectories*. *Nucl. Phys. B484*, 375–402 (1997), [hep-lat/9606004](#).
- [124] C. Urbach. *Untersuchung der Reversibilitätsverletzung im Hybrid Monte Carlo Algorithmus*. Master’s thesis, Freie Universität Berlin, Fachbereich Physik, 2002.
- [125] R. Gupta, G. W. Kilcup, and S. R. Sharpe. *Tuning the hybrid monte carlo algorithm*. *Phys. Rev. D38*, 1278 (1988).
- [126] S. Fischer et al. *A Parallel SSOR Preconditioner for Lattice QCD*. *Comp. Phys. Commun.* 98, 20–34 (1996), [hep-lat/9602019](#).
- [127] R. C. Brower, T. Ivanenko, A. R. Levi, and K. N. Orginos. *Chronological inversion method for the Dirac matrix in hybrid Monte Carlo*. *Nucl. Phys. B484*, 353–374 (1997), [hep-lat/9509012](#).

BIBLIOGRAPHY

- [128] R. C. Brower, A. R. Levi, and K. Orginos. *Extrapolation methods for the Dirac inverter in hybrid Monte Carlo*. Nucl. Phys. Proc. Suppl. 42, 855–857 (1995), [hep-lat/9412004](#).
- [129] M. Lüscher. *Solution of the Dirac equation in lattice QCD using a domain decomposition method*. Comput. Phys. Commun. 156, 209–220 (2004), [hep-lat/0310048](#).
- [130] K. Jansen. *Recent developments in fermion simulation algorithms*. Nucl. Phys. Proc. Suppl. 53, 127–133 (1997), [hep-lat/9607051](#).
- [131] P. de Forcrand and T. Takaishi. *Fast fermion Monte Carlo*. Nucl. Phys. Proc. Suppl. 53, 968–970 (1997), [hep-lat/9608093](#).
- [132] R. Frezzotti and K. Jansen. *A polynomial hybrid Monte Carlo algorithm*. Phys. Lett. B402, 328–334 (1997), [hep-lat/9702016](#).
- [133] R. Frezzotti and K. Jansen. *The PHMC algorithm for simulations of dynamical fermions. I: Description and properties*. Nucl. Phys. B555, 395–431 (1999), [hep-lat/9808011](#).
- [134] R. Frezzotti and K. Jansen. *The PHMC algorithm for simulations of dynamical fermions. II: Performance analysis*. Nucl. Phys. B555, 432–453 (1999), [hep-lat/9808038](#).
- [135] S. Weinberg. *Phenomenological Lagrangians*. Physica A96, 327 (1979).
- [136] J. Gasser and H. Leutwyler. *Chiral perturbation theory to one loop*. Ann. Phys. 158, 142 (1984).
- [137] J. Gasser and H. Leutwyler. *Chiral perturbation theory: expansions in the mass of the strange quark*. Nucl. Phys. B250, 465 (1985).
- [138] G. Rupak and N. Shoresh. *Chiral perturbation theory for the Wilson lattice action*. Phys. Rev. D66, 054503 (2002), [hep-lat/0201019](#).
- [139] S. Aoki. *New phase structure for lattice qcd with wilson fermions*. Phys. Rev. D30, 2653 (1984).
- [140] S. Aoki. *A solution to the U(1) problem on a lattice*. Phys. Rev. Lett. 57, 3136 (1986).
- [141] S. Aoki. *On the phase structure of QCD with Wilson fermions*. Prog. Theor. Phys. Suppl. 122, 179–186 (1996), [hep-lat/9509008](#).

