

SIMON KUBERSKI

STANDARD MODEL PARAMETERS IN THE HEAVY
QUARK SECTOR FROM THREE-FLAVOR LATTICE
QCD

2020

Theoretische Physik

STANDARD MODEL PARAMETERS IN THE HEAVY QUARK
SECTOR FROM THREE-FLAVOR LATTICE QCD

Inauguraldissertation zur Erlangung des Doktorgrades
der Naturwissenschaften im Fachbereich Physik
der Mathematisch-Naturwissenschaftlichen Fakultät
der Westfälischen Wilhelms-Universität Münster

vorgelegt von
SIMON KUBERSKI
aus Bielefeld

– 2020 –

DEKAN:	Prof. Dr. Gerhard Wilde
ERSTER GUTACHTER:	Prof. Dr. Jochen Heitger
ZWEITER GUTACHTER:	Priv.-Doz. Dr. Georg Bergner
TAG DER MÜNDLICHEN PRÜFUNG:	28.09.2020
TAG DER PROMOTION:	28.09.2020

ABSTRACT

Despite its great success, the Standard Model of particle physics fails to explain experimental observations as the baryon asymmetry of the universe or the existence of dark matter. Since the comparison of theoretical predictions and experimental findings does not reveal deviations that are statistically significant, many corrections to the Standard Model that incorporate so-far unexplained phenomena have been ruled out.

The flavor physics sector is tested in indirect searches for new physics beyond the Standard Model. To constrain theoretical predictions in this sector, the contributions of low-energy Quantum Chromodynamics (QCD) have to be determined to high precision. Lattice QCD provides the framework to determine Standard Model parameters *ab initio* with controlled systematics and without model-dependent assumptions.

The large spread of energy scales which are present in observables involving charm and bottom quarks leads to theoretical and numerical challenges in lattice calculations. In our work, we lay the foundations to determine Standard Model parameters in the heavy quark sector of three-flavor QCD. We perform a non-perturbative improvement of the discretized theory to minimize systematic uncertainties in the determination of the mass of the charm quark from lattice simulations with $2+1$ dynamical flavors. In the b quark sector, we use Heavy Quark Effective Theory (HQET) to describe strong interactions involving the heavy quark. We prepare the renormalization of the effective theory to next-to-leading order via a non-perturbative finite-volume matching with QCD and extract low-energy couplings of Heavy Meson Chiral Perturbation Theory, an effective theory describing the physics of heavy-light mesons.

The findings of our work will help to improve theoretical predictions in the heavy quark sector of the Standard Model based on results from lattice calculations with carefully controlled systematic effects.

ZUSAMMENFASSUNG

Trotz seines großen Erfolges kann das Standardmodell der Teilchenphysik nicht dazu verwendet werden, experimentelle Beobachtungen wie die Baryonenasymmetrie des Universums oder die Existenz dunkler Materie zu erklären. Da der Vergleich theoretischer Vorhersagen und experimenteller Befunde keine statistisch signifikanten Abweichungen offenbart, wurden viele Korrekturen des Standardmodells, welche bisher nicht erklärte Phänomene beinhalten, ausgeschlossen.

Der Flavorsektor wird durch indirekte Suchen nach neuer Physik jenseits des Standardmodells getestet. Um die theoretischen Vorhersagen in diesem Bereich einzuschränken, müssen niederenergetische Beiträge der QCD mit hoher Präzision bestimmt werden. Quantenchromodynamik (QCD) auf dem Gitter schafft die Voraussetzungen um Parameter des Standardmodells ab initio mit kontrollierten systematischen Unsicherheiten und ohne modellabhängige Annahmen bestimmen zu können.

Die große Bandbreite an Energieskalen, welche in Observablen die Charm- und Bottom-Quarks beinhalten vorhanden sind, führt zu theoretischen und numerischen Herausforderungen in Gitterrechnungen. In unserer Arbeit legen wir die Grundlagen um Standardmodellparameter im Sektor schwerer Quarks der Drei-Flavor QCD zu bestimmen. Wir verbessern die diskretisierte Theorie nichtperturbativ um systematische Unsicherheiten in der Bestimmung der Charm-Quarkmasse auf Grundlage von Gitterrechnungen mit $2 + 1$ dynamischen Quarkflavors zu minimieren. Im b-Quarksektor verwenden wir die effektive Theorie schwerer Quarks (HQET) um starke Wechselwirkungen mit schweren Quarks zu beschreiben. Wir bereiten die Renormierung der effektiven Theorie zur ersten Ordnung mittels eines Matchings mit QCD in kleinem Volumen vor und extrahieren niederenergetische Kopplungskonstanten der chiralen Störungstheorie schwerer Mesonen, welche eine effektive Theorie der Physik der schwer-leichten Mesonen darstellt.

Die Ergebnisse unserer Arbeit werden dabei helfen theoretische Vorhersagen des Standardmodells im Bereich schwerer Quarks auf der Grundlage der Resultate aus Gitterrechnungen mit gewissenhaft kontrollierten systematischen Effekten zu verbessern.

CONTENTS

1	INTRODUCTION	1
I THEORY		
2	QCD ON THE LATTICE	7
2.1	Gauge action	8
2.1.1	On-shell improvement	9
2.1.2	Lüscher-Weisz gauge action	10
2.2	Fermion action	11
2.2.1	Wilson fermions	12
2.2.2	Improved Wilson fermions	13
3	THE SCHRÖDINGER FUNCTIONAL	17
3.1	Continuum formulation	17
3.1.1	Gluon action	17
3.1.2	Quarks in the Schrödinger functional	18
3.2	Lattice formulation	20
3.3	Schrödinger functional correlation functions	22
3.3.1	Boundary sources	22
3.3.2	Expectation values	23
3.3.3	Correlation functions	23
3.3.4	Flavor off-diagonal bilinears	25
3.4	Addendum: Open Boundary Conditions	25
3.4.1	Motivation	25
3.4.2	Setup	26
3.4.3	Boundary improvement	26
4	OPERATOR IMPROVEMENT AND RENORMALIZATION	29
4.1	Improvement of the action: Revisited	29
4.1.1	Mass renormalization	29
4.1.2	Improvement of the coupling	30
4.2	Operator improvement	30
4.2.1	Renormalization of improved operators	31
4.2.2	Improvement of local currents	31
4.3	Improved quark masses	32
4.3.1	Renormalization of the bare subtracted quark mass	32
4.3.2	The PCAC relation	33
4.3.3	Current quark masses	34
4.3.4	Relations between renormalized quark masses	35
4.3.5	The ratio-difference method	37
5	ALGORITHMIC TECHNIQUES	39
5.1	Simulation algorithms	39
5.1.1	Hybrid Monte Carlo	40
5.1.2	Integration schemes	41
5.1.3	HMC with dynamical quarks	42

5.1.4	Rational Hybrid Monte Carlo	43
5.1.5	Reweighting	44
5.2	Frequency splitting	45
5.2.1	Hasenbusch splitting	45
5.2.2	Splitting the rational determinant	46
5.2.3	Twisted mass reweighting	46
5.3	Solver	47
5.3.1	Conjugate Gradient solver	48
5.3.2	Multi-shift Conjugate Gradient solver	48
5.3.3	Even-odd preconditioning	49
5.3.4	Block preconditioning	50
5.3.5	Deflation acceleration	51
5.3.6	Distance preconditioning	54
5.4	Errors in Monte Carlo data	55
5.4.1	Correlations	55
5.4.2	Slow modes	57
6	EXTRACTION OF PHYSICAL OBSERVABLES	59
6.1	Meson masses	59
6.1.1	The spectral decomposition of correlation functions	59
6.1.2	Correlation functions and open boundary conditions	60
6.2	The generalized eigenvalue problem	61
6.3	Fitting	63
6.3.1	Fitting correlated data	64
6.4	Gradient flow observables	65
6.4.1	The Wilson flow	66
6.4.2	The Zeuthen flow	67
6.4.3	Setting a scale from the flow	67
6.4.4	The gradient flow coupling in the Schrödinger functional	68
6.4.5	The topological charge from the gradient flow	70
6.5	Smearing	72
6.5.1	APE smearing	72
6.5.2	HYP smearing	73
6.5.3	Gaussian smearing	74
6.5.4	Smearing from three-dimensional fermions	75
6.5.5	Smearing from three-dimensional bosons	76
7	HEAVY QUARK EFFECTIVE THEORY	79
7.1	Continuum HQET	79
7.1.1	Derivation of the HQET Lagrangian	80
7.1.2	Heavy Quark Symmetry	83
7.2	HQET on the lattice	84
7.2.1	Lattice action	84
7.2.2	Expectation values at next-to-leading order	86
7.3	Non-perturbative matching with QCD	87
7.3.1	Matching in finite volume	87

7.3.2	Determination of the matching parameters . . .	88
7.3.3	Quark mass dependence	90
7.4	HQET observables	91
7.4.1	The axial current in HQET	91
7.4.2	The B meson mass	92
7.4.3	The GEVP in HQET	93
II THE LINE OF CONSTANT PHYSICS FOR A NON-PERTURBATIVE MATCHING OF HQET AND QCD		
8	SIMULATIONS ALONG A LINE OF CONSTANT PHYSICS	97
8.1	The line of constant physics: QCD	97
8.1.1	The running of the coupling	98
8.1.2	Interpolation of the coupling at fixed L/a	99
8.1.3	Target precision	102
8.1.4	Tuning towards the critical hopping parameter	103
8.1.5	Step scaling to the matching volume	105
8.1.6	The matching volume	107
8.2	The line of constant physics: HQET	108
8.2.1	Tuning in the matching volume	109
8.2.2	Step scaling towards large volume simulations .	110
8.2.3	Contact to the large-volume simulations	112
8.3	The running of the mass	114
8.3.1	The step scaling function of the mass	114
8.3.2	Determination of RGI quark masses	115
8.3.3	Determination of the step scaling function . . .	116
8.4	Conclusions	118
9	ALGORITHMIC EXPERIENCES	119
9.1	Solvers	119
9.1.1	Quark mass dependence	120
9.1.2	Tuning of the deflated solver	121
9.2	RHMC	126
9.3	Scaling of autocorrelation times	129
9.3.1	The topological charge in finite volume	129
III RENORMALIZATION AND IMPROVEMENT FROM FINITE-VOLUME SIMULATIONS		
10	STRATEGY	135
10.1	Estimators from time slice averages	135
10.2	Quark mass parametrization	137
10.2.1	Redefinition of the estimators	137
10.2.2	Polynomial interpolations	139
10.3	The line of constant physics	140
11	QUARK MASS IMPROVEMENT IN THE STRONGLY COUPLED REGIME OF QCD	145
11.1	Gauge ensembles	145
11.1.1	Topological charge	146
11.2	Results	148

11.2.1	Mass parametrizations and determination of the estimators	148
11.2.2	Chiral interpolations	152
11.2.3	Interpolation in the coupling	156
11.2.4	Errors and correlations	161
11.3	Ambiguity checks	162
11.4	Estimators from time slice averages	165
11.5	Conclusions	166
12	QUARK MASS RENORMALIZATION FOR THE MATCHING OF QCD AND HQET	171
12.1	PCAC masses	173
12.2	Estimators	175
12.3	Ambiguities	179
12.4	Renormalization of the pseudoscalar density and the axial current	185
12.4.1	Renormalization of the pseudoscalar density	185
12.4.2	Renormalization of the axial current	186
12.5	Hopping parameters for constant renormalized quark masses	189
12.6	Verification of improvement	192
12.7	Conclusions	199
IV	THE MASS OF THE CHARM QUARK	
13	CLS ENSEMBLES	203
13.1	The chiral trajectory	205
13.2	Quark mass shifts	206
14	DETERMINATION OF THE CHARM QUARK MASS	209
14.1	Calibration of the measurements	209
14.2	The renormalized charm quark mass from current quark masses	212
15	LATTICE COMPUTATION	215
15.1	Correlation functions	215
15.2	Meson masses	217
15.2.1	Fit ranges	218
15.2.2	Extraction of the effective mass	222
15.3	Current quark masses	226
15.3.1	Fit ranges	228
15.4	Distance preconditioning	232
15.4.1	Tuning procedure	234
15.4.2	Impact on physical observables	234
15.5	Stochastic sources	237
15.6	Chiral-continuum extrapolation	240
15.7	Preliminary result and discussion	243

V HQET IN LARGE VOLUME	
16 HEAVY QUARK EFFECTIVE THEORY WITH OPEN BOUND- ARY CONDITIONS	247
16.1 Setup of the measurements	247
16.2 Optimization of the variational basis	248
16.2.1 Definition of a suitable basis	249
16.2.2 Reduction of the number of smearing procedures	251
16.2.3 Variational basis from smearing method	253
16.3 Boundary effects	255
16.3.1 Boundary states	256
16.3.2 Asymmetry in the correlation matrix	258
17 COMPUTATION OF THE $B^*B\pi$ COUPLING	259
17.1 Heavy Meson Chiral Perturbation Theory	260
17.2 Lattice computation	261
17.3 Setup	264
17.4 Results	265
17.4.1 Plateaus	265
17.4.2 Chiral-continuum extrapolation	266
17.5 Discussion	268
17.6 The matrix element of the first radial excitations	270
VI CONCLUSIONS	
18 CONCLUSIONS	275
A CONVENTIONS	279
A.1 The special unitary group	279
A.2 Dirac matrices	280
A.3 Lattice derivatives	281
B TABLES	283
B.1 The line of constant physics for a non-perturbative matching of HQET and QCD	283
B.2 Renormalization and improvement from finite-volume simulations	287
B.3 The mass of the charm quark	291
B.4 HQET in large volume	296
BIBLIOGRAPHY	297

INTRODUCTION

Four fundamental forces are sufficient to describe the laws of physics, namely the long ranged gravitational and electromagnetic forces and the short ranged weak and strong forces. The strong, weak and electromagnetic forces can be embedded in the Standard Model of particle physics where quantum field theories are used to describe elementary particles and the fundamental interactions among them.

Since the Standard Model has been developed, it has been confirmed experimentally by the discovery of the underlying particles: quarks, leptons, gauge bosons and, most prominently, the higgs boson. Numerous high precision tests did not reveal deviations between theoretical predictions and experimental findings. Whereas the development of this model might be the greatest achievement of modern physics, its accuracy leads to conceptual challenges regarding the construction of new theories.

Some experimental observations in particle physics cannot be explained within the Standard Model. The Standard Model describes baryonic matter which contributes about 5% to the energy content of the universe. It does not explain the origin of dark matter or dark energy which amount to the remaining 95%. Furthermore it fails to explain the baryon asymmetry of the universe and the experimentally confirmed non-vanishing neutrino masses. Since every model that is developed to explain one or more of these phenomena has implications on the interactions within the Standard Model, the agreement of Standard Model predictions and observations to high precision has ruled out a large number of such models.

The search for new physics beyond the Standard Model is performed in two different regimes. Direct searches for new particles are the main approach of high-energy experiments as the Large Hadron Collider. An upper bound on the energies in reach is given by the immense cost and technological constraints that are connected with the construction of new particle colliders. Therefore, the indirect search for new physics might lead to the first evidence of physics beyond the Standard Model. Here, low-energy processes are investigated to high precision to hunt for deviations from theoretical predictions.

The calculation of such predictions poses theoretical and numerical challenges. The perturbative treatment of Quantum Chromodynamics (QCD), which is the theory of strong interactions, fails at low energies due to the strength of the interactions in this regime. Lattice QCD is the natural ab initio method to perform computations in the low-energy sector of QCD. The formulation and regularization of the theory on

a finite space-time grid allows to describe the interactions between hadrons which, due to confinement [1], are the particles that take part in physical processes at low energies.

The flavor physics sector of low-energy QCD is believed to be a promising field for the search for new physics. For example, flavor-changing neutral currents are highly suppressed in the Standard Model and therefore sensitive to contributions from beyond the Standard Model theories. Currently, there are a few quantities in the sector of heavy quarks where discrepancies at the 3σ level are found. An example is given by the element $|V_{cb}|$ of the Cabbibo-Kobayashi-Maskawa (CKM) matrix. The determination of $|V_{cb}|$ from inclusive semi-leptonic $b \rightarrow c$ decays, an overview is given in [2], does not coincide with the determination from exclusive semi-leptonic $B \rightarrow D\ell\nu$ decays, e.g., by the Belle collaboration [3]. To compute the latter, the experimentally determined decay rates are combined with lattice QCD computations of hadronic matrix elements.

The calculation of hadronic quantities involving bottom quarks poses severe challenges for lattice QCD. These have their origin in the hierarchy of scales involved in the computation. To control all systematic effects, the requirement

$$L^{-1} \ll m_\pi \approx 140 \text{ MeV} \ll m_B \approx 5 \text{ GeV} \ll a^{-1}, \quad (1.1)$$

where L is the physical extent of the lattice and a is the lattice spacing, has to be fulfilled. This is not possible on modern computers. To circumvent this fundamental issue, we resort to an effective theory, namely Heavy Quark Effective Theory (HQET), to describe strong interactions involving bottom quarks.

HQET allows us to integrate out the heavy degrees of freedom and to subsequently perform an expansion of the Lagrange density in the inverse of the heavy quark mass m_h . The treatment of the light degrees of freedom remains unchanged. The non-perturbative matching of the effective theory with QCD allows us to renormalize quantities of interest such that the continuum limit of the lattice results can be taken. To conserve the predictive power of the theory, this matching is performed in finite-volume at $L \approx 0.5 \text{ fm}$ where relativistic bottom quarks can be included in the simulation of QCD.

To describe physics at the scale of the bottom quark, we take the next-to-leading order of HQET, i.e., terms of $O(1/m_h)$ into account. To decouple the effects of the next-to-leading order from discretization effects, we have to remove all effects of $O(a)$ from the discretized theory. The controlled removal of such effects goes under the name of Symanzik improvement [4, 5] and we apply it to the Lagrangian and all observables included in our calculations. Based on the renormalized effective theory, we are able to determine physical observables in the b quark sector such as the bottom quark mass, decay constants of B mesons and form factors of semi-leptonic B decays.

In the lattice regularization of the QCD Lagrangian, L and a correspond to the infrared and ultraviolet cut-off.

We will also use HQET to determine the leading order coupling constant of Heavy Meson Chiral Perturbation Theory ($\text{HM}\chi\text{PT}$). This effective theory describes heavy-light mesons in the limit of massless light and infinitely heavy quarks. An expansion in this theory guides chiral extrapolations of observables that have been determined on the lattice at unphysically large quark masses. Therefore, the precise knowledge of the leading order coupling helps to reduce the error on heavy quark observables from lattice QCD.

In contrast to bottom quarks, relativistic charm quarks can be simulated in large-volume lattice QCD. However, mass dependent cut-off effects of $\mathcal{O}(am_c^2)$ lead to systematic uncertainties in the continuum extrapolation of lattice results obtained from Wilson quarks. The improvement of the theory à la Symanzik allows to cancel the leading order cut-off effects and to improve the scaling towards the continuum limit. This enables us to determine the charm quark mass in $2+1$ flavor lattice QCD. The knowledge of bottom and charm quark mass is per se of interest since both are parameters of the Standard Model. Their precise knowledge is also important for the search for new physics since they enter theoretical predictions such as Higgs branching ratios to charm and bottom quarks and the inclusive semi-leptonic decay rate of B mesons [6]. The hopping parameters for physical charm quark masses on the CLS ensembles will be employed for the computation of heavy-charm correlation functions in our HQET calculations.

This work is structured as follows. In the first part we will give an introduction to the theoretical and computational foundations of our work. We will describe the discretized action that we use to simulate three-flavor QCD, explain its formulation in the Schrödinger functional and the Symanzik improvement of the theory. Afterwards we will introduce algorithmic techniques that enable us to perform our calculations and outline how they are used to optimize the performance of lattice calculations. Subsequently we will introduce a number of observables which are determined at different stages of our analysis and present their definition in our setup. We conclude this part with an introduction to HQET, the formulation of the effective theory on the lattice and an outline of the non-perturbative matching in finite-volume.

In the second part we will describe the simulation of gauge ensembles for the non-perturbative matching of QCD and HQET. We will explain the tuning that was done to perform the simulations along a line of constant physics and we will extract information on the running of the strong coupling constant and quark masses in the weakly coupled regime of non-perturbative QCD. In a separate chapter we will present some of the algorithmic experiences that we have made in our simulations of large lattices in finite volume. These can be used as a guideline for future simulations in the Schrödinger functional.

Afterwards we will turn to the operator improvement and renormalization that is necessary to determine heavy quark observables. We will outline our strategy to determine the improvement coefficients $b_A - b_P$ and b_m as well as the renormalization constant Z . This strategy is first applied in the coupling region of large-volume simulations, where the parameters can be used to renormalize and improve quark masses. We then employ our strategy in the region of the finite-volume matching and use the improvement coefficients and renormalization constants to tune to fixed renormalized heavy quark masses for the matching procedure. To ensure the validity of our results, we perform checks by a large number of variations in the definitions of our observables. We conclude this part by testing the removal of cut-off effects by the parameters b_m and Z in the matching region.

The determination of the charm quark mass is performed on large-volume ensembles generated by members of the CLS effort. We describe these ensembles and detail our strategy to compute the renormalization group invariant charm quark mass. Afterwards we explain in detail how we extracted the bare quantities from measurements on a large set of ensembles and present the results. This is followed by a discussion on different techniques that have been used to obtain the results with high precision. The chiral-continuum extrapolation of the charm quark mass is presented and a preliminary result at the physical point is given.

In the last part of this work, we discuss the computation of HQET observables on large-volume CLS ensembles. We present our investigations of the properties of these observables on lattices with open boundary conditions and describe our computations of the correlation functions that will be combined with the matching parameters to obtain renormalized observables in the effective theory. Afterwards we will introduce the coupling \hat{g} of $\text{HM}\chi\text{PT}$ and explain how we extract it on the CLS ensembles. We will conclude with a chiral-continuum extrapolation of \hat{g} , quote a result for the coupling in the chiral limit and give an outlook on future improvements of this calculation.

Part I
THEORY

QCD ON THE LATTICE

QCD is the theory of strongly interacting particles. In the continuum, its Lagrange density is given by

$$\mathcal{L}_{\text{QCD}} \equiv \mathcal{L}_{\text{YM}} + \mathcal{L}_f \quad (2.1)$$

$$= -\frac{1}{2} \text{Tr} [F_{\mu\nu} F^{\mu\nu}] + \sum_f \bar{\psi}_f (i\gamma_\mu D^\mu - m_f) \psi_f, \quad (2.2)$$

where the Yang-Mills Lagrangian \mathcal{L}_{YM} describes the interaction of gauge fields via the field strength tensor $F_{\mu\nu}$ and the fermion fields ψ_f of flavor f couple to the gauge fields via the covariant derivative D^μ .

There are many different ways to discretize the QCD action in order to simulate the theory on the lattice. In the following discussion, we split the action of QCD into the gauge action S_g and the fermion action S_f . Discretizing the gauge action is done straight-forwardly and does not pose many problems. This is different in the case of the fermionic action. All of the currently used fermion discretizations have advantages and disadvantages and the choice of a specific setup has to be adapted to the problems under investigation and the available computing resources. One shortcoming is common to all discretizations: They cannot circumvent the Nielsen-Ninomiya no-go theorem [8–10] which states that actions with the discretized Dirac operator D in even dimensions cannot at the same fulfill all of the following criteria [11]:

1. Locality of D .
2. Correct continuum limit as $a \rightarrow 0$.
3. No fermion doublers, i.e., the Fourier transformed operator $\tilde{D}(p)$ is invertible everywhere except $p_\mu = 0$.
4. Respect chiral symmetry on the lattice, i.e., the relation $\{\gamma_5, D\} = 0$ holds.

The so-called overlap fermions come as close as possible to satisfy these criteria. The corresponding Dirac operator fulfills a lattice version of chirality, the Ginsparg-Wilson equation [12], namely

$$\{\gamma_5, D\} = aD\gamma_5D. \quad (2.3)$$

Unfortunately, the cost of the simulation of overlap fermions is magnitudes larger than for other discretizations and therefore not practical for many applications. This is opposed to the widely used Wilson fermions which are cheap to simulate, but break chiral symmetry explicitly at finite lattice spacing.

For a derivation of the QCD Lagrangian we refer to textbooks on quantum field theory, e.g., [7].

In this chapter, we do not give an overview of all popular discretizations, but focus on the action which is used in all simulations related to our work. We start with the gauge action and explain the idea of Symanzik improved theories, before we discuss Wilson fermions.

In the following we assume that the reader is familiar with the basics of lattice simulations as they are taught in Master's courses on lattice field theory. Therefore, we do not explain the path integral formalism, the Wick rotation of the path integral to Euclidean time and the correspondence of lattice field theories and statistical mechanics. We also assume that the definition of ensemble averages that are defined from a sampling of gauge configurations according to the Boltzmann weight given by the lattice action and that approach the quantum mechanical expectation values for large sample sizes is known. Excellent introductions to these topics can be found in [13–17]

2.1 GAUGE ACTION

As a first step towards a discretized action of QCD, we like to express the gauge part of the theory, i.e., the action of Yang-Mills theory [18], on the lattice. Following [19], we start with a smooth continuum gauge field $A_\mu(x)$, $x \in \mathcal{R}$ and superimpose a lattice with lattice sites $x = an$, $n \in \mathcal{Z}^4$, where the small distance a between two lattice sites is called lattice spacing. In order to approximate $A_\mu(x)$ by a lattice gauge field $U_\mu(n)$, we define the latter as the parallel transporter from $an + a\hat{\mu}$ to an via the path ordered exponential

$$U_\mu(n) = \mathcal{T} \exp \left(a \int_0^1 dt A_\mu(an + a\hat{\mu} - ta\hat{\mu}) \right). \quad (2.4)$$

For small values of a , this lattice gauge field locally approximates $A_\mu(x)$. If we now define the plaquette as closed loop of four lattice gauge fields

$$U_{\mu\nu}(n) = U_\mu(n)U_\nu(n + a\hat{\mu})U_\mu(n + a\hat{\nu})^{-1}U_\nu(n)^{-1}, \quad (2.5)$$

which is a gauge invariant object, we have the ingredients to formulate the Wilson gauge action [1]

$$\begin{aligned} S_W(U) &= \frac{1}{g_0^2} \sum_n \sum_{1 \leq \mu < \nu \leq 4} \text{ReTr} [\mathbf{1} - U_{\mu\nu}(n)] \\ &\equiv \frac{1}{g_0^2} \sum_p \text{Tr} [\mathbf{1} - U(p)] \end{aligned} \quad (2.6)$$

for the group $SU(N)$ with the bare coupling g_0^2 . In the second line of eq. (2.6), we introduced the sum \sum_p over all oriented plaquettes $U(p)$. An expansion of the plaquette operator at small values of a yields

$$\text{Tr} [\mathbf{1} - U(p)] = -\frac{a^4}{2} \sum_{\mu, \nu} \text{Tr} [F_{\mu\nu} F_{\mu\nu}] + O(a^5) \quad (2.7)$$

with the field strength tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \quad (2.8)$$

and we see that we, to first order, restore the Yang-Mills action in the continuum limit. A detailed derivation with a similar formalism can be found in [13].

2.1.1 On-shell improvement

Symanzik [4, 5, 20] introduced the concept of improvement of lattice theories. The main idea is to consider the discretized theories as effective theories for the continuum one. In the limit of a vanishing lattice spacing a , full and effective theory coincide but at finite lattice spacing there are many different effective theories. The construction of an effective, i.e., discretized theory proceeds similar to a Taylor expansion in the lattice spacing and if the leading coefficients are known, they can be explicitly subtracted from the theory. The action of such an effective theory can be expressed by

$$S_{\text{eff}} = \int d^4x [\mathcal{L}_0(x) + a\mathcal{L}_1(x) + a^2\mathcal{L}_2(x) + \dots], \quad (2.9)$$

where \mathcal{L}_0 is the continuum Lagrange density and the higher order terms \mathcal{L}_k include operators of dimension $4 + k$.

Following Lüscher and Weisz [19, 21, 22], we can expand a lattice operator $\mathcal{O}(n)$ in polynomials of the gauge field $A_\mu(x)$ and its derivatives to obtain

$$\mathcal{O}(0) \sim \sum_{k=0}^{\infty} a^k \mathcal{O}_k(0). \quad (2.10)$$

We have seen the leading order of this expansion for the plaquette operator in eq. (2.7), where the dimension of the operator, i.e., the smallest k for which $\mathcal{O}_k \neq 0$, is four. Taking into account that $\mathcal{O}(n)$ transforms as a scalar field under the symmetry groups of the lattice, we see that all odd orders of a have to vanish and the next-to-leading operators are of dimension six [19].

For a generic gauge action, the approximate scaling of a physical quantity m towards the continuum limit $a \rightarrow 0$ is

$$a \frac{d}{da} m(a, \bar{g}(a)) = O(a^2), \quad (2.11)$$

where the function $\bar{g}(a)$ can be calibrated such that m matches the value of an external input \hat{m}_0 which, e.g., can be obtained from experiment. For a spectral quantity m_0 we set

$$m_0(a, \bar{g}(a)) = \hat{m}_0. \quad (2.12)$$

From universality, we expect that a large class of actions has the identical continuum limit and all of these actions differ by $O(a^2)$ effects [19]. Spectral quantities, e.g., meson masses are unaffected by a local gauge transformation of the lattice gauge fields.

A non-perturbatively improved gauge action is constructed such that the error term in eq. (2.11) is reduced to $O(a^4)$ and the scaling towards the continuum limit is faster compared to the unimproved theory. Using perturbative improvement up to k loops, the leading remaining terms are of order $O(g_0^{2(k+1)} a^2)$.

2.1.2 Lüscher-Weisz gauge action

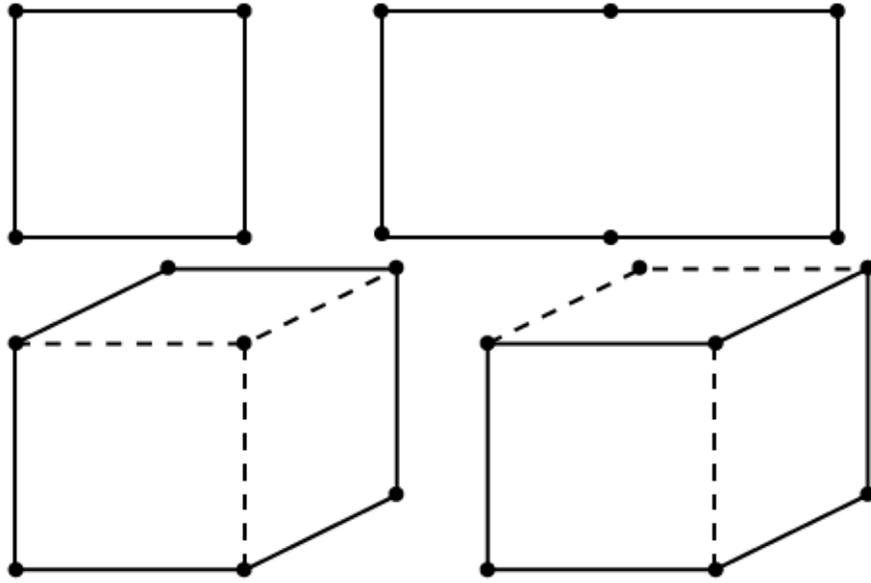


Figure 2.1: Gauge loops contributing to the improved gauge action. The loops in the first line are planar, the other two extend in three dimensions. Dots correspond to lattice sites and the dashed lines are drawn to guide the eye. Based on [19].

An on-shell improved gauge action can be constructed by a perturbative analysis of the coefficients of the next-to-leading order operators of the expansion of the plaquette operator which are of dimension six [19, 23]. In principle, three different six-link loops have to be considered, in addition to the simple plaquette of the Wilson gauge action. Their form is shown in figure 2.1. As it can be seen, a planar rectangle and two further three dimensional structures are added as elementary structures L_i . The action is constructed as extension of (2.6) via

$$S_g(\mathbf{U}) = \frac{1}{9_0} \sum_{i=0}^3 c_i(g_0^2) \sum_{l \in L_i} \mathbf{U}(l) \quad (2.13)$$

As for the plaquettes, loops which only differ by their orientation are considered equal.

where we replaced the sum over all plaquettes by a sum over all loops within the four elementary structures. Every structure is multiplied by a weight and their normalization is chosen to be

$$c_0(g_0^2) + 8c_1(g_0^2) + 8c_2(g_0^2) + 16c_3(g_0^2) = 1. \quad (2.14)$$

The derivation of the coefficients, as it is done in [24] up to corrections of order $O(g_0^4 a^2)$, leads to

$$c_0(g_0^2) = \frac{5}{3} + 0.2370g_0^2, \quad c_1(g_0^2) = -\frac{1}{12} - 0.02521g_0^2, \quad (2.15)$$

$$c_2(g_0^2) = -0.00441g_0^2, \quad c_3(g_0^2) = 0. \quad (2.16)$$

Taking only the tree-level values, $c_i \equiv c_i(0)$, we can construct a tree-level improved action from plaquettes and planar rectangles, since the weights of the non-planar loops vanish. We choose tree-level improvement over full $O(a)$ improvement since a full $O(a^2)$ improvement becomes too expensive and complicated as soon as we include quarks in the theory. Nevertheless, we expect an improved scaling by the inclusion of the tree-level values. We can rewrite eq. (2.13) to look similar to the Wilson action, eq. (2.6), by adding the weighted sum over the rectangles,

$$S_{LW}(U) = \frac{\beta}{6} \left(c_0 \sum_p \text{Tr}[\mathbf{1} - U(p)] + c_1 \sum_r \text{Tr}[\mathbf{1} - U(r)] \right). \quad (2.17)$$

We have introduced the bare inverse gauge coupling $\beta = 2N/g_0^2$ for an $SU(N)$ gauge group. In the case of QCD which is under investigation in our work, the gauge fields live in the group $SU(3)$. This so-called Lüscher-Weisz action defined by eq. (2.17) is chosen in all of our simulations within the bulk of the lattice. We will come back to the boundaries at a later stage.

The positivity of the Wilson gauge action has been shown in the transfer matrix formalism [26]. The theory satisfies the Osterwalder-Schrader axioms [27, 28]. This is not the case for theories including operators of dimension six, where the violation of positivity is a lattice artifact [29, 30]. As long as distances significantly larger than a are considered the effect can be ignored. The violation of positivity is expected to vanish in the continuum limit. Special care has to be taken, when the variational method, which is introduced in section 6.2, is used at short distances.

An other example for an improved gauge action is the Iwasaki [25] gauge action.

Significantly larger distances of $O(3a)$.

2.2 FERMION ACTION

Having set up the gluonic part of the QCD action, we now turn to the fermionic action. We will introduce Wilson fermions and discuss their properties, especially their influence on the chiral symmetry of the Dirac operator. Afterwards we will apply the concept of improvement to the fermion action.