- [142] G. Münster. *On the phase structure of twisted mass lattice QCD*. JHEP 09, 035 (2004), [hep-lat/0407006](#).
- [143] S. R. Sharpe and J. M. S. Wu. *The phase diagram of twisted mass lattice QCD*. Phys. Rev. D70, 094029 (2004), [hep-lat/0407025](#).
- [144] E.-M. Ilgenfritz, W. Kerler, M. Müller-Preußker, A. Sternbeck, and H. Stüben. *A numerical reinvestigation of the Aoki phase with $N(f) = 2$ Wilson fermions at zero temperature*. Phys. Rev. D69, 074511 (2004), [hep-lat/0309057](#).
- [145] A. Sternbeck, E.-M. Ilgenfritz, W. Kerler, M. Müller-Preußker, and H. Stüben. *The Aoki phase for $N(f) = 2$ Wilson fermions revisited*. Nucl. Phys. Proc. Suppl. 129, 898–900 (2004), [hep-lat/0309059](#).
- [146] S. Aoki, S. Boettcher, and A. Gocksch. *Spontaneous breaking of flavor symmetry and parity in the Nambu-Jona-Lasinio model with Wilson fermions*. Phys. Lett. B331, 157–164 (1994), [hep-lat/9312084](#).
- [147] S. Aoki, A. Ukawa, and T. Umemura. *Finite temperature phase structure of lattice QCD with Wilson quark action*. Phys. Rev. Lett. 76, 873–876 (1996), [hep-lat/9508008](#).
- [148] Y. Iwasaki, K. Kanaya, S. Sakai, and T. Yoshie. *Quark confinement in multi-flavor quantum chromodynamics*. Nucl. Phys. Proc. Suppl. 30, 327–330 (1993), [hep-lat/9211035](#).
- [149] S. Aoki. *Phase structure of lattice QCD with Wilson fermion at finite temperature*. Nucl. Phys. Proc. Suppl. 60A, 206–219 (1998), [hep-lat/9707020](#).
- [150] S. Aoki et al. *Non-trivial phase structure of $N(f) = 3$ QCD with $O(a)$ -improved Wilson fermion at zero temperature*. Nucl. Phys. Proc. Suppl. 106, 263–265 (2002), [hep-lat/0110088](#).
- [151] F. Farchioni et al. *Lattice spacing dependence of the first order phase transition for dynamical twisted mass fermions*. (2005), [hep-lat/0506025](#). Accepted for publication in Phys. Lett. B.
- [152] D. J. Gross and A. Neveu. *Dynamical symmetry breaking in asymptotically free field theories*. Phys. Rev. D10, 3235 (1974).
- [153] S. Aoki and K. Higashijima. *The recovery of the chiral symmetry in lattice Gross-Neveu model*. Prog. Theor. Phys. 76, 521 (1986).
- [154] T. Izubuchi, J. Noaki, and A. Ukawa. *Two-dimensional lattice Gross-Neveu model with Wilson fermion action at finite temperature and chemical potential*. Phys. Rev. D58, 114507 (1998), [hep-lat/9805019](#).





BIBLIOGRAPHY

- [155] J. Noaki, T. Izubuchi, and A. Ukawa. *Two-dimensional Gross-Neveu model with Wilson fermion action at finite temperature and density*. Nucl. Phys. Proc. Suppl. 73, 483–485 (1999), [hep-lat/9809071](#).
- [156] K Nagai. *Two-dimensional Gross-Neveu model with Wilson twisted mass fermions*. private communication.
- [157] A. Ali Khan et al. *Chiral properties of domain-wall quarks in quenched QCD*. Phys. Rev. D63, 114504 (2001), [hep-lat/0007014](#).
- [158] K. Orginos. *Chiral properties of domain wall fermions with improved gauge actions*. Nucl. Phys. Proc. Suppl. 106, 721–723 (2002), [hep-lat/0110074](#).
- [159] Y. Aoki et al. *Domain wall fermions with improved gauge actions*. Phys. Rev. D69, 074504 (2004), [hep-lat/0211023](#).
- [160] K. Jansen and K.-I. Nagai. *Reducing residual-mass effects for domain-wall fermions*. JHEP 12, 038 (2003), [hep-lat/0305009](#).
- [161] S. Aoki et al. *Bulk first-order phase transition in three-flavor lattice QCD with $O(a)$ -improved Wilson fermion action at zero temperature*. (2004), [hep-lat/0409016](#).
- [162] Y. Iwasaki. *Renormalization group analysis of lattice theories and improved lattice action. 2. four-dimensional nonabelian $SU(N)$ gauge model*. UTHEP-118.
- [163] Y. Iwasaki. *Renormalization group analysis of lattice theories and improved lattice action: two-dimensional nonlinear $O(N)$ sigma model*. Nucl. Phys. B258, 141–156 (1985).
- [164] T. Takaishi. *Heavy quark potential and effective actions on blocked configurations*. Phys. Rev. D54, 1050–1053 (1996).
- [165] P. de Forcrand et al. *Search for effective lattice action of pure QCD*. Nucl. Phys. Proc. Suppl. 53, 938–941 (1997), [hep-lat/9608094](#).
- [166] P. Weisz. *Continuum limit improved lattice action for pure Yang-Mills theory. 1*. Nucl. Phys. B212, 1 (1983).
- [167] P. Weisz and R. Wohlert. *Continuum limit improved lattice action for pure Yang-Mills theory. 2*. Nucl. Phys. B236, 397 (1984).
- [168] F. Farchioni et al. *The phase structure of lattice QCD with Wilson quarks and renormalization group improved gluons*. Eur. Phys. J. C42, 73–87 (2005), [hep-lat/0410031](#).