2.2.1 Wilson fermions

Wilson [1] introduced the discretized Dirac operator

$$D'_{W,f} = \frac{1}{2} \sum_{\mu=0}^3 (\gamma_{\mu}(\nabla_{\mu}^* + \nabla_{\mu}) - a\nabla_{\mu}^* \nabla_{\mu}) + m_{0,f}, \quad (2.18)$$

with the covariant forward and backward derivatives ∇_{μ} and ∇_{μ}^* and the gamma matrices γ_{μ} defined in appendix A. The bare quark mass $m_{0,f}$ of a fermion flavor f can be expressed in terms of the hopping parameter κ_f via

$$m_{0,f} = \frac{1}{2a} \left(\frac{1}{\kappa_f} - 8 \right) \quad (2.19)$$

and we will frequently use this form. After an appropriate re-scaling of the fields we can parametrize the Dirac operator by the hopping parameter and write

$$D'_{W,f} = \mathbf{1} - \kappa_f \sum_{\mu=0}^3 \left[(1 - \gamma_{\mu}) U_{\mu}(n) \delta_{n, n+a\hat{\mu}} + \delta_{n, n-a\hat{\mu}} (1 + \gamma_{\mu}) U_{\mu}^{\dagger}(n) \right]. \quad (2.20)$$

In this way, we see explicitly that the Dirac operator only includes self- and nearest-neighbor interactions. The discretized fermionic action can now be constructed from the Dirac operator and the spinors $\psi, \bar{\psi}$ via

$$S'_{W,f} = a^4 \sum_n \bar{\psi}(n) D'_{W,f} \psi(n). \quad (2.21)$$

We briefly introduced the Nielsen-Ninomiya no-go theorem at the beginning of this chapter and can now discuss its implications on D'_{W} . In principle, only the first term of eq. (2.18) alone is a straight-forward lattice version of the continuum Dirac operator of a massless quark. A Dirac operator which consists only of this term is local, shows the correct continuum limit and respects chiral symmetry which is a fundamental symmetry of QCD with massless quarks. An investigation of the Fourier transform of the free Dirac operator constructed from the first term shows that \hat{D}'_{W} in four dimensions vanishes at the 2^4 corners of the Brillouin zone [11]. A discretized theory defined from this Dirac operator has 2^4 flavors of Dirac fermions in the continuum limit.

Wilson circumvented this problem by the introduction of the second term which is an irrelevant operator since it vanishes in the continuum. It adds a term of order $O(a^{-1})$ to the Fourier transform which vanishes for $p_{\mu} = 0$, but is non-vanishing for $ap_{\mu} = \pi$. The unphysical doublers receive an infinitely large contribution in the continuum limit and decouple from the theory.

The restoration of the correct continuum limit using Wilsons formulation of the Dirac operator has a downside. At finite lattice spacing, the action including $-a\nabla_\mu^*\nabla_\mu$ is not invariant under the $SU(N_f)$ chiral symmetry of massless QCD. We will discuss consequences of the broken chiral symmetry concerning the renormalization of quark masses in chapter 4.

The bare quark mass $m_{0,f}$ vanishes in massless QCD with preserved chiral symmetry.

2.2.2 Improved Wilson fermions

The inclusion of the Wilson term in the Dirac operator adds an explicit term of order $O(a)$ to the fermion action. In contrast to pure Yang-Mills theory, we are faced with discretization errors that vanish only linearly towards the continuum limit. As in the case of the gluon action, it is possible to determine all operators which belong to the next-to-leading Lagrangian in eq. (2.9).

Since this next-to-leading term is of order $O(a)$, we search for operators of order five which are gauge invariant and respect all symmetries of the lattice theory. It turns out that the Lagrangian \mathcal{L}_1 has to be a linear combination of the following five operators [31]

$$\begin{aligned}\mathcal{O}_1 &= \bar{\psi}\sigma_{\mu\nu}F_{\mu\nu}\psi \\ \mathcal{O}_2 &= \bar{\psi}D_\mu D_\mu\psi + \bar{\psi}\overleftarrow{D}_\mu\overleftarrow{D}_\mu\psi \\ \mathcal{O}_3 &= m\text{Tr}[F_{\mu\nu}F_{\mu\nu}] \\ \mathcal{O}_4 &= m\left(\bar{\psi}\gamma_\mu D_\mu\psi - \bar{\psi}\overleftarrow{D}_\mu\gamma_\mu\psi\right) \\ \mathcal{O}_5 &= m^2\bar{\psi}\psi\end{aligned}\tag{2.22}$$

The left action of the Dirac operator is defined in appendix A.

which are based on the gauge and spinor fields. This set can be reduced to three operators. We want to achieve on-shell improvement and this means that correlation functions, in which all arguments are separated by physical distances, are improved. In this case we can use the field equations $(\gamma_\mu D_\mu + m)\psi = 0$, to show that the operators in eq. (2.22) are linearly dependent. Using the relations

$$\mathcal{O}_1 - \mathcal{O}_2 + 2\mathcal{O}_5 = 0\tag{2.23}$$

$$\mathcal{O}_4 + 2\mathcal{O}_5 = 0\tag{2.24}$$

On-shell quantities are, e.g., masses, decay constants and physical matrix elements.

we can eliminate \mathcal{O}_2 and \mathcal{O}_4 from the set of operators [32]. The $O(a)$ Lagrangian of the effective theory is therefore a linear combination of \mathcal{O}_1 , \mathcal{O}_3 and \mathcal{O}_5 , with appropriately chosen coefficients $c_i(g_0^2)$. When we recall the form of the gauge action in eq. (2.7), we see that a term proportional to \mathcal{O}_3 is already present in an action including quarks and gluons. The same is true for \mathcal{O}_5 , where the corresponding term is part of the Wilson Dirac operator. The introduction of these two operators can therefore be done by a redefinition of the bare parameters g_0^2 and m_0 which anyway undergo renormalization.

We are left with \mathcal{O}_1 as the only operator that we have to add to the action with the appropriate parameter, in order to obtain an action

In their honor, the parameter multiplying \mathcal{O}_1 is denoted by c_{SW} .

with leading corrections of order $O(a^2)$. This was first derived by Sheikholeslami and Wohlert [33]. The modified Dirac operator reads

$$D_{W,f} = \frac{1}{2} \sum_{\mu=0}^3 (\gamma_{\mu}(\nabla_{\mu}^* + \nabla_{\mu}) - a\nabla_{\mu}^* \nabla_{\mu}) + m_{0,f} + a c_{SW} \sum_{\mu,\nu=0}^3 \frac{i}{4} \sigma_{\mu\nu} \hat{F}_{\mu\nu}, \quad (2.25)$$

where the definition of the discretized gluon field strength tensor $\hat{F}_{\mu\nu}$ is chosen as in [31], to be

$$\hat{F}_{\mu\nu}(n) = \frac{1}{8a^2} [Q_{\mu\nu}(n) - Q_{\nu\mu}(n)] \quad (2.26)$$

This name is based on the geometrical form of $Q_{\mu\nu}$.

with the so-called clover plaquettes

$$Q_{\mu\nu}(n) = U_{\mu}(n)U_{\nu}(n+a\hat{\mu})U_{\mu}(n+a\hat{\nu})^{-1}U_{\nu}(n)^{-1} + U_{\nu}(n)U_{\mu}(n-a\hat{\mu}+a\hat{\nu})^{-1}U_{\nu}(n-a\hat{\mu})^{-1}U_{\mu}(n-a\hat{\mu}) + U_{\mu}(n-a\hat{\mu})^{-1}U_{\nu}(n-a\hat{\mu}-a\hat{\nu})^{-1}U_{\mu}(n-a\hat{\mu}-a\hat{\nu})U_{\nu}(n-a\hat{\nu}) + U_{\nu}(n-a\hat{\nu})^{-1}U_{\mu}(n-a\hat{\nu})U_{\nu}(n+a\hat{\mu}-a\hat{\nu})U_{\mu}(n)^{-1}. \quad (2.27)$$

The coefficient c_{SW} has first been determined perturbatively by Sheikholeslami and Wohlert [33]. In our simulations in the three-flavor theory, we use the non-perturbative determination of [34] which is parametrized by

$$c_{SW}(g_0^2) = \frac{1 - 0.1921g_0^2 - 0.1378g_0^4 + 0.0717g_0^6}{1 - 0.3881g_0^2}. \quad (2.28)$$

By using the non-perturbative determination, we achieve an improvement from leading discretization effects of order $O(a)$ to $O(a^2)$.

With this definition of the Dirac operator, we can now specify the action of 2 + 1 flavor QCD which is used in our simulations. The fermion action is written

$$S_W = a^4 \sum_{f=1}^3 \sum_n \bar{\psi}(n) D_{W,f} \psi(n). \quad (2.29)$$

where we now included a sum over the three flavors up, down and strange. We take the masses of the light quarks, up and down, degenerate

$$m_{0,u} \equiv m_{0,d} \equiv m_{0,l} \quad (2.30)$$

in all considered simulations. It has been shown [35] that the effects of the isospin mass splitting between physical up and down quark are of the order of QED effects. As long, as these QED effects are not taken

When several quark flavors are included, the scalar mass term in the Dirac operator is promoted to a diagonal matrix in flavor space.

In our simulations in the Schrödinger functional, we use three mass degenerate quarks.

into account, mass degenerate light quarks are a good approximation to QCD. As we will see in chapter 5, it is cheaper to simulate a pair of mass degenerate quarks than single quarks. On top, we reduce the number of parameters that have to be tuned, in order to obtain the correct results in the continuum limit.

The full action for $N_f = 2 + 1$ flavors of $O(a)$ improved quarks and tree-level improved gluons is then defined by

$$S \equiv S_{LW} + S_W \tag{2.31}$$

with the actions from eqs. (2.17) and (2.29). The only thing left to do is to specify the boundary conditions and we will do this in the next chapter.

THE SCHRÖDINGER FUNCTIONAL

The use of the Schrödinger representation in quantum field theory was brought forward by Symanzik [36], who studied the two-dimensional ϕ^4 theory and showed that the wave functionals which result from the use of the Schrödinger picture can be renormalized to all orders. An introduction to his paper can be found in [37]. Symanzik's conclusion, translated to the language of lattice field theory, implies that the Schrödinger functional converges in the continuum limit.

In our work, we use the Schrödinger functional representation of lattice QCD for finite-volume simulations. Our introduction of the general setup of non-abelian gauge theories in the Schrödinger functional will be along the lines of [38]. The introduction of quarks to the Schrödinger functional was done in [39]. Pedagogical introductions can be found in [40, 41].

The Schrödinger functional is utilized for finite-volume simulations of QCD since its boundary conditions are well suited for perturbative calculations and Monte Carlo simulations. As we will see later when we introduce quark fields in the Schrödinger functional, the boundary conditions also offer the possibility to perform simulations at vanishing renormalized quark masses making it the perfect basis for the determination of renormalization constant in a massless renormalization scheme.

3.1 CONTINUUM FORMULATION

Before we specify the discretized theory that has been used in our simulations, we introduce the Schrödinger functional in the continuum.

3.1.1 Gluon action

We express the matrix elements of the Euclidean time evolution operator $e^{\mathcal{H}T}$, where \mathcal{H} is the Hamilton operator, between two gauge invariant states via a functional integral over all gauge fields $A_\mu(x)$ in four dimensions $\mathbb{T} \times \mathbb{L}^3$ with $0 \leq x_0 \leq T$. In the spatial directions, we assume periodic boundary conditions. In the temporal direction, we impose Dirichlet boundary conditions for the vector potentials,

$$A_k(x) = \begin{cases} C_k^\wedge(x), & \text{at } x_0 = 0 \\ C_k^\prime(x), & \text{at } x_0 = T \end{cases}, \quad (3.1)$$

where C, C' are classical gauge potentials. From these boundary conditions, we can see that we formulate our theory on a four-dimensional cylinder in space-time. In eq. (3.1), A^Λ denotes the gauge transformation of A which is defined by

$$A_k^\Lambda(x) = \Lambda(x)A_k(x)\Lambda^{-1}(x) + \Lambda(x)\partial_k\Lambda^{-1}(x), \quad \Lambda \in \text{SU}(N). \quad (3.2)$$

We can now define the Schrödinger functional as the Euclidean partition function

$$\mathcal{Z}[C', C] = \int \mathcal{D}[\Lambda] \int \mathcal{D}[A] e^{-S_G[A]} \quad (3.3)$$

with the boundary conditions as stated above. S_G is the Yang-Mills gauge action and the integral measures are defined by

$$\mathcal{D}[A] = \prod_{x, \mu, \alpha} dA_\mu^\alpha(x), \quad \mathcal{D}[\Lambda] = \prod_x d\Lambda(x). \quad (3.4)$$

With these functional integrals, we integrate over all components of the Euclidean field and all gauge transformations. \mathcal{Z} is invariant under gauge transformations of the boundary fields.

So far, we have not defined how to choose the boundary fields C, C' . A proper choice of these fields allows to induce a color background field. In this case, constant Abelian fields are chosen at the boundaries. In our simulations, we only consider the case of vanishing background fields. This is achieved by setting the boundary fields to zero, i.e.,

$$A_k(0, \mathbf{x}) = A_k(T, \mathbf{x}) = 0. \quad (3.5)$$

Symanzik [36] has shown to all orders in perturbation theory that the Schrödinger functional is finite after standard renormalization of the field theory and the addition of boundary counter-terms of dimension $d \leq 3$. Since no gauge invariant local field with $d \leq 3$ exists in Yang-Mills theory, no additional counter term is necessary when the theory is formulated in the Schrödinger functional. A proof to all orders for Yang-Mills theory and QCD is missing, but from one-loop results and various numerical studies it is believed that Symanzik's statement is also true in this case [41].

3.1.2 Quarks in the Schrödinger functional

To study QCD in the Schrödinger functional, we have to formulate the boundary conditions for quark fields. This was done in [39] and we collect the results here. With the definition of the projector

$$P_\pm = \frac{1}{2}(1 \pm \gamma_0) \quad (3.6)$$

In Yang-Mills theory, the integral over all gauge transformations can be reformulated into an integral over all topological classes [38].

which projects onto the upper or lower components of a spinor field, we can formulate the Dirichlet boundary conditions for quark ψ and anti-quark $\bar{\psi}$ fields,

$$\begin{aligned} P_+ \psi(0, \mathbf{x}) &= \rho, & \bar{\psi}(0, \mathbf{x}) P_- &= \bar{\rho}, \\ P_- \psi(T, \mathbf{x}) &= \rho', & \bar{\psi}(T, \mathbf{x}) P_+ &= \bar{\rho}'. \end{aligned} \quad (3.7)$$

These boundary conditions have to be formulated for every quark flavor.

Since the Dirac operator is a first order differential operator, the specification of boundary conditions for half of the components is sufficient. The partition sum of the Schrödinger functional now includes integrals over the quark and anti-quark fields,

$$\mathcal{Z}[C', \rho', \bar{\rho}', C, \rho, \bar{\rho}] = \int \mathcal{D}[A] \int \mathcal{D}[\psi] \mathcal{D}[\bar{\psi}] e^{-S[A, \psi, \bar{\psi}]} \quad (3.8)$$

and the action has been extended by the fermion action S_F ,

$$\begin{aligned} S[A, \psi, \bar{\psi}] &= S_G[A] + S_F[A, \psi, \bar{\psi}] \quad (3.9) \\ S_F[A, \psi, \bar{\psi}] &= \int d^4x \bar{\psi}(x) [\gamma_\mu D_\mu + m] \psi(x) \\ &\quad - \int d^3x [\bar{\psi}(x) P_- \psi(x)]|_{x_0=0} \\ &\quad - \int d^3x [\bar{\psi}(x) P_+ \psi(x)]|_{x_0=T}. \end{aligned} \quad (3.10)$$

Compared to the usual formulation of this action, we have added two terms with integrations over the spatial components at the time boundaries. From this form we see that Schrödinger functional boundary conditions affect the action only directly at the boundaries. The bulk is unchanged compared to (anti-) periodic boundary conditions in the time direction. The dependence of the Schrödinger functional on the boundary conditions is implicit in the last two terms of the fermion action.

As for the pure Yang-Mills action, we have to ask ourselves, whether the boundaries change the renormalization pattern. With the inclusion of quark and anti-quark fields, it is now possible to construct gauge invariant counter terms with dimension $d = 3$. The inclusion of these counter terms amounts to a multiplicative renormalization of the boundary fields $\rho, \rho', \bar{\rho}, \bar{\rho}'$ and if we choose to set all of them to zero, no further renormalization has to be done in addition to the case without boundaries [39].

We set these fields to zero in all our simulations.

Another, maybe even more important, feature of the Schrödinger functional was derived in [39]. In small volume, the boundaries effect the spectrum of the Dirac operator and induce a gap below the smallest eigenvalue. This leads to a significantly reduced condition number of the Dirac operator and allows to simulate quarks with vanishing mass. If one has in mind that the simulation of masses as small as the physical light quark masses is very expensive in large volume,

For the free and massless Dirac operator, the smallest eigenvalue is $(\pi/2T)^2$.

We will have a look at the effect of a reduced condition number in chapter 5.

massless quarks are an astonishing feature of the Schrödinger functional. Setting the quark masses to zero, after an appropriate tuning of the hopping parameter, allows to investigate mass-independent renormalization patterns on the lattice.

These phase angles can also be implemented in the definition of the covariant derivative. In this case, periodic boundary conditions are used.

So far, we have not specified the boundary conditions of the (anti-)quark fields in the space directions. We can choose them to be periodic, as we did it for the gauge fields, but the most general choice for the spinors includes the phase angles $\theta_k, k = 1, 2, 3$ via

$$\psi(x + L\hat{k}) = e^{i\theta_k}\psi(x), \quad \bar{\psi}(x + L\hat{k}) = e^{-i\theta_k}\bar{\psi}(x), \quad (3.11)$$

where \hat{k} is the unit vector in direction k . Periodic boundary conditions are recovered from $\theta_k = 0$. We take all three theta angles to be the same, $\theta_k = \theta$. It was shown to one-loop in [43] that setting $\theta = \pi/5$ increases the spectral gap in the Dirac operator by about a factor of two. A similar behavior is expected in our simulations.

In our simulations, we use $\theta = 0$ or $\theta = 0.5$ [42] for the sea quarks, depending on the project.

3.2 LATTICE FORMULATION

When we formulate the Schrödinger functional on the lattice, we can use the results derived in chapter 2 for the bulk and only have to define how we proceed at and beyond the boundaries. It is convenient to extend the fields formally beyond the boundaries and to set the spinor fields in this region to zero, i.e.,

$$\psi(x) = 0 = \bar{\psi}(x) \quad \text{if } x_0 < 0 \text{ or } x_0 > T, \quad (3.12)$$

and

$$P_- \psi(0, x) = 0 = P_+ \psi(T, x), \quad (3.13)$$

$$\bar{\psi}(0, x) P_+ = 0 = \bar{\psi}(T, x) P_- . \quad (3.14)$$

The gauge field variables beyond the boundary are set to unity. In this way it is possible to write the action as sum over all space-time points without any constraint for the time coordinate [41].

In the previous section we have argued that the renormalization pattern does not change, when boundaries are included. If we formulate the theory on lattice, we are faced with $O(a)$ effects near the boundaries and, as in the bulk, we like to improve the leading cut-off effects [38]. For the gauge fields, there are different ways to handle the loops at the boundaries and the specific choice decides, how the improvement has to be done. In our work, we stick to choice B of [44] which we summarize here.

We add weights to the loops in the Lüscher-Weisz gauge action, eq. (2.17), and write

$$S_{LW}^{\text{SF}}(\mathbf{U}) = \frac{\beta}{6} \left(c_0 \sum_{\mathbf{p}} w_0(\mathbf{p}) \text{Tr} [\mathbf{1} - \mathbf{U}(\mathbf{p})] + c_1 \sum_{\mathbf{r}} w_1(\mathbf{r}) \text{Tr} [\mathbf{1} - \mathbf{U}(\mathbf{r})] \right). \quad (3.15)$$

Within option B, we take over the notation of [45] and define the weights to be

$$w_0(p) = \begin{cases} 1/2, & \text{all links in } p \text{ are on the time boundary} \\ c_t(g_0^2), & p \text{ has one link on the time boundary} \\ 1, & \text{otherwise} \end{cases}, \quad (3.16)$$

and

$$w_1(r) = \begin{cases} 1/2, & \text{all links in } r \text{ are on the time boundary} \\ 3/2, & r \text{ has two links on the time boundary} \\ 1, & \text{otherwise} \end{cases}. \quad (3.17)$$

In this way, we have one free coefficient, $c_t(g_0^2)$ which has to be tuned, to achieve $O(a)$ improvement. Unfortunately, no non-perturbative determination of this coefficient exists and we have to stick to tree-level or one-loop perturbative improvement. For the one-loop improvement, we use formula (3.6) of [45],

$$c_t(g_0^2) = 1 + c_t^{(1)} g_0^2 + O(g_0^4). \quad (3.18)$$

In the following we document how this one-loop coefficient has been derived from the literature. From [43], we get the decomposition of the one-loop coefficient into a gauge and a fermion part,

$$c_t^{(1)} = c_t^{(1,0)} + N_f c_t^{(1,1)} \quad (3.19)$$

and from [43] and [46], we find

$$c_t^{(1,0)} = A_1/2 \quad \text{with} \quad A_1 = -0.005940(2) \quad (3.20)$$

$$c_t^{(1,1)} = 0.0191410(1). \quad (3.21)$$

This one-loop coefficient is the correct one, if a corresponding one-loop coefficient for the rectangles $c_t^{R(1)} = 2c_t^{(1)}$ is included in the weight in eq.(3.17), for r with two links on the time boundary. Since we set $w_1(r)$ to the tree-level weight $3/2$ and therefore choose $c_t^{R(1)} = 0$, we use the first equation in (3.13) of [46] to obtain

$$c_t^{(1)} = \frac{c_t^{(1)}}{c_0} = \frac{c_t^{(1,0)}}{c_0} + N_f \frac{c_t^{(1,1)}}{c_0} \quad (3.22)$$

which was used together with $N_f = 3$ to derive eq. (3.18).

A similar boundary improvement has to be performed for the spinor fields. The inclusion of the corresponding operators [31], effectively amounts to adding a term to the $O(a)$ improved Wilson Dirac operator (2.25),

$$D_{W,f}^{\text{SF}} = D_{W,f} + \delta D_{\text{bnd}} \quad (3.23)$$

We have used both versions in different projects and will specify the action in the corresponding chapter.

and its action on a quark field is defined as [45]

$$\delta D_{\text{bnd}}\psi(x) = (\tilde{c}_t - 1)\frac{1}{a}(\delta_{x_0, a} + \delta_{x_0, T-a})\psi(x). \quad (3.24)$$

As it can be seen the addition of the term in eq. (3.24) only affects the time slices at the boundary. In the bulk of the lattice, the action remains the same. For \tilde{c}_t , we use the formula from eq. (3.6) in [45],

$$\tilde{c}_t = 1 + \tilde{c}_t^{(1)}g_0^2 + O(g_0^4), \quad \tilde{c}_t^{(1)} = -0.01505 \quad (3.25)$$

which is valid for our action and independent of the number of flavors.

The discussion in this section applies to the formulation of the Schrödinger functional using Abelian boundary fields. If the fields are chosen differently, another set of improvement coefficients, the space-like coefficients c_s and \tilde{c}_s , has to be included. The operators which have to be multiplied by these coefficients are defined in [31, 44]. Since we set the boundary fields to zero in our simulations, we do not have to consider them.

3.3 SCHRÖDINGER FUNCTIONAL CORRELATION FUNCTIONS

A slightly different path to define the quark fields at the boundaries which results in the same setup, is taken in [47].

One of the features of simulations in the Schrödinger functional is the possibility to construct sources which are not placed in the bulk of the lattice but directly on the boundaries. These allow for new types of correlation functions.

3.3.1 Boundary sources

We will define a set of sources \mathcal{O}^a as in [48] by differentiating the Schrödinger functional with respect to the boundary values of the quark fields given in eq. (3.7). The boundary (anti-) quark fields are defined via functional derivatives of the Boltzmann factor, by [31]

A definition of the derivatives can be found in Appendix C of [31].

$$\begin{aligned} \zeta(x) &= \frac{\delta}{\delta\bar{\rho}(x)}, & \bar{\zeta}(x) &= -\frac{\delta}{\delta\rho(x)}, \\ \zeta'(x) &= \frac{\delta}{\delta\bar{\rho}'(x)}, & \bar{\zeta}'(x) &= -\frac{\delta}{\delta\rho'(x)}. \end{aligned} \quad (3.26)$$

Inside the Schrödinger functional, these derivatives which are taken before the boundary fields are set to zero have the effect of inserting (anti-) quark fields close to the boundaries. The operators \mathcal{O}^a are then defined by

$$\mathcal{O}^a = \int_0^L d^3y d^3z \zeta(y) \gamma_5 \frac{1}{2} \tau^a \zeta(z) \quad (3.27)$$

and due to the integrals in the three space directions (which become sums on the lattice), these boundary fields are projected onto their zero momentum components. The Pauli matrices τ^a are defined in Appendix A.

3.3.2 Expectation values

The expectation value of any product of fields \mathcal{O} in the Schrödinger functional on the lattice is defined by [31]

$$\langle \mathcal{O} \rangle = \left\{ \frac{1}{Z} \int \mathcal{D}[U] \mathcal{D}[\psi] \mathcal{D}[\bar{\psi}] \mathcal{O} e^{-S[U, \psi, \bar{\psi}]} \right\}_{\rho=\rho'=\bar{\rho}=\bar{\rho}'=0} \quad (3.28)$$

and we see how the derivatives in the boundary operators, eq. (3.26), act on the exponential before the boundary values are set to zero.

3.3.3 Correlation functions

Based on the boundary fields and operator insertions in the bulk of the lattice, we are now able to construct three different types of correlation functions:

- boundary-to-boundary correlation functions, where a state is created at $x_0 = 0$ and propagates to the other boundary at $x_0 = T$ where it is annihilated,
- forward boundary-to-bulk correlation functions, where the state is created at $x_0 = 0$ and is annihilated in the bulk,
- backwards boundary-to-bulk correlation functions, where the state is created at $x_0 = T$ and is annihilated in the bulk.

In [49], it is shown that any correlation function of bulk and boundary quark fields can be calculated, when Wick's theorem with the appropriate two-point contractions is used.

We define the lattice boundary operators

$$\begin{aligned} \mathcal{O}^a &= \frac{a^6}{L^3} \sum_{y,z} \bar{\zeta}(y) \gamma_5 \frac{1}{2} \tau^a \zeta(z), & \mathcal{O}'^a &= \frac{a^6}{L^3} \sum_{y,z} \bar{\zeta}'(y) \gamma_5 \frac{1}{2} \tau^a \zeta'(z), \\ \mathcal{O}_k^a &= \frac{a^6}{L^3} \sum_{y,z} \bar{\zeta}(y) \gamma_k \frac{1}{2} \tau^a \zeta(z), & \mathcal{O}'_k{}^a &= \frac{a^6}{L^3} \sum_{y,z} \bar{\zeta}'(y) \gamma_k \frac{1}{2} \tau^a \zeta'(z), \end{aligned} \quad (3.29)$$

where the sums are taken over all sites on the boundary time slice, $k = 1, 2, 3$. In combinations with the bulk operators

$$X^a(x) = \bar{\psi}(x) \Gamma \frac{1}{2} \tau^a \psi(x) \quad (3.30)$$

we can construct Schrödinger functional correlation functions. In the definition of bulk operators in eq. (3.30) X is a symbol for the scalar, vector, tensor, pseudo-scalar or axial-vector currents and the corresponding Γ -combinations are

$$\begin{aligned} S : \Gamma &= \mathbf{1}, & V_\mu : \Gamma &= \gamma_\mu, & T_{\mu\nu} : \Gamma &= \gamma_\mu \gamma_\nu, \\ P : \Gamma &= \gamma_5, & A_\mu : \Gamma &= \gamma_\mu \gamma_5. \end{aligned} \quad (3.31)$$

We will use the same combinations of gamma matrices in chapter 14.

We now introduce the two-point correlation functions used in our projects with conventions based on [50]. We start with the boundary-to-boundary correlation functions

$$f_1 \equiv -\frac{1}{2}\langle\mathcal{O}'^a\mathcal{O}^a\rangle, \quad k_1 \equiv -\frac{1}{6}\langle\mathcal{O}'_k^a\mathcal{O}_k^a\rangle, \quad (3.32)$$

continue with the forward boundary-to-bulk functions

$$f_A(x_0) \equiv -\frac{a^3}{2}\sum_x\langle A_0^a(x)\mathcal{O}^a\rangle, \quad (3.33)$$

$$f_P(x_0) \equiv -\frac{a^3}{2}\sum_x\langle P^a(x)\mathcal{O}^a\rangle, \quad (3.34)$$

$$k_V(x_0) \equiv -\frac{a^3}{6}\sum_x\langle V_k^a(x)\mathcal{O}_k^a\rangle, \quad (3.35)$$

$$k_T(x_0) \equiv -\frac{a^3}{6}\sum_x\langle T_{k0}^a(x)\mathcal{O}_k^a\rangle, \quad (3.36)$$

and conclude with the backwards boundary-to-bulk functions

$$g_A(T-x_0) \equiv -\frac{a^3}{2}\sum_x\langle\mathcal{O}'^a A_0^a(x)\rangle, \quad (3.37)$$

$$g_P(T-x_0) \equiv -\frac{a^3}{2}\sum_x\langle\mathcal{O}'^a P^a(x)\rangle, \quad (3.38)$$

$$l_V(T-x_0) \equiv -\frac{a^3}{6}\sum_x\langle\mathcal{O}'_k^a V_k^a(x)\rangle, \quad (3.39)$$

$$l_T(T-x_0) \equiv -\frac{a^3}{6}\sum_x\langle\mathcal{O}'_k^a T_{k0}^a(x)\rangle. \quad (3.40)$$

In all cases, the sum over isospin indices a and spatial indices k is implied. To smooth out statistical fluctuations, we can use the time-reversal symmetry and average over forward and backward boundary-to-bulk correlation functions with the same Γ -structure.

For the computation of the renormalization constant Z_A , we also need four-point boundary-to-boundary correlation functions with two insertions in the bulk. We define them for general Γ -structures X, Y as

$$F_{XY}(x_0, y_0) = -\frac{a^6}{6}\sum_{x,y}\epsilon^{abc}\epsilon^{cde}\langle\mathcal{O}'^d X^a(x)Y^b(y)\mathcal{O}^e\rangle, \quad (3.41)$$

with the totally antisymmetric tensor ϵ^{abc} .

Most computations of Schrödinger functional correlation functions in this work have been done with the `sfcf` program [52], some with a modified version of `openQCD` [53].

*These functions have
been introduced in
[51].*

3.3.4 Flavor off-diagonal bilinears

In many applications, we consider off-diagonal bilinears when we work with currents based on different quark flavors. In this case we switch to the basis τ^+, τ^-, τ^3 of the flavor $SU(2)$ symmetry which is defined in eq. (A.5).

We consider quarks of different flavors $i, j = 1, 2, \dots$ which only differ by their mass. If we write the bispinors as $\psi(x) = (\psi^i(x), \psi^j(x))^T$, we can use eq. (3.30) to define the operators

$$X^\pm = X^1 \pm iX^2 \quad (3.42)$$

and express them in terms of the components ψ^i via

$$X^+ = \bar{\psi}^i(x) \Gamma \psi^j(x) \equiv X^{ij}, \quad (3.43)$$

$$X^- = \bar{\psi}^j(x) \Gamma \psi^i(x) \equiv X^{ji}. \quad (3.44)$$

The same construction applies for the boundary operators and the correlation functions which we have defined above. We will explicitly indicate the use of this basis by using the corresponding indices as superscript. We do not consider flavor diagonal bilinears in this work.

Sometimes, we also consider quarks of different flavor but equal mass.

3.4 ADDENDUM: OPEN BOUNDARY CONDITIONS

Closely connected to Schrödinger functional boundary conditions, in the sense that the theoretical and practical setup is done similarly, are the so-called open boundary conditions. Since these are used in the large-volume simulations which we use to extract physical observables as the mass of the charm quark, we introduce them in this section.

3.4.1 Motivation

With the advance of simulations with dynamical quarks towards smaller and smaller lattice spacings, a problem became apparent which is connected to the topological properties of QCD on a $T \times L^3$ torus. In the continuum theory, disconnected topological sectors emerge. On the lattice, with non-zero lattice spacing, there is a finite probability to tunnel between these sectors. Since we want to sample all sectors according to their weight in the path integral this tunneling is a necessary feature. Towards the continuum limit, the probability to change the topological sector, i.e., to pass the topological barrier, is rapidly decreasing [54–57].

This is problematic for two reasons. If we are not able to sample all sectors appropriately in a finite simulation time, we lose the theoretical control over the predictions made from our simulations. In addition, the suppression of changes of the topological charge, the so-called freezing, leads to extremely long-ranged autocorrelations

We will look at the statistical analysis of autocorrelation times in section 5.4.

of the charge. These correlations can also influence and increase the autocorrelation times of other observables.

The freezing of the topological charge is overcome by the introduction of open boundary conditions which allow the charge to “flow out of the lattice”. It was shown that the theory with open boundary conditions is equivalent to other setups since the transfer matrix is unchanged, when the open boundary conditions are introduced [58, 59]. The corresponding Wilson-Dirac operator is γ_5 -hermitian.

3.4.2 Setup

As in the Schrödinger functional, we have to specify boundary conditions for the gauge, quark and antiquark fields at the boundaries at $x_0 = 0$ and $x_0 = T$. All fields satisfy periodic boundary conditions in the space directions. Formally, the open boundary conditions are defined by

$$F_{0k}(0, \mathbf{x}) = F_{0k}(T, \mathbf{x}) = 0 \quad \text{for all } k = 1, 2, 3 \quad (3.45)$$

with the gauge-field tensor

$$F_{\mu\nu}(\mathbf{x}) = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \quad (3.46)$$

for the gauge fields and

$$P_+ \psi(0, \mathbf{x}) = P_- \psi(T, \mathbf{x}) = 0 \quad (3.47)$$

$$\bar{\psi}(0, \mathbf{x}) P_- = \bar{\psi}(T, \mathbf{x}) P_+ = 0 \quad (3.48)$$

for the quark and anti-quark fields [58]. The latter are the same as for the Schrödinger functional with vanishing background field. For the gauge fields, we switch from Dirichlet to Neumann boundary conditions, compared to the Schrödinger functional. All statements regarding the renormalizability of the theory in the treatment of the Schrödinger functional hold for open boundary conditions.

Open boundary conditions emerge in the continuum limit, if we simply set all links that stick out of the lattice to zero [58]. These are the links in negative time direction at $x_0 = 0$ and in positive time direction at $x_0 = T$.

3.4.3 Boundary improvement

The boundary improvement with open boundary conditions is similar to the case of Schrödinger functional boundary conditions and proceeds via the same sets of operators that have to be added at the boundaries. The improvement coefficients themselves differ from the ones in the Schrödinger functional, since the boundaries are set up differently.

For the CLS simulations, the weights in the Lüscher-Weisz action defined in eq. (2.17) are chosen to be

$$w_0(p) = \begin{cases} \frac{1}{2}c_G, & \text{all links in } p \text{ are on the time boundary} \\ 1, & \text{otherwise} \end{cases}, \quad (3.49)$$

and

$$w_0(r) = \begin{cases} \frac{1}{2}c_G, & \text{all links in } r \text{ are on the time boundary} \\ 1, & \text{otherwise} \end{cases}. \quad (3.50)$$

The improvement coefficient is set to its tree-level value $c_G = 1$, since higher orders or non-perturbative determinations are not available. This choice amounts to a weight of $1/2$ for all spatial loops on the time boundaries.

The improvement coefficient for the quark fields with open boundary conditions which is called c_F , enters as \bar{c}_t in eq. (3.24). It is set to its tree-level value $c_F = 1$ in the CLS simulations [60]. The perturbative improvement at tree-level leaves us with boundary effects of order $O(ag_0^2)$.

OPERATOR IMPROVEMENT AND RENORMALIZATION

We looked at the Symanzik improvement program [36] and its implications on our setup of discretized gluon and fermion actions in chapter 2. In this chapter, we will consider the implications of improvement on lattice operators and investigate how improvement is connected with renormalization.