- [169] R. Frezzotti and G. C. Rossi. *Twisted-mass lattice QCD with mass non-degenerate quarks*. Nucl. Phys. Proc. Suppl. 128, 193–202 (2004), [hep-lat/0311008](#).
- [170] K. Jansen and C. Liu. *Implementation of Symanzik’s improvement program for simulations of dynamical Wilson fermions in lattice QCD*. Comput. Phys. Commun. 99, 221–234 (1997), [hep-lat/9603008](#).
- [171] R.W. Freund. in Numerical Linear Algebra, L. Reichel, A. Ruttan and R.S. Varga (eds.) page p. 101 (1993).
- [172] U. Glässner et al. *How to compute Green’s functions for entire mass trajectories within Krylov solvers*. (1996), [hep-lat/9605008](#).
- [173] B. Jegerlehner. *Multiple mass solvers*. Nucl. Phys. Proc. Suppl. 63, 958–960 (1998), [hep-lat/9708029](#).

List of publications

Most of the results presented in this work are already published in one of the references listed below. The selection has been done with focus on those results where the author contributed significantly to the implementation, the simulations and the analysis. This concerns in particular the references [P2-P5], [P8], [P10], [P11] and [P12].

- [P1]  Collaboration: K. Jansen, C. McNeile, C. Michael, K. Nagai, M. Papinutto, J. Pickavance, A. Shindler, C. Urbach, I. Wetzorke, *Flavour Breaking Effects of Wilson twisted mass fermions*, accepted for publication in Phys. Lett. **B** (2005), [hep-lat/0507032](#).
- [P2]  Collaboration: K. Jansen, M. Papinutto, A. Shindler, C. Urbach, I. Wetzorke, *Quenched Scaling of Wilson twisted mass fermions*, accepted for publication in JHEP (2005), [hep-lat/0507010](#).
- [P3] F. Farchioni, K. Jansen, I. Montvay, E.E. Scholz, L. Scorzato, A. Shindler, N. Ukita, C. Urbach, U. Wenger, I. Wetzorke, *Lattice Spacing Dependence of the First Order Phase Transition for Dynamical Twisted Mass Fermions*, accepted for publication in Phys. Lett. **B** (2005), [hep-lat/0506025](#).
- [P4] C. Urbach, K. Jansen, A. Shindler, U. Wenger, *HMC algorithm with multiple time scale integration and mass preconditioning*, accepted for publication in Comp. Phys. Com. (2005), [hep-lat/0506011](#).
- [P5]  Collaboration: K. Jansen, M. Papinutto, A. Shindler, C. Urbach, I. Wetzorke, *Light quarks with twisted mass fermions*, Phys.Lett. **B619** (2005) 184-191, [hep-lat/0503031](#),
- [P6]  Collaboration: W. Bietenholz, S. Capitani, T. Chiarappa, N. Christian, M. Hasenbusch, K. Jansen, K-I. Nagai, M. Papinutto, L. Scorzato, S. Shcheredin, A. Shindler, C. Urbach, U. Wenger, I. Wetzorke, *Going chiral: overlap versus twisted mass fermions*, JHEP **0412** (2004) 044, [hep-lat/0411001](#).
- [P7] F. Farchioni, K. Jansen, I. Montvay, E.E. Scholz, L. Scorzato, A. Shindler, N. Ukita, C. Urbach, I. Wetzorke, *The phase structure of lattice QCD with Wilson quarks and renormalization group improved gluons*, Eur.Phys.J.**C42** (2005) 73-87, [hep-lat/0410031](#).
- [P8] F. Farchioni, R. Frezzotti, K. Jansen, I. Montvay, G.C. Rossi, E.E. Scholz, A. Shindler, N. Ukita, C. Urbach, I. Wetzorke, *Twisted mass quarks and the phase structure of lattice QCD*, Eur.Phys.J.**C39** (2005) 421-433, [hep-lat/0406039](#).