The basis for our discussion was laid in [31, 48, 61, 62]. A detailed analysis of the three-flavor theory has been done in [63]. Everything that is stated in this chapter is explained in great detail in these references and we try to focus on the key points which will be most important in the context of this work.

Within this chapter, we work to order $O(a)$. This means that we neglect all higher order terms starting at $O(a^2)$ without further notice.

4.1 IMPROVEMENT OF THE ACTION: REVISITED

We begin this chapter with a discussion of some implications of the discretization of the action according to eqs. (2.17) and (2.29) on the improvement of operators.

4.1.1 Mass renormalization

In the continuum theory, chiral symmetry is restored if all bare quark masses vanish. Based on a discussion of chiral Ward identities one can infer that quark masses only have to be renormalized multiplicatively and not additively. Therefore, the renormalized quark mass has to vanish when the bare quark mass vanishes and we can write

$$m_R = Z_m m_0. \quad (4.1)$$

As explained when we introduced the Wilson Dirac operator, chiral symmetry is broken explicitly by the introduction of the Wilson term, even at vanishing bare quark mass. This leads to a finite renormalized quark mass at vanishing bare mass. To restore the definition of a multiplicatively renormalized quark mass, the bare subtracted quark mass

$$m_{q,f} = m_{0,f} - m_{cr} \equiv \frac{1}{2a} \left(\frac{1}{\kappa_f} - \frac{1}{\kappa_{cr}} \right) \quad (4.2)$$

is introduced. The critical quark mass m_{cr} parametrizes the shift of the chiral point by the introduction of the Wilson term. Since

We suppress the explicit dependence of m_{cr} on g_0^2 .

this critical quark mass is not known a priori, the breaking of chiral symmetry introduces the need for an extensive tuning procedure in order to simulate massless quarks. As apparent from eq. (4.2), we can also parameterize the bare subtracted quark mass by the hopping parameter κ_f and the critical hopping parameter κ_{cr} .

For the more general case of N_f flavors of quarks with non-degenerate masses, we define the quark mass matrix via

$$M_q = \text{diag}(m_{q,1}, m_{q,2}, \dots, m_{q,N_f}). \quad (4.3)$$

The renormalization pattern is even more involved in this case. We will come back to the full pattern later in this chapter.

4.1.2 Improvement of the coupling

When we introduced improved Wilson fermions in section 2.2.2, we have seen that the inclusion of the operators \mathcal{O}_3 and \mathcal{O}_5 amounts to a redefinition of the bare mass and coupling. This does not pose any problem, as long as a mass-dependent renormalization scheme is used. In this work, as it is done in the references stated at the beginning of this chapter, we use a mass-independent scheme. As a result, a re-parametrization of the bare theory has an impact on the improvement. Instead of the bare coupling g_0^2 , we consider the effective bare coupling

$$\tilde{g}_0^2 = g_0^2 \left(1 + a b_g \frac{\text{Tr}[M_q]}{N_f} \right) \quad (4.4)$$

which is then used to define a renormalized coupling via [31]

$$g_R^2 = Z_g(\tilde{g}_0^2, a\mu) \tilde{g}_0^2. \quad (4.5)$$

Here, we introduced the renormalization constant for the coupling Z_g which depends on the effective coupling and the scale μ . The parameter b_g vanishes at tree-level and the one-loop coefficient is known from [31]. If the bare coupling is held fixed and $\text{Tr}[M_q]$ is varied, the effective coupling and the lattice spacing change. We take different approaches to circumvent this behavior, depending on the simulation setup. Therefore, we do not have to further specify $b_g(g_0^2)$.

4.2 OPERATOR IMPROVEMENT

When we improve local operators on the lattice, we proceed as in the improvement of the action. For order $O(a)$ improvement, we have to search for gauge invariant dimension five operators with the same quantum numbers as the original operator \mathcal{O} which are not canceled by the field equations. These operators can be split into two classes. Operators that amount to a redefinition of the bare local operator are

handled in the next subsection. Operators \mathcal{P} that are not present in the unimproved definition are explicitly included. In general, we write

$$\mathcal{O}_I^{ij} \equiv \mathcal{O}^{ij} + a c_{\mathcal{O}} \mathcal{P}^{ij} \quad (4.6)$$

and introduce a parameter $c_{\mathcal{O}}$ that has to be determined non-perturbatively in order to achieve full $O(a)$ improvement.

4.2.1 Renormalization of improved operators

In [63] the complete pattern for the renormalization of improved operators based on the gamma structures of eq. (3.31) is given. There, the cases of flavor diagonal, flavor off-diagonal and flavor singlet operators are treated separately. In this work, we only consider non-singlet, flavor off-diagonal operators and the general renormalization pattern of these operators is given by

$$\mathcal{O}_R^{ij} = Z_{\mathcal{O}} [1 + a b_{\mathcal{O}} m_{q,ij} + a \bar{b}_{\mathcal{O}} \text{Tr}[M_q]] \mathcal{O}_I^{ij}, \quad (4.7)$$

with definition (4.6) for the improved operator and

$$m_{q,ij} \equiv \frac{1}{2} (m_{q,i} + m_{q,j}). \quad (4.8)$$

In eq. (4.7) we introduced another pair of improvement coefficients. Corresponding formulas for eqs. (4.6) and (4.7) with parameters $\bar{c}_{\mathcal{O}}, d_{\mathcal{O}}, \bar{d}_{\mathcal{O}}$ and $Z_{\mathcal{O}}$ exist for singlet operators. To achieve a full $O(a)$ improvement of the theory, a large amount of work and computer time has to be invested to determine all of the necessary coefficients. The renormalization pattern simplifies significantly in the chiral limit.

4.2.2 Improvement of local currents

Since we will need them in our work, we specify the improvement of the local currents, defined in eq. (3.31). Following [31], we arrive at

$$P_I^{ij} = P^{ij} \quad (4.9)$$

$$(A_I)_{\mu}^{ij} = A_{\mu}^{ij} + a c_A \partial_{\mu} P^{ij} \quad (4.10)$$

$$S_I^{ij} = S^{ij} \quad (4.11)$$

$$(V_I)_{\mu}^{ij} = V_{\mu}^{ij} + a c_V \partial_{\nu} T_{\mu\nu} \quad (4.12)$$

$$(T_I)_{\mu\nu}^{ij} = T_{\mu\nu}^{ij} + a c_T [\partial_{\mu} V_{\nu}^{ij} - \partial_{\nu} V_{\mu}^{ij}]. \quad (4.13)$$

We now give an overview of the status of the non-perturbative determinations of the necessary improvement coefficients and renormalization constants in QCD with three quark flavors. The scale dependent renormalization constant $Z_P(g_0^2, \mu)$ has been determined for the hadronic scale of the CLS ensembles in [64]. For Z_A there exist two different determinations. The first one was performed in the Schrödinger

functional [65], the second one in the chirally rotated Schrödinger functional [66], where a significantly better precision could be achieved. In [66], values for the renormalization constant are provided at the bare coupling values of the CLS simulations. An interpolation formula is given in eq. (C.5) of that reference. The improvement factor c_A has been determined in [67] and its parametrization is given by

$$c_A = -0.006033 g_0^2 \left[1 + \exp \left(p_0 + \frac{p_1}{g_0^2} \right) \right], \quad (4.14)$$

with $p_0 = 9.2056, \quad p_1 = -13.9847$

The interpolation formula is constrained by data in the region $g_0^2 \in [3.3, 3.81]$ and by perturbation theory for $g_0^2 \rightarrow 0$.

The vector current renormalization constant Z_V and the improvement coefficient c_V have been determined on the CLS ensembles in [68] and a determination in the Schrödinger functional is in progress [69]. The same is true for $Z_T(g_0^2, \mu)$ and c_T [70].

As given by eq. (4.7), the renormalization of operators in the improved theory requires the determination of the b and \bar{b} coefficients. A preliminary account of the non-perturbative determination of many of these coefficients using the coordinate space method on CLS ensembles has been given in [71, 72]. The vector improvement coefficients have also been determined in [68] and [73]. In our work, we have determined b_S and the combination $b_A - b_P$ non-perturbatively [74]. We will see later why this particular combination is useful.

$b_S = -2b_m$ [63].

In principle, all of these coefficients depend on \tilde{g}_0^2 rather than on g_0^2 . Since the redefinition of the coupling affects the b -terms at $O(\alpha^2)$, we do not have to differentiate between the different definitions. However, this is different in the determination of renormalization constants.

4.3 IMPROVED QUARK MASSES

Our main objective is not the calculation of currents but rather the determination of physical observables built from them. The renormalization and improvement pattern from this section can then be applied to the observable constructed from the currents. This applies, e.g., to current quark masses which we will introduce in the next section.

4.3.1 Renormalization of the bare subtracted quark mass

When we introduced the bare subtracted quark masses in eq. (4.2), we did not specify the renormalization pattern. Following [63], we start with the renormalization of non-singlet mass combinations

The Gell-Mann matrices λ are defined in app. A.

$$\text{Tr} [\lambda M_q]_R = Z_m \left[(1 + \alpha \bar{b}_m \text{Tr}[M_q]) \text{Tr} [\lambda M_q] + \alpha b_m \text{Tr} [\lambda M_q^2] \right]. \quad (4.15)$$

The notation $\text{Tr}[\lambda\mathcal{O}]$ where the trace is taken over all flavor indices and λ is one of the Gell-Mann matrices includes all non-singlet bilinears. The corresponding formula for the singlet mass combinations is

$$\text{Tr}[M_q]_R = Z_m r_m \left[(1 + a\bar{d}_m \text{Tr}[M_q]) \text{Tr}[M_q] + a d_m \text{Tr}[M_q^2] \right], \quad (4.16)$$

where we expressed the singlet renormalization constant by $r_m = \bar{Z}_m/Z_m$. The renormalization constant Z_m is scale dependent but the ratio r_m is not.

With these two formulas and the condition of a diagonal mass-matrix, we can express the renormalized subtracted quark mass via [63],

$$\begin{aligned} m_{i,R} \equiv Z_m \left\{ \left[m_{q,i} + (r_m - 1) \frac{\text{Tr}[M_q]}{N_f} \right] + a B_i \right\}, \\ B_i = b_m m_{q,i}^2 + \bar{b}_m m_{q,i} \text{Tr}[M_q] \\ + (r_m d_m - b_m) \frac{\text{Tr}[M_q^2]}{N_f} + (r_m \bar{d}_m - \bar{b}_m) \frac{\text{Tr}[M_q]^2}{N_f}. \end{aligned} \quad (4.17)$$

The term B_i vanishes in the chiral limit but the term proportional to the ratio of singlet to non-singlet renormalization constant remains. This means that the chiral point where the renormalized quark mass vanishes receives an additional shift. The origins and physical interpretations of the different terms in eq. (4.17) are explained in detail in section 2 of our publication [74].

4.3.2 The PCAC relation

The Lagrange density of continuum QCD is invariant under a number of symmetry transformations. Some symmetries hold only approximately or in a certain limit of the theory. In this section, we consider the variation of the theory under a chiral rotation of the quark fields

$$\delta\psi(x) = \omega^\alpha(x) \frac{1}{2} \tau^\alpha \gamma_5 \psi(x), \quad \delta\bar{\psi}(x) = \omega^\alpha(x) \bar{\psi}(x) \gamma_5 \frac{1}{2} \tau^\alpha, \quad (4.18)$$

where we assume to have an isospin doublet of quarks with equal mass m . In these variations, ω^α is a smooth function which vanishes outside some bounded region R . We can use this transformation to define a renormalized quark mass on the lattice. In this derivation, we follow the pedagogical introduction in [32].

Since the integration measure in the path integral is invariant under the transformation in eq. (4.18), we can derive the equation

$$\langle \delta S \mathcal{O} \rangle = \langle \mathcal{O} \rangle \quad (4.19)$$

and with the explicit form of the variation of the action,

$$\delta S = \int d^4x \omega^a [-\partial_\mu A_\mu^a + 2mP^a], \quad (4.20)$$

which includes the axial current A_μ and the pseudo-scalar density P , we are able to derive the relation

$$\langle \partial_\mu A_\mu^a(x) \mathcal{O} \rangle = 2m \langle P^a(x) \mathcal{O} \rangle. \quad (4.21)$$

This relation holds if the operator \mathcal{O} has no support on the region R . If this was not the case, additional contact terms would appear. As we can see from eq. (4.21), the axial current is conserved for massless quarks. Therefore, this relation is called the partially conserved axial current (PCAC) relation. In this relation, we see the explicit breaking of the chiral symmetry by quark masses.

A corresponding relation can be derived on the lattice [75, 76] which is more difficult, because the Wilson term breaks chiral symmetry even at vanishing bare quark mass. With the properly renormalized but non-improved axial current $(A_R)_\mu^a$ and density P_R^a , we can write eq. (4.21) on the lattice as

$$\langle \tilde{\partial}_\mu (A_R)_\mu^a(x) \mathcal{O} \rangle = 2m_R \langle P_R^a(x) \mathcal{O} \rangle + O(a). \quad (4.22)$$

Here, $\tilde{\partial}_\mu$ is one of the lattice derivatives defined in Appendix A. On the lattice, chiral symmetry is broken by the quark mass and by lattice effects. The size of these effects is a measure for the violation of chiral symmetry. If we improve the currents $(A_R)_\mu^a$ and P_R^a we expect the discretization effects to be of order $O(a^2)$ and in fact relation (4.22) is used to determine c_{SW} and c_A such that linear effects in a are absent [31, 49, 62].

The so-called chiral Ward identities which have been used to define a renormalized quark mass via eq. (4.22) have many applications in lattice QCD. Especially in the field of non-perturbative renormalization and improvement, axial and vector variations can be used to derive a number of important identities. One of these applications is the determination of the critical hopping parameter κ_{cr} which has to be tuned to restore the chiral symmetry of the Wilson Dirac operator. If the mass in eq. (4.22) vanishes κ_{cr} is tuned correctly.

Going into all details would be beyond the scope of this work. Nevertheless, we stress an important point. Since the PCAC relation is derived locally, we expect boundary effects to be absent in the bulk of the lattice. The PCAC relation is an operator relation and valid on every time slice separately.

4.3.3 Current quark masses

Since the renormalized quark mass is present in eq. (4.22), we can use this equation to determine a renormalized and improved quark

mass from the axial current and the pseudoscalar density. If we switch to the basis τ^\pm, τ_3 of the current algebra, we can rewrite the PCAC relation to

$$\langle \tilde{\partial}_\mu (A_R)_\mu^{ij}(x) \mathcal{O}^{ij} \rangle = (m_{R,i} + m_{R,j}) \langle P_R^{ij}(x) \mathcal{O}^{ij} \rangle + O(\alpha^2), \quad (4.23)$$

where we have used the renormalized operators according to eq. (4.7). From this relation, we can define the quark mass to be

$$m_{R,ij} \equiv \frac{m_{R,i} + m_{R,j}}{2} \quad (4.24)$$

$$= \frac{Z_A}{Z_P} m_{ij} [1 + (b_A - b_P) \alpha m_{q,ij} + (b_A - b_P) \alpha \text{Tr}[M_q]] \quad (4.25)$$

with the bare improved current quark mass

$$m_{ij} = \frac{\langle (\tilde{\partial}_0 A_0^{ij} + \alpha c_A \partial_0^* \partial_0 P^{ij}) \mathcal{O}^{ij} \rangle}{2 \langle P^{ij} \mathcal{O}^{ij} \rangle}. \quad (4.26)$$

In principle, the operator \mathcal{O} can be chosen freely, as long as it is defined in a different region as A and P . In the Schrödinger functional, we can use the boundary sources and define a quark mass from the correlation functions f_A and f_P .

4.3.4 Relations between renormalized quark masses

We now have two different definitions of improved and renormalized quark masses. One is the definition from the subtracted quark masses, eq. (4.17) and the other is the definition from the current quark masses, eq. (4.25). For full $O(\alpha)$ improvement, a number of improvement coefficients needs to be known. Both definitions have to coincide in the continuum limit but on the lattice they differ in their cut-off effects. Therefore, one or the other definition might be beneficial, depending on the context.

If we work with (larger than) physical quark masses, the definition from the subtracted quark mass might be inferior for two reasons. The critical hopping parameter might be known to insufficient precision and the leading order term proportional to the average quark mass has to be taken into account. On the other hand, if these parameters are known, the hopping parameter for a fixed renormalized quark mass can be determined without further measurements.

Since we have two expressions for the renormalized quark mass, we can equate both and derive

$$\begin{aligned}
m_{ij} &= \frac{Z_m Z_P}{Z_A} \left\{ \left[m_{q,ij} + (r_m - 1) \frac{\text{Tr}[M_q]}{N_f} \right] + a B_{ij} \right\}, \quad (4.27) \\
B_{ij} &= b_m \frac{m_{q,i}^2 + m_{q,j}^2}{2} + (b_A - b_P) m_{q,ij}^2 \\
&\quad + \left(\bar{b}_m - (b_A - b_P) \frac{(r_m - 1)}{N_f} - (\bar{b}_A - \bar{b}_P) \right) m_{q,ij} \text{Tr}[M_q] \\
&\quad + (r_m d_m - b_m) \frac{\text{Tr}[M_q^2]}{N_f} \\
&\quad + (r_m d_m - \bar{b}_m - (r_m - 1)(\bar{b}_A - \bar{b}_P)) \frac{\text{Tr}[M_q]^2}{N_f}.
\end{aligned}$$

When we introduce the renormalization constant

$$Z \equiv \frac{Z_m Z_P}{Z_A} \quad (4.28)$$

and solve eq. (4.27) for the bare subtracted quark mass, we arrive at

$$m_{q,ij} = \frac{m_{ij}}{Z} - (r_m - 1) \frac{\text{Tr}[M_q]}{N_f} - a B_{ij}. \quad (4.29)$$

This expression can be used to eliminate the bare subtracted quark mass in eq. (4.25),

$$\begin{aligned}
m_{R,ij} &= \frac{Z_A}{Z_P} m_{ij} \left[1 + \frac{(b_A - b_P)}{Z} a m_{ij} \right. \\
&\quad \left. + \left((\bar{b}_A - \bar{b}_P) - (b_A - b_P) \frac{(r_m - 1)}{N_f} \right) a \text{Tr}[M_q] \right], \quad (4.30)
\end{aligned}$$

where we, as always, ignored all higher order terms. In exchange for the elimination of $m_{q,ij}$ we have introduced the parameter r_m . In order to completely remove the dependence on κ_{cr} , we use eq. (4.29) to rewrite the quark mass matrix [77] and obtain

$$\begin{aligned}
m_{R,ij} &= \frac{Z_A}{Z_P} m_{ij} \left[1 + (\bar{b}_A - \bar{b}_P) a m_{ij} \right. \\
&\quad \left. + \left(\frac{(\bar{b}_A - \bar{b}_P)}{Z r_m} - (\bar{b}_A - \bar{b}_P) \frac{(r_m - 1)}{r_m N_f} \right) a M_{\text{sum}} \right], \quad (4.31)
\end{aligned}$$

with

$$\bar{b}_A - \bar{b}_P \equiv \frac{(b_A - b_P)}{Z} \quad (4.32)$$

$$\begin{aligned}
M_{\text{sum}} &\equiv m_{12} + m_{23} + \cdots + m_{(N_f-1)N_f} + m_{N_f 1} \\
&= Z r_m \text{Tr}[M_q] + O(a). \quad (4.33)
\end{aligned}$$

In eq. (4.31), all dependence on the critical hopping parameter has been removed.

So far we have not specified the improvement of valence quarks, i.e., quarks that are not part of the action and therefore not included in dynamical fermion loops. In Appendix A of [74] it was shown that the discussion of this chapter can be extended to a theory with N_f sea and N_{val} valence quarks. The quark matrix contributions remain unchanged when partially quenched valence quarks are introduced and the critical hopping parameter is defined by the point where all N_f sea quark masses vanish.

Both have to be regularized à la Wilson.

4.3.5 The ratio-difference method

We can combine the two definitions for renormalized quark masses, eq. (4.17) and eq. (4.25), in order to construct a third definition with different cut-off effects. We refer to the construction in this section as ratio-difference method [78].

We define a ratio of current quark masses r and a difference of subtracted quark masses d via

$$r_{ij} \equiv \frac{m_i}{m_j}, \quad d_{ij} \equiv am_{q,i} - am_{q,j}, \quad (4.34)$$

where we introduced

$$m_i \equiv m_{ii}, \quad (4.35)$$

as a shorthand for a current quark mass defined from two quarks of different flavor i and i' but equal mass. In this way, we stick to non-singlet, non-diagonal entries of the quark matrix. This setup is chosen such that the multiplicative renormalization of the current quark masses cancels, while the additive renormalization of the bare subtracted quark masses via the critical hopping parameter and the leading order dependence on $\text{Tr}[M_q]$ is canceled in the difference.

Taking into account the mass-dependent improvement coefficients, we define the improved ratios and differences by

$$r_{l,ij} = r_{ij} [1 + (b_A - b_P)d_{ij}], \quad (4.36)$$

$$d_{l,ij} = d_{ij} \left[1 + b_m d_{ij} \frac{r_{ij} + 1}{r_{ij} - 1} + a b_m \text{Tr}[M_q] \right]. \quad (4.37)$$

From these definitions we construct the renormalized quark mass via

$$m_{R,i} = Z_m \frac{r_{l,ij} d_{l,ij}}{r_{l,ij} - 1}, \quad (4.38)$$

and we can express the mass renormalization factor Z_m via

$$Z_m = \frac{Z_A Z}{Z_P}. \quad (4.39)$$

$\text{Tr}[M_q]$ can be eliminated in favor of $M_{\text{sum}}/(Zr_m)$ as in eq. (4.31).

ALGORITHMIC TECHNIQUES

The simulation of full QCD, even when the top quark is considered as decoupled from the theory, is currently prohibitively expensive and can not be accomplished just by an increase in computing power. Our ability to simulate QCD with up to four distinct flavors, in boxes which are large enough to have finite size effects under control and with lattice spacings that allow for a safe extrapolation to the continuum limit, is based on significant algorithmic improvements in the field during the last years. Some of these will be highlighted in this chapter. Especially the simulation of quarks with physical masses seemed to be prohibitively expensive at the beginning of this century [79]. This problem has been overcome by the development of new, quark mass independent solvers [80–86]. As all simulations considered in this thesis have been done using the openQCD package [53, 59], we focus on the algorithms implemented in this program. Some of the techniques are summarized in [87]. The details of the implementations can be found in the documentation of openQCD [53].

5.1 SIMULATION ALGORITHMS

The calculation of the path integral of QCD cannot be done analytically due to the complexity of the non-linear equations. Even the discretized version cannot be determined analytically, due to the large number of integrations in the path integral. Therefore, the numerical solution is determined using importance sampling and Markov Chain Monte Carlo (MCMC). The distribution of the representative ensemble of fields is determined from the action of QCD. We assume that the reader is familiar with the basics of this technique, pedagogical introductions can be found in [13, 14, 88].

We note that the simulation via MCMC leads to configurations of gauge fields which are correlated in the simulation time, since each configuration is based on its predecessor. The transition probability from one configuration U to another U' , denoted by $T(U \rightarrow U')$, satisfies the so-called detailed balance property

$$P(U)T(U \rightarrow U') = P(U')T(U' \rightarrow U), \quad (5.1)$$

when a Metropolis accept/reject step [89] is used. In this equation, $P(U)$ is the probability weight of configuration U in the probability distribution.

5.1.1 Hybrid Monte Carlo

The simulation of quark loops is non trivial since the Grassmann-valued fields cannot be represented straightforwardly in the computer. It can be achieved using the Hybrid Monte Carlo (HMC) [90–92] algorithm. An excellent introduction is given in [87] and in the following we briefly recall the basics of the algorithm. We start with the definition of the HMC algorithm for the gauge action and afterwards describe the inclusion of quark loops.

The partition sum of the theory is extended by momenta $\pi(x, \mu) = \pi^\alpha(x, \mu)T^\alpha$ such that

$$Z = \int D[\mathbf{U}] e^{-S(\mathbf{U})} \rightarrow \int D[\pi] D[\mathbf{U}] e^{-\frac{1}{2}(\pi, \pi) - S(\mathbf{U})}, \quad (5.2)$$

with

$$(\pi, \pi) \equiv \sum_{x, \mu} \pi_\mu^\alpha(x), \pi_\mu^\alpha(x) \quad (5.3)$$

Here, $S(\mathbf{U})$ is the gauge action defined from the fields \mathbf{U} . This change in the definition of the partition sum does not alter the expectation values of observables since it amounts to a multiplication of the original sum with a constant factor. The exponent in the integral is the Hamiltonian

$$H(\pi, \mathbf{U}) = \frac{1}{2}(\pi, \pi) + S(\mathbf{U}). \quad (5.4)$$

The system now obeys Hamilton's equations of motion

$$\partial_t \pi_\mu(x) = -T^\alpha \partial_{x, \mu}^\alpha S(\mathbf{u}) \equiv -F_\mu(x), \quad (5.5)$$

$$\partial_t \mathbf{U}_\mu(x) = \pi_\mu(x) \mathbf{U}_\mu(x) \quad (5.6)$$

in terms of a fictitious Monte Carlo time t , not to be confused with the Euclidean time. F is the molecular-dynamics force. From the classical equations of motion and Liouville's theorem we know that the Hamiltonian and the phase space volume are conserved. If the equations of motions were solved exactly, the HMC algorithm would be exact in the sense that

$$\partial_t H = 0 \quad \Leftrightarrow \quad \Delta H = H(\pi, \mathbf{U}) - H(\pi', \mathbf{U}') = 0 \quad (5.7)$$

for some configuration (π', \mathbf{U}') based on (π, \mathbf{U}) . In practice, we cannot integrate exactly but have to resort to numerical methods. The algorithm becomes exact again when the evolution is combined with a Metropolis step with the acceptance probability $\exp(-\Delta H)$. The HMC algorithm then comprises three steps which are repeated for a number of trajectories:

1. Generate random momenta with a probability density proportional to $\exp\{-\frac{1}{2}(\pi, \pi)\}$.

A definition of the derivative $\partial_{x, \mu}^\alpha$ can, e.g., be found in [93].

2. Integrate the equations of motion from time $t = 0$ to some time $t = \tau$, called trajectory length, and determine U' from U and π .
3. Accept the new configuration with $P_{\text{acc}} = \min[1, \exp(-\Delta H)]$.

The trajectory length is given in molecular dynamics units (MDU).

5.1.2 Integration schemes

The numerical integration of the molecular-dynamics equations (5.6) can be done according to different integration schemes. They differ in their associated integration error which is compensated by the accept/reject step.

We divide the integration interval $[0, \tau]$ in N_0 steps of size ϵ_0 and determine the elementary update steps via Taylor expansion to be

$$\mathcal{J}_\pi(\epsilon_0) : \pi \rightarrow \pi - \epsilon_0 F \quad (5.8)$$

$$\mathcal{J}_U(\epsilon_0) : U \rightarrow \exp(\epsilon_0 \pi) U. \quad (5.9)$$

In the so-called leapfrog integrator, the update steps are connected via

$$\text{LPFR}(\epsilon_0) = \mathcal{J}_\pi\left(\frac{1}{2}\epsilon_0\right) \mathcal{J}_U(\epsilon_0) \mathcal{J}_\pi\left(\frac{1}{2}\epsilon_0\right) \quad (5.10)$$

and N_0 such steps can be combined to integrate from 0 to τ . Omelyan, Mryglod, and Folk [95] introduced $O(100)$ different integration schemes, of which two are used in our simulations. For the second order integrator, a redundant parameter λ is introduced. The update scheme is

$$\text{OMF2}(\epsilon_0) = \mathcal{J}_\pi(\epsilon_0 \lambda) \mathcal{J}_U\left(\frac{1}{2}\epsilon_0\right) \mathcal{J}_\pi(\epsilon_0(1 - 2\lambda)) \mathcal{J}_U\left(\frac{1}{2}\epsilon_0\right) \mathcal{J}_\pi(\epsilon_0 \lambda). \quad (5.11)$$

Since the force has to be evaluated twice as often as with the leapfrog method when this update step is used we could think that this integrator performs worse. It turns out that the norm of the error coefficients is ten times smaller which allows using longer step sizes [96] when OMF2 instead of LPFR is used. This leads to a performance gain by about 50% [97]. Values of λ between 1/6 and 0.2 have been shown to give a good performance [96, 97]. The leapfrog integrator and the OMF2 integrator yield violations of the Hamiltonian conservation of $O(\epsilon_0^2)$.

In our simulations we also use the fourth order integrator OMF4 which has 5 tunable update steps and reduces the integration error to $O(\epsilon_0^4)$ thus allowing even smaller step sizes. In openQCD, these are given by eqs. (63) and (71) of [95]. The violation of the Hamiltonian conservation as a function of the step-size for the integrators LPFR, OMF2 and OMF4 is investigated in figure 4 of [98].

A trajectory length of $\tau = 2$ MDU has been shown to be a good choice considering the stability and cost of QCD simulations [94].

We use $\lambda = 1/6$ in our simulations.

5.1.3 HMC with dynamical quarks

Until now, we considered HMC only for an unspecified action. Now, we look at the application to dynamical fermions. For QCD with a doublet of mass-degenerate sea quarks, we can express the action as

$$S(\mathbf{U}) = S_g(\mathbf{U}) - \ln \left(\det(D^\dagger D) \right). \quad (5.12)$$

As stated before, the numeric representation of the Grassmann-valued quark fields is non-trivial. We can circumvent this issue by using pseudo-fermion fields $\phi(x)$ [99]. These fields carry a Dirac and a color index, but their components are complex numbers rather than Grassmann numbers. We can use the pseudo-fermion fields to express the quark determinant. Using the transformation $\phi = D\eta$, we can write

$$\int D[\eta] D[\eta^\dagger] e^{-(\eta, \eta)} = \int D[\phi] D[\phi^\dagger] \frac{1}{\det(D^\dagger D)} e^{-(D^{-1}\phi, D^{-1}\phi)}. \quad (5.13)$$

It can be seen that the determinant of $D^\dagger D$ is the Jacobian of this transformation. This property can be used to express this determinant via

$$\det(D^\dagger D) = \frac{1}{Z_\phi} e^{-(\phi, (D^\dagger D)^{-1}\phi)} \equiv \frac{1}{Z_\phi} e^{-S_{\text{pf}}(\mathbf{U}, \phi)}, \quad (5.14)$$

where Z_ϕ is the partition sum including the pseudo-fermion fields. We can now write the effective action as

$$S_{\text{eff}}(\mathbf{U}) = S_g(\mathbf{U}) + S_{\text{pf}}(\mathbf{U}, \phi) \quad (5.15)$$

and revisit the steps in the HMC algorithm:

1. Generate random momenta with probability density proportional to $\exp\{-\frac{1}{2}(\pi, \pi)\}$ and generate pseudo-fermion fields according to the probability density $\exp\{-S_{\text{pf}}\}$. This can be done via a normally distributed field η and $\phi = D\eta$.
2. Integrate the equations of motion from time $t = 0$ to $t = \tau$ and determine \mathbf{U}' from \mathbf{U} and π based on S_{eff} . ϕ remains unchanged.
3. Accept the new configuration with $P_{\text{acc}} = \min[1, \exp(-\Delta H_{\text{eff}})]$.

The force F in the molecular-dynamics integration can be split up into a gauge force F_0 and a quark force F_1 . It requires two inversions of the Dirac operator to determine F_1 which takes most of the computing time in the HMC algorithm [87].

The approximation of mass-degenerate up and down quarks is valid, as long as QED effects which have a similar impact on physical observables as the isospin breaking, are neglected [35].

$$H_{\text{eff}} = \frac{1}{2}(\pi, \pi) + S_{\text{eff}}$$

5.1.4 Rational Hybrid Monte Carlo

The inclusion of further quarks to the sea cannot be done straightforwardly with the HMC algorithm. Adding another mass-degenerate pair of heavier quarks can be done (see, e.g., [100]), but this is not even close to an approximation of the strange and charm quark in QCD. Therefore, it is highly desirable to simulate single quark flavors. The simulation of single quark flavors via HMC does not work, since the positivity of D cannot be guaranteed. The determinant of the strange-quark Dirac operator

$$\frac{m_c}{m_s} \approx 13.7$$

$$\det(D_s) = \pm |\det(D_s)| \quad (5.16)$$

is only guaranteed to be positive, if chiral symmetry is exactly preserved on the lattice [87]. The Dirac operator can then be replaced by the non-negative and hermitian operator

$$|Q_s| = (Q_s^2)^{1/2}, \quad Q_s = \gamma_5 D_s \quad (5.17)$$

and its determinant can be expressed by pseudo-fermion fields. Chiral symmetry is not preserved when Wilson fermions are used, as it is the case in our simulations. Nevertheless, since the strange quark is comparably heavy one can assume that the spectral gap of the Dirac operator is large enough such that the configurations with a negative determinant have a negligible weight in the partition sum [101]. In a recent publication [102] it was pointed out that this assumption may not hold under certain circumstances and a correction via reweighting has to be performed in this case.

The calculation of the integral for the strange determinant can be performed with the Polynomial Hybrid Monte Carlo (PHMC) [103–105] or the Rational Hybrid Monte Carlo (RHMC) [106, 107] algorithm. In our simulations, we use the latter.

RHMC relies on the rational function

$$R_{n,\epsilon}(y) = A \frac{(y + a_1)(y + a_3) \dots (y + a_{2n-1})}{(y - a_2)(y - a_4) \dots (y - a_{2n})} \quad (5.18)$$

with degree $[n, n]$ which approximates $1/\sqrt{y}$ in the interval $[\epsilon, 1]$ with the smallest possible relative deviation

$$\delta = \max_{[\epsilon, 1]} |1 - \sqrt{y} R_{n,\epsilon}(y)|. \quad (5.19)$$

The ratio in eq. (5.18) is called the optimal rational approximation or Zolotarev approximation. The coefficients $A > 0$ and $a_1 > a_2 > \dots > a_{2n} > 0$ that minimize δ , as well as δ itself can be determined analytically [108]. R can then be used to represent the strange quark determinant via

$$\det(D_s) = W_s R^{-1} \quad (5.20)$$

In the application, it must be ensured by a proper scaling that the approximation is valid in the whole spectral range.

where $R = R(D_s^\dagger D)$ is appropriately scaled to approximate the inverse in the complete spectral range of $D_s^\dagger D$, denoted by $[r_a, r_b]$. The factor

$$W_s = \det(D_s R) \quad (5.21)$$

is the so-called reweighting factor, needed to correct for the approximation error δ .

The determinant can now, similar to eq. (5.14), be represented via two pseudo-fermion fields ϕ_1, ϕ_2 via

$$\det(Q_s) \propto \int D[\phi_1] D[\phi_2] e^{-S_{\text{pf}}(\mathbb{U}, \phi_1, \phi_2)} \quad (5.22)$$

$$S_{\text{pf}} = (\phi_1, (|Q_s| R)^{-1} \phi_1) + (\phi_2, R \phi_2). \quad (5.23)$$

Since the first term is constant up to deviation of size δ and therefore nearly independent of the gauge field, it does not have to be included in the molecular-dynamics Hamilton function, as long it is included in the final accept/reject step [87]. Apart from this modification, the molecular-dynamics evolution can proceed as in the HMC algorithm.