- [P9] $\chi_{\text{L}}^{\text{F}}$ Collaboration: T. Chiarappa, K. Jansen, K.-I. Nagai, M. Papinutto, L. Scorzato, A. Shindler, C. Urbach, U. Wenger, I. Wetzorke, *Comparison between overlap and twisted mass fermions towards the chiral limit*, Nucl.Phys.Proc.Suppl.**140** (2005) 683, hep-lat/0409109.
- [P10] $\chi_{\text{L}}^{\text{F}}$ Collaboration: T. Chiarappa, K. Jansen, K.-I. Nagai, M. Papinutto, L. Scorzato, A. Shindler, C. Urbach, U. Wenger, I. Wetzorke, *Comparing iterative methods for overlap and twisted mass fermions*, Nucl.Phys.Proc.Suppl.**140** (2005) 853, hep-lat/0409107.
- [P11] F. Farchioni, C. Urbach, R. Frezzotti, K. Jansen, I. Montvay, G.C. Rossi, E.E. Scholz, A. Shindler, N. Ukita, I. Wetzorke, *Exploring the phase structure of lattice QCD with twisted mass quarks*, Nucl.Phys.Proc.Suppl.**140** (2005) 240, hep-lat/0409098.
- [P12] $\chi_{\text{L}}^{\text{F}}$ Collaboration: K. Jansen, A. Shindler, C. Urbach, I. Wetzorke, *Scaling test for Wilson twisted mass QCD*, Phys. Lett. **B586** (2004) 432-438, hep-lat/0312013.
- [P13] $\chi_{\text{L}}^{\text{F}}$ Collaboration: W. Bietenholz, S. Capitani, T. Chiarappa, M. Hasenbusch, K. Jansen, M. Müller-Preussker, K.I. Nagai, M. Papinutto, S. Shcheredin, A. Shindler, C. Urbach, I. Wetzorke, *Extracting physics from an unphysical situation: light mesons in a small box*, published in: NIC Symposium 2004, NIC Series **Volume 20** (2004) 117-127, physics/0309072.

Danksagung

An dieser Stelle möchte ich die Gelegenheit nutzen mich bei allen zu bedanken, die zum Zustandekommen dieser Arbeit beigetragen haben.

Karl Jansen bin ich zu besonderem Dank verpflichtet. Sein Vertrauen, seine Unterstützung und seine vielen Ideen und inhaltlichen Anregungen haben wesentlich zum Gelingen dieser Arbeit beigetragen.

Danken möchte ich auch Volkard Linke für seine Unterstützung und die angenehmen Arbeitsbedingungen. Auch wäre diese Arbeit ohne meine Stelle in seiner Arbeitsgruppe an der Freien Universität Berlin nicht möglich gewesen.

Moreover, I am indebted to Andrea Shindler, Urs Wenger and Ines Wetzorke for their support and many important discussions. It was a pleasure to work with you!

Diese Arbeit ist im Rahmen zweier Kollaborationen entstanden, der \mathbb{X}_F und der Hamburg-Münster-Zeuthen Kollaboration. Mein Dank gilt allen Mitgliedern dieser Gruppen, und insbesondere Thomas Chiarappa, Nils Christian, Federico Farchioni, Roberto Frezzotti, Istvan Montvay, Kei-ichi Nagai, Mauro Papinutto, Giancarlo Rossi, Enno E. Scholz und Luigi Scorzato.

Die in dieser Arbeit präsentierten Ergebnisse waren nur mit großem Rechenzeitaufwand möglich. Die dafür nötigen Ressourcen wurden von den Rechenzentren des NIC, DESY Zeuthen, HLRN and FZ Karlsruhe bereitgestellt. Mein Dank gilt allen Mitarbeitern dieser Rechenzentren und insbesondere Peter Wegner für ihre Unterstützung. In addition, I thank Roberto Frezzotti and Claudio Destri for providing me access to a PC-cluster in Milano.

Nicht unerwähnt lassen möchte ich die DFG, NIC und DESY, sowie das INFN, die mir mit finanzieller Unterstützung die Teilnahme an wichtige Konferenzen, Workshops und einen Forschungsaufenthalt in Mailand ermöglicht haben.

Für die kritische Durchsicht dieses Manuskripts und viele Anregungen und Kommentare möchte ich Volkard Linke, Andrea Shindler, Axel Thimm, Urs Wenger und insbesondere Karl Jansen meinen Dank ausdrücken.

Nicht zuletzt gebührt meiner Familie und im besonderen Katja Schiffers mein herzlicher Dank: Meiner Familie für ihren bedingungslosen Rückhalt, meiner Freundin Katja für ihre Zuneigung und weil sie mir in den letzten Jahren immer zur Seite gestanden hat.