5.1.5 Reweighting

When the gauge field evolution is performed with the RHMC algorithm, the simulated action deviates from the action of the theory by the approximation error of the Zolotarev approximation. This error is corrected with the reweighting factor from eq. (5.21). When reweighting is applied, the expectation value of an observable \mathcal{O} is defined as [109]

$$\langle \mathcal{O} \rangle_w = \frac{\langle \mathcal{O} W \rangle}{\langle W \rangle} \quad (5.24)$$

where $\langle \cdot \rangle$ is the usual expectation value and W is the reweighting factor. From now on, we assume that this replacement is done for all expectation values if it is necessary and omit the subscript w .

The reweighting factor W_s from eq. (5.21) can be determined via stochastic estimation. With the random noise fields η_i , the factor W_s on a single configuration is given by

$$W_s = \sum_{j=1}^N \exp \left\{ - \left(\eta_j, (1 + Z)^{-1/2} \eta_j \right) \right\}, \quad (5.25)$$

where

$$Z = D^\dagger D R^2 - 1. \quad (5.26)$$

If the number of poles is chosen appropriately, the fluctuations in the reweighting factor are guaranteed to be small and one noise field is sufficient to determine W_s precisely [108].

We choose the number of poles N such that $12N_e \delta^2 \leq 10^{-4}$, where N_e is the number of even lattice points and δ was introduced in eq. (5.19).

5.2 FREQUENCY SPLITTING

The inclusion of the strange quark adds a third force to the molecular-dynamics equations which has to be determined with iterative methods. The step size ϵ_0 , the strength of the forces and the acceptance rate are closely connected. If the force is larger in magnitude, it requires smaller step sizes in order to achieve the same acceptance rate.

It turns out that the quark forces are an order of magnitude smaller than the force deriving from the gauge field [59]. This lead to the idea of using different integration step sizes for different forces [110, 111].

If we consider the gauge force F_0 , a quark force F_1 with step sizes

$$\epsilon_0 = \frac{\tau}{N_0 N_1}, \quad \epsilon_1 = \frac{\tau}{N_1} \quad (5.27)$$

and the leapfrog integrator (5.10), we can construct a so-called hierarchical integration scheme:

$$\text{LPFR}_0(\epsilon_0) = J_{\pi,0} \left(\frac{1}{2}\epsilon_0 \right) J_{\mathbf{U}}(\epsilon_0) J_{\pi,0} \left(\frac{1}{2}\epsilon_0 \right) \quad (5.28)$$

$$\text{LPFR}_1(\epsilon_1) = J_{\pi,1} \left(\frac{1}{2}\epsilon_0 \right) [\text{LPFR}_0(\epsilon_0)]^{N_0} J_{\pi,0} \left(\frac{1}{2}\epsilon_1 \right), \quad (5.29)$$

where the notation of the step J_π introduced in eq. (5.8) has been replaced by

$$J_{\pi,k}(\epsilon_k) = \pi \rightarrow \pi - \epsilon_k F_k. \quad (5.30)$$

This scheme can be generalized to any number of forces, using different integration schemes for every level. Hierarchical integration schemes can lead to even larger performance gains, if frequency splitting of the quark determinant is used.

5.2.1 Hasenbusch splitting

The so-called Hasenbusch frequency splitting [112] is used to split up the determinant of the mass-degenerate pair of light quarks via

$$\begin{aligned} |\det(D)|^2 &= \det \left(DD^\dagger + \mu_n^2 \right) \det \left(\frac{DD^\dagger}{DD^\dagger + \mu_1^2} \right) \\ &\quad \times \prod_{i=1}^{n-1} \det \left(\frac{DD^\dagger + \mu_i^2}{DD^\dagger + \mu_{i+1}^2} \right) \end{aligned} \quad (5.31)$$

with tunable parameters $\mu_n < \mu_{n-1} < \dots < \mu_1$. In the HMC algorithm, the factorized determinants are again represented by pseudo-fermions with the action

$$\begin{aligned} S_{\text{pf}}(\mathbf{U}, \phi) &= \left(\phi_n, (DD^\dagger + \mu_n^2)^{-1} \phi_n \right) \\ &\quad + \sum_{i=1}^{n-1} \left(\phi_i, (DD^\dagger + \mu_{i+1}^2)(DD^\dagger + \mu_i^2)^{-1} \phi_i \right) \\ &\quad + \left(\phi_0, \phi_0 + \mu_1 (DD^\dagger)^{-1} \phi_0 \right). \end{aligned} \quad (5.32)$$

It was found for $n = 2$ that the fluctuation of the force is reduced such that the step-size can be increased by a factor of two to achieve a similar acceptance rate [113]. Later it was found that the dominant contribution comes from the first factor in eq. (5.31) [114]. The different forces deriving from the factorization can be integrated on different time scales to further increase the performance.

Tuning the parameters μ_i is non-trivial. If several factors are used, a hierarchy according to the scheme $\mu_n = 1$ and $\frac{\mu_i}{\mu_{i-1}} = 10$ leads to a good performance [59]. From the shadow Hamiltonian approach, it was found that the cost do not scale with ΔH , but with its variance [115, 116]. The shadow Hamiltonian can be used to optimize the parameters μ_i beyond the simple rule given above [117, 118].

5.2.2 Splitting the rational determinant

Frequency splitting can also be introduced for the rational determinant [107]. In this case, the rational function R given by eq. (5.18) is factorized in several terms

$$P_{k,l} = \prod_{i=k}^l \frac{y + a_{2i-1}}{y + a_{2i}} \quad (5.33)$$

leading to

$$\det(R_{n,\epsilon}^{-1}) \propto \det(P_{1,j}^{-1}) \det(P_{j+1,k}^{-1}) \cdots \det(P_{m+1,n}^{-1}). \quad (5.34)$$

As in the other cases, this factorization can be simulated using an according number of pseudo-fermion fields in the molecular-dynamics evolution. The action involved in the integration can be expressed by

$$S_{\text{pf}}^{\text{RHMC}} = S_{\text{pf}}^{(1,j)} + S_{\text{pf}}^{(j+1,k)} + \cdots + S_{\text{pf}}^{(m+1,n)} \quad (5.35)$$

where the single terms are given by

$$S_{\text{pf}}^{(i,j)} = \left(\phi_2^{(i,j)}, P_{i,j} \phi_2^{(i,j)} \right). \quad (5.36)$$

Here, we only look at the part of the action that is taken into account in the integration as described in eq. (5.23). As it is the case for Hasenbusch frequency splitting, this the factorization of the rational determinant allows to separate forces with different scales in the molecular-dynamics evolution. These forces can be integrated on different time scales and treated with different solvers.

5.2.3 Twisted mass reweighting

The Wilson Dirac operator is not protected against eigenvalues smaller than the bare quark mass, since chiral symmetry is broken. Therefore, the field space is divided by surfaces of zero eigenvalues which are

barriers of infinite action. If the molecular-dynamics evolution was done exactly, these barriers could not be crossed [60]. This problem can be overcome by adding a small twisted-mass term μ_0 in the action which has to be compensated for by another reweighting factor W_0 [59, 109]. Effectively, this amounts to applying the transformation

$$m_0 \rightarrow m_0 + i\mu_0\gamma_5 \quad (5.37)$$

to the Dirac operator of the pair of light quarks.

Equation (5.31) is then changed to

$$\begin{aligned} |\det(D)|^2 &= \det\left(DD^\dagger + \mu_n^2\right) \det\left(\frac{DD^\dagger + \mu_0^2}{DD^\dagger + 2\mu_0^2}\right) \\ &\times \prod_{i=0}^{n-1} \det\left(\frac{DD^\dagger + \mu_i^2}{DD^\dagger + \mu_{i+1}^2}\right). \end{aligned} \quad (5.38)$$

A larger value of μ_0 tends to stabilize the algorithm, but it also increases the fluctuation of the reweighting factor. It is desirable to have little fluctuations because otherwise the precision for observables with small correlation to W_0 is reduced [60, 119].

5.3 SOLVER

The Dirac equation on the lattice is a coupled system of non-linear equations which cannot be solved analytically. Numerical methods have to be used to invert the Dirac matrix. During the integration of the molecular-dynamics equations, we have to perform this inversion for every pseudo-fermion field. While this looks like a giant overhead, the use of appropriate solvers can lead to a significant speedup and thereby enable us to simulate light quarks on fine lattices.

We illustrate the challenge in this parameter region by an investigation of the condition number $\kappa(D)$ of the Dirac operator which is defined by

$$\kappa(D) = \|D\| \|D^{-1}\|, \quad (5.39)$$

where $\|\cdot\|$ is an appropriate norm for D . We can express κ via [87]

$$\kappa(D) = \left(\frac{r_b}{r_a}\right)^{1/2} \propto (am)^{-1} \quad (5.40)$$

where we denote the extremal eigenvalues of $D^\dagger D$ by r_a and r_b . The accuracy of the solution is limited by this condition number. Many algorithmic improvements in iterative solvers are achieved a transformation of D such that the condition number of the inverted operator is reduced.

In the HMC algorithm, a performance gain for subsequent solves can be achieved by the chronological inversion method [121]. Here,

We do not expect this problem to be present in simulations in the Schrödinger functional, due to the gap below the lowest eigenvalues.

A possible norm is given in Appendix B of [120]

Since the Wilson Dirac operator has only nearest-neighbor interactions, all solvers in this section are well suited for the inversion of sparse matrices

the solution from the last n solves is extrapolated to the current Monte Carlo time and used as starting point for the new solve. With this setup, the solution can be found using less iterations.

In the following we will describe the solvers that are part of the openQCD package, since these have been used in the course of our simulations. From the technical point of view, the solvers implemented in this package are state-of-the-art.

5.3.1 Conjugate Gradient solver

The Conjugate Gradient (CG) [122] solver is one variant of many so-called Krylov subspace solvers. Derivations of the algorithm can be found in [123, 124] and its use for lattice field theory is described in [14, 88].

The main idea of the algorithm is to solve the equation

$$A \cdot x = b \quad (5.41)$$

by minimizing the function

$$f(x) = \frac{1}{2}x \cdot A \cdot x - b \cdot x \quad (5.42)$$

which is minimal, when its gradient

$$\nabla f = A \cdot x - b \quad (5.43)$$

vanishes. The minimization is obtained by an iterative construction of vectors p_k which minimize $f(x_k + \alpha_k p_k)$ where

$$x_k = x_{k-1} + \alpha_{k-1} p_{k-1}. \quad (5.44)$$

The vectors p_k are orthogonal to $A \cdot p_i$ with $i \neq k$ and build up a vector space, the so-called Krylov subspace

$$\mathbb{K}^k = (p_0, \dots, p_k). \quad (5.45)$$

As a result, x_{k+1} minimizes f over the whole space \mathbb{K}^k [124]. Since the matrix A in eq. (5.41) has to be positive definite and symmetric for this procedure to work, the conjugate gradient solver can be used to solve

$$D^\dagger D \psi = \eta. \quad (5.46)$$

5.3.2 Multi-shift Conjugate Gradient solver

The Multi-Shift Conjugate Gradient (MSCG) solver [125, 126] is an extension to the CG solver. It allows to solve $n + 1$ equations

$$A \cdot x = b \quad (A + s_i) \cdot x = b, \quad i \in [1, n] \quad (5.47)$$

at once [127]. This can be used for the determination of valence quark propagators with different quark masses and (even more important) in the RHMC algorithm. Here, the solver can be used in the determination of the forces from (5.36), based on the factors in (5.33) [108].

5.3.3 Even-odd preconditioning

As mentioned in the introduction of this section, a decrease of the condition number of the Dirac matrix leads to an increase of the performance of inversions. The reduction of $\kappa(D)$ can be achieved by a transformation of the Dirac matrix. This technique is called preconditioning. Nowadays, there exist a number of preconditioning algorithms which use information on the system of equations to transform them in a profitable manner.

Even-odd preconditioning [120] is inspired by the structure of the Wilson Dirac operator. It is based on the fact that only nearest-neighbor interactions are present in the Dirac operator. A lattice point is classified odd or even based on the parity of the sum of its coordinates $x_0 + x_1 + x_2 + x_3$. Afterwards, every quark field ψ can be split up into two parts

$$\psi = \psi_e \psi_o, \quad \text{where} \quad \psi_e(x)|_{x \text{ odd}} = 0 \quad \psi_o(x)|_{x \text{ even}} = 0 \quad (5.48)$$

and if the lattice points are labeled such that the even points come first, the Dirac operator can be written in the block form

$$D = \begin{pmatrix} D_{ee} & D_{eo} \\ D_{oe} & D_{oo} \end{pmatrix}. \quad (5.49)$$

Here, the first column acts on the even points and the second column on the odd ones. We can solve the two equations in the system

$$\begin{pmatrix} D_{ee} & D_{eo} \\ D_{oe} & D_{oo} \end{pmatrix} \begin{pmatrix} \psi_e \\ \psi_o \end{pmatrix} = \begin{pmatrix} \eta_e \\ \eta_o \end{pmatrix} \quad (5.50)$$

to arrive at the equation

$$(D_{ee} - D_{eo} D_{oo}^{-1} D_{oe}) \psi_e = \eta_e - D_{eo} D_{oo}^{-1} \eta_o \quad (5.51)$$

for ψ_e . The solution of this equation can be used to determine ψ_o via

$$\psi_o = D_{oo}^{-1} (\eta_o - D_{oe} \psi_e). \quad (5.52)$$

The operator

$$\hat{D} = D_{ee} - D_{eo} D_{oo}^{-1} D_{oe}, \quad (5.53)$$

which is the Schur complement of D [88], is called the even-odd preconditioned Dirac operator and we can factorize the determinant of the original operator into

$$\det(D) = \det(D_{oo}) \det(\hat{D}). \quad (5.54)$$

As a consequence of even odd-preconditioning, only D_{oo} has to be inverted. The complexity reduces even further since $\det(D_{oo})$ can be factorized into a product of determinants on the odd sites of the lattice [120].

Even-odd preconditioning can be combined with the algorithmic techniques for solvers presented above. The change of the definition of the Dirac operator then has to be translated into the presented formulae.

5.3.4 Block preconditioning

The idea of even-odd preconditioning can be extended to blocks of lattice points. The Schwarz alternating procedure can be used for such preconditioning [128, 129].

Here, the lattice is divided into equal sized blocks Λ_i . For every single block, we can define Λ_i as the set of points within the block and Λ_i^* as the set of points outside the block. We denote with $\partial\Lambda_i$ the boundary points *inside* the block and with $\partial\Lambda_i^*$ the boundary points *outside* the block. The Dirac matrix can then, similar to eq. (5.49), be written in the block form

$$D = \begin{pmatrix} \tilde{D}_{\Lambda_i} & \tilde{D}_{\partial\Lambda_i} \\ \tilde{D}_{\partial\Lambda_i^*} & \tilde{D}_{\Lambda_i^*} \end{pmatrix}. \quad (5.55)$$

Following [129], we can define matrices acting on the whole lattice as

$$D_{\Lambda_i} = \begin{pmatrix} \tilde{D}_{\Lambda_i} & 0 \\ 0 & 0 \end{pmatrix}, \quad D_{\Lambda_i^*} = \begin{pmatrix} 0 & 0 \\ 0 & \tilde{D}_{\Lambda_i^*} \end{pmatrix}, \quad \dots \quad (5.56)$$

in order to write the Dirac operator as

$$D = D_{\Lambda_i} + D_{\Lambda_i^*} + D_{\partial\Lambda_i} + D_{\partial\Lambda_i^*}. \quad (5.57)$$

The algorithm of the Schwarz alternating procedure now finds a global solution by iteratively solving the equation

$$D\psi'(x)|_{x \in \Lambda_i} = \eta(x), \quad \psi'(x)|_{i \notin \Lambda_i} = \psi(x) \quad (5.58)$$

block by block. Here, the field $\psi(x)$ which is defined on the whole lattice is the solution before block Λ_i has been visited and $\psi'(x)$ the solution afterwards. For the start of the procedure, $\psi(x) = 0$ can be

This is the same as imposing Dirichlet boundary conditions on $\partial\Lambda_i$ [128].

As in the Schrödinger functional, D_{Λ_i} is not expected to become ill-conditioned, due to the infrared cut-off on gluon and quark modes by the boundary conditions [128].

chosen. We can now divide the blocks into the subset of even blocks Ω and the subset of all odd blocks Ω^* ,

$$\Lambda_i \in \Omega \quad \forall i \text{ odd} \quad (5.59)$$

$$\Lambda_i \in \Omega^* \quad \forall i \text{ even.} \quad (5.60)$$

The process of solving (5.58) leaves the field outside of Λ_i unchanged. This change affects the following solution on Λ_i and on the neighboring blocks. All other blocks in the same set $\Omega^{(*)}$ are unaffected. Therefore, it is possible to update all blocks in Ω or Ω^* simultaneously. An update on Ω , followed by an update on Ω^* is called a Schwarz cycle.

The rate of convergence of this algorithm is not good enough to compete with other algorithms, but it can be chosen as preconditioning procedure for a Krylov solver. In contrast to even-odd preconditioning, we do not precondition the Dirac matrix but rather compute an approximate solution for the Dirac equation. Therefore, it is sufficient to use a small number of Schwarz cycles n_{cy} and inside these cycles only a small number of solver iterations n_{mr} for eq. (5.58). The speed can be further enhanced by the use of single precision instead of double precision for this “sloppy” solve. Afterwards, a Krylov solver can be used to compute the precise solution. This can be done with significantly less effort than in the non-preconditioned case. In openQCD, the minimal residual algorithm is used for the sloppy and the GCR algorithm for the precise solve. The latter is described in full detail in [129] and guarantees the correct solution, even if the preconditioning is done inexactly.

In addition to n_{cy} and n_{mr} , the block size is another tunable parameter. Larger blocks lead to a better preconditioning but also consume more time in the block solve. If the blocks fit on the local lattices on single CPU cores, a lot of communication can be avoided, since only the boundaries of the blocks have to be communicated. This is especially important for large lattices which are split in many local lattices.

Block preconditioning can be combined with even-odd preconditioning for the block solves given by eq. (5.58). In this case, the Dirac operator has to be adjusted appropriately.

5.3.5 Deflation acceleration

The most recently developed technique used in our simulations is the so-called deflation acceleration [80, 81]. The general idea is to project the source on a subspace which is orthogonal to the eigenvectors of the smallest eigenvalues of the Dirac operator in order to decrease the condition number of the system. Information on these eigenvectors is required to perform the projection.

*$n_{cy} = 4$ and
 $n_{mr} = 5$ are
exemplary values.*

*For the formulation
of an
SAP-preconditioned
HMC algorithm see
[130].*

If one uses the projection in large volume, one encounters a problem which is connected to a property of the Dirac operator of QCD. The Banks-Casher relation states that the density of eigenvalues λ_k of the massless Dirac operator at the origin is proportional to the quark condensate in the chiral limit [131]. In finite space-time, this leads to an increase of the number of low modes which is proportional to the volume V . With the term low modes we describe modes with eigenvalues below some boundary M . The computational effort for an exact projection increases with $O(V^2)$ [87].

When chiral symmetry is broken, e.g., when using Wilson quarks, the eigenvalues are not protected from fluctuations towards zero.

The eigenvalues of the massive operator $D^\dagger D$ are given by $\alpha_k = m^2 + \lambda_k^2$. A decrease in the quark mass thus results in lower eigenvalues, an increased condition number and therefore in a higher computational effort. Inexact, domain-decomposed deflation removes the strong dependence of the computational effort on the quark mass and at the same time avoids the $O(V^2)$ scaling.

The general idea of deflation is outlined in [87] and the exact procedure which is used in our simulations, is described in Appendix A of [81]. Here, we collect the basic ingredients.

The projection of the Dirac equation onto the eigenvalues α_k is defined by

$$P'\psi = \sum_{k=1}^N v_k(v_k, \psi), \quad (5.61)$$

where v_k denote the eigenvectors. The calculation of these eigenvectors becomes very expensive for large volumina. Therefore, the so-called inexact deflation is used. Here, a projection onto an orthonormal set of N fields ϕ_k via

The construction of the fields ϕ_k is outlined below.

$$P\psi = \sum_{k=1}^N \phi_k(\phi_k, \psi) \quad (5.62)$$

is considered. This projector is used to define the left and right projectors

$$P_L = 1 - DP(PDP)^{-1}P \quad (5.63)$$

$$P_R = 1 - P(PDP)^{-1}PD \quad (5.64)$$

which are employed to decouple the Dirac equation into a "little" system

$$\Leftrightarrow D(1 - P_R)\psi = (1 - P_L)\eta \quad (5.65)$$

$$D\psi_{\parallel} = \eta_{\parallel} \quad (5.66)$$

and a deflated system

$$\Leftrightarrow DP_R\psi = P_L\eta \quad (5.67)$$

$$D\psi_{\perp} = \eta_{\perp}. \quad (5.68)$$

The solution of the little system is easy to compute and the condition number of the deflated system is significantly smaller, if the low modes are efficiently projected out.

The process is refined by dividing the lattice into non-overlapping blocks Λ , as it was done in section 5.3.4. With the orthonormal fields $\phi_1^\wedge \dots \phi_{N_s}^\wedge$ on the blocks Λ , the projection

$$P = \sum_{\Lambda} P_{\Lambda} \quad P_{\Lambda} \psi = \sum_{k=1}^{N_s} \phi_k^\wedge (\phi_k^\wedge, \psi) \quad (5.69)$$

is used. Since the dimension of the blocks is proportional to the volume, the computational effort is lowered from $O(N_s V^2)$ to $O(N_s V)$.

The first step when solving the Dirac equation with a deflated solver is the generation of the deflation subspace. This is achieved by iteratively constructing low-precision solutions of $\phi_{k,i} = D^{-1} \phi_{k,i-1}$, where the global fields $\phi_{k,0}$ are chosen to be random fields. Within this step, the small Dirac equation has to be solved a number of times. The block fields ϕ_k^\wedge can be built by projecting the global fields to the blocks followed by an orthonormalization process on the single blocks.

After the generation of the deflation subspace, the deflated system can be solved with the GCR solver. In openQCD, the deflated system is block-preconditioned using the methods of section 5.3.4, before it is solved. Whereas deflation projects away the low modes, the SAP preconditioning reduces the high-mode components [80].

Deflation only leads to a performance gain, if all low modes are sufficiently well approximated by the projection subspace. Since the number of low modes is significantly larger than the number of deflation fields N_s , this is not guaranteed a priori. Lüscher [80] investigated numerically the so-called property of local coherence of the low modes. He found that all low modes, when projected to the blocks, are contained in the relatively small deflation subspaces, up to small deficits that depend on the block size.

The performance of the deflated solver depends on many different parameters and their interplay may differ, based on the problem at hand. Therefore, an extensive parameter tuning may lead to a significant gain in the performance. In chapter 9, we detail some of the experiences that we have made in the parameter tuning.

If the deflated solver is used in the HMC algorithm, the deflation subspace does not have to be generated from scratch for every solve. Instead, it can be used for some time $\Delta\tau$, before it loses its efficiency. An update of the fields ϕ_k^\wedge restores the quality of the subspace [81].

A larger value of N_s improves the projection but increases the cost.

The hopping parameter used in this solve does not have to coincide with the hopping parameters in the action.

5.3.6 Distance preconditioning

If an iterative solver is used to obtain a solution for the Dirac equation, the stopping criterion relies on a global residuum r_{gl} . The iterative procedure is stopped as soon as the condition

$$r_{gl} \equiv \left| \sum_z D_{x,z} S_z^{(n)} - \eta_x^{(y)} \right|_V < r^* \quad \text{with} \quad D \equiv D[U] + m_0, \quad (5.70)$$

is fulfilled for some fixed bound r^* . Here, $\eta_x^{(y)}$ is a source which is non-zero on time-slice y_0 and $S_z^{(n)}$ is the solution after n iterations. D is the Dirac operator for a valence quark with bare mass m_0 . The norm is determined on the volume $V = T \times L^3$.

This definition of the stopping criterion can lead to problems when heavy bare masses m_0 are considered. The norm of the quark propagator obtained in the solution of the Dirac equation decays exponentially such that

$$|S(x_0)| = \exp\left(-\frac{am_v}{2}|x_0 - y_0|\right) \quad (5.71)$$

where m_v is the mass of the pseudoscalar meson from the (valence) quark propagator. When this mass is large, the norm decays rapidly such that deviations $|DS - \eta|$ at large source-sink separations $x_0 - y_0$ are too small to have any influence on eq. (5.70). The residuum is not sensitive to changes of $O(100\%)$ at large source-sink separations [132].

If the value of r_{gl} is lowered, the number of time slices that have an influence on the stopping criterion is increased. However, machine precision sets a bound on the minimal solver residuum and for lattices with large time extent, a large part of the time slices cannot be used for the calculation of physical observables.

This problem can be cured by the use of distance preconditioning [133]. In the implementation of [134], the Dirac equation is altered to a preconditioned one. The preconditioning matrix P is a unit matrix in spin, color and spatial coordinates. In the time coordinates, it is a diagonal matrix and we define it via

$$P = \text{diag}(p_i), \quad p_i = \exp\left(\alpha|y_0 - x_0^{(i)}|\right), \quad x_0^{(i)}/a = i. \quad (5.72)$$

We transform the Dirac equation by the multiplication of P from the left and insert a factor of $P^{-1}P$. This leads to

$$(PDP^{-1})(PS) = P\eta \quad (5.73)$$

$$\Leftrightarrow PS = (PD^{-1}P^{-1})(P\eta) \quad (5.74)$$

which is solved instead of the original Dirac equation. Subsequently, the solution of the original problem is obtained from (PS) by a multiplication with P^{-1} .

The use of an appropriate choice for the parameter α counteracts the exponential decrease of the norm of the propagator. This leads to an increased cost of the inversion. Therefore, we monitor the quality of the solution under a variation of α . We define the local residuum via

$$r_{\text{loc}}(x_0, y_0) = \frac{\left| \sum_z D_{x,z} S_z^{(n)} - \eta_x^{(y)} \right|_{L^3}}{\left| S_x^{(n)} \right|_{L^3}}, \quad (5.75)$$

where the norm is taken on each time slice separately. We use the local residuum to judge the quality of the solution on single time slices. The dependence on the source position y_0 is made explicit in eq. (5.75).

By tuning α to an appropriate value, we obtain a solution which fulfills $r_{\text{gl}} < r^*$ and $r_{\text{loc}}(x_0, y_0) < r^*$ for $x_0 < t^*$. Since we have to discard a number of time slices close to the boundary due to cut-off effects, we choose $t^* = \frac{7}{8}T$ in our measurements on CLS ensembles. The tuning of α such that the condition for r_{loc} is fulfilled can be done for a single source on one configuration for each ensemble. In principle different bounds for r_{gl} and r_{loc} can be chosen. In this case, the bound on r_{loc} determines the range of time slices that can be used for computations of physical observables.

5.4 ERRORS IN MONTE CARLO DATA

The results obtained from our simulation converge to the true values in the limit of infinitely large statistics. For the realistic case of a finite number of gauge configurations and measurements, we have to quantify the statistical error on our ensemble averages. Depending on approximations, theoretical or numerical constraints, we have to assign an additional systematic error to our results. By definition, we cannot know the size of this systematic error, but we will carefully estimate its size.

In this section, we will discuss the determination of the statistical errors of our data. We will take a look at characteristic features of data from Markov chain Monte Carlo and our tools to determine the errors as reliable as possible.

5.4.1 Correlations

Since the gauge configurations used to calculate physical observables are generated in a Markov chain, subsequent configurations are correlated and this correlation is also present for observables determined on the configurations. We have to take these correlations into account, when we determine the statistical errors of ensemble averages.

Assuming that our configurations are thermalized and have been generated in a finite number of Monte Carlo steps to sample the

We omit the generalization of the following formulae to N_r independent MC runs. It is incorporated in the derivations in [135].

distribution according to the weight from the path integral, we define the ensemble average of an observable O_α via

$$\langle O_\alpha \rangle = \bar{O}_\alpha + \delta\bar{O}_\alpha, \quad \bar{O}_\alpha = \frac{1}{N} \sum_{i=1}^N O_{\alpha,i}, \quad (5.76)$$

where N is the number of configurations and $O_{\alpha,i}$ is the observable on configuration i . The statistical error $\delta\bar{O}_\alpha$ has to be determined based on an investigation of the autocorrelation function of the observables.

In the following, we will present the results of [56, 135–138]. In these references a Taylor expansion of functions of primary observables O_α has been used to derive the methods to determine the statistical error and the error of the error of Monte Carlo observables. The autocorrelation function of two observables O_α, O_β is defined by

$$\Gamma_{\alpha\beta}(t) = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K [O_{\alpha,i+t} - \langle O_{\alpha,i} \rangle] [O_{\beta,i+t} - \langle O_{\beta,i} \rangle] \quad (5.77)$$

and it can be determined from the results in the Monte Carlo chain. Upper values for K and t are given by N .

All results for derived observables trivially hold for primary observables, as well.

In most cases, we are not interested in the determination of primary observables O_α but in the calculation of complicated functions of primary observables. The autocorrelation function of these derived observables $F(O_\alpha)$ can be defined from $\Gamma_{\alpha\beta}$ via

$$\Gamma_F(t) = \sum_{\alpha,\beta} F_\alpha \Gamma_{\alpha\beta}(t) F_\beta, \quad F_\alpha = \frac{\partial F}{\partial \langle O_\alpha \rangle}. \quad (5.78)$$

The partial derivatives F_α can be determined numerically or, more accurately, via automatic differentiation as described in [138].

From Γ_F we can define the naive variance σ_F^2 and the normalized autocorrelation function $\rho_F(t)$ via

$$\sigma_F^2 = \Gamma_F(0), \quad \rho_F(t) = \frac{\Gamma_F(t)}{\Gamma_F(0)}. \quad (5.79)$$

If no correlation is present in the data, the statistical error can be determined directly from σ_F^2 and N . As a measure for the strength of the autocorrelation we introduce the integrated autocorrelation time

$$\tau_{\text{int}}(F) = \frac{1}{2} + \sum_{t=1}^{\infty} \rho_F(t), \quad (5.80)$$

which can be employed to determine the Monte Carlo error of the derived observable F via

$$\delta\bar{F} = \sqrt{2\tau_{\text{int}}(F) \frac{\sigma_F^2}{N}}. \quad (5.81)$$

In the presence of autocorrelations, the statistics is therefore effectively reduced by a factor of $\sqrt{2\tau_{\text{int}}(F)}$.

We are limited to finite statistics, and therefore we cannot perform the sum $\rho_F(t)$ to infinity as given in eq. (5.80). Since the absolute size of ρ_F decreases exponentially with t while its error remains roughly constant, we have to truncate the sum at some value W [137]. In [135], the summation window W is chosen such that the systematic error due to the truncation and the statistical error are balanced and the recipe for the automatic windowing procedure is given in eqs. (50–52) of that reference.

We will refer to the above method for the error estimation as Γ method. Other methods as the Jackknife and the Bootstrap procedure together with binning [137, 139, 140] which are shown to perform inferior [135, 138] in the presence of slow modes are not considered in this thesis.

5.4.2 Slow modes

The autocorrelation function $\Gamma_{\alpha\beta}$ can be shown to have the spectral decomposition [56, 58]

$$\Gamma_{\alpha\beta}(t) = \sum_{n=0}^{\infty} \text{Re} \left\{ c_{\alpha n} c_{\beta n} \lambda_n^{|t|} \right\}, \quad |\lambda_n| = e^{-1/\tau_n} \quad (5.82)$$

with the so-called exponential autocorrelation times $\tau_0 \geq \tau_1 \geq \dots$ which are properties of the algorithm. Based on this representation, we can give an upper limit for the integrated autocorrelation time which is given by

$$\tau_{\text{int}}(O_\alpha) \leq \tau_0. \quad (5.83)$$

The couplings $c_{\alpha n}$ of an observable to the modes λ_n are not known a priori. If an observable couples to a slow mode, i.e., a mode with a large value τ_i and the window for the determination of τ_{int} is not of the order of τ_i , the determination of the statistical error is not save. In [56] a strategy to estimate the statistical error in the presence of slow modes has been derived.

Henceforth, we will refer to the exponential autocorrelation time of the slowest mode τ_0 as τ_{exp} . The size of τ_{exp} can be estimated using prior information on the Monte Carlo chain. For this task, we investigate the integrated autocorrelation times of observables which are known to couple strongly to the slow MC modes. Examples are the squared topological charge Q^2 and the gradient flow coupling \bar{g}_{GF}^2 . For observables determined on the large-volume CLS simulations, we use information from detailed analyses on large statistics to estimate τ_{exp} [56, 60].

When the HMC algorithm is used, the coefficients λ_n and $c_{\alpha n}$ are real.

Only parity even observables couple to $\Gamma_{\alpha\beta}$ in the case of the HMC algorithm.

Using this prior knowledge, we can add an exponential tail to the integrated autocorrelation time and redefine it by [56, 138]

$$\tau_{\text{int}}(F) = \frac{1}{2} + \sum_{t=1}^W \rho_F(t) + \tau_{\text{exp}} \rho_F(W+1), \quad (5.84)$$

where W is determined by the point where ρ_F is compatible with zero within one standard deviation. Even when using this formula to estimate the error, the length of a Monte Carlo run has to be significantly larger than the exponential autocorrelation time, to provide a save error estimation.

In [60], the lower bound $N \geq \tau_{\text{exp}}$ is given.

EXTRACTION OF PHYSICAL OBSERVABLES

In the last chapter we have described the techniques that are used in our work to generate a set of gauge configurations according to the probability weight given by the Boltzmann factor in the Wick rotated path integral. We have also described the solvers that are used to numerically solve the Dirac equation. In this chapter we will explain how we define certain physical observables on the lattice and discuss the techniques that are used for the extraction based on the gauge configurations. A special focus is laid on techniques that suppress systematic effects and improve the statistical precision such as the generalized eigenvalue problem or smearing techniques.

6.1 MESON MASSES

In large-volume simulations hadron masses are used to calibrate the discretized theory, i.e., to eliminate the bare parameters of the Lagrangian in favor of physical observables. Since most meson masses can be extracted to a high precision and are relatively easy to compute, they are well suited for this task. The basis for the extraction of energies of physical states of the Hamiltonian is given by the spectral decomposition of two-point correlation functions.

6.1.1 The spectral decomposition of correlation functions

Denoting the expectation value of a two-point Green function as

$$C(t) = \langle 0 | O(t + t_i) O^\dagger(t_i) | 0 \rangle = \langle O(t + t_i) O^\dagger(t_i) \rangle, \quad (6.1)$$

where a state is created at time t_i and annihilated at time $t + t_i$, we can deduce its spectral decomposition by the insertion of a complete set of eigenstates of the corresponding Hamiltonian H with a proper normalization [141],

$$\mathbf{1} = \frac{\sum_n |n\rangle \langle n|}{2E_n L^3}. \quad (6.2)$$

Here E_n are the energies of the eigenstates of H with $E_1 < E_2 < \dots$ and L^3 is the spatial volume. This leads to

$$\begin{aligned} C(t) &= \sum_n \langle 0 | \hat{O} | n \rangle \frac{e^{-E_n t}}{2E_n L^3} \langle n | \hat{O}^\dagger | 0 \rangle = \sum_n \frac{|\langle 0 | \hat{O} | n \rangle|^2}{2E_n L^3} e^{-E_n t} \\ &\equiv \sum_n |C_n|^2 e^{-E_n t}, \end{aligned} \quad (6.3)$$

where \hat{O} is the operator acting in Hilbert space which corresponds to the interpolator O . For large times $t \rightarrow \infty$, all states with $E_n > E_1$ are exponentially suppressed and we can approximate the correlation function by

$$C(t) = |C_1|^2 e^{-E_1 t} + O(e^{-\Delta_n t}), \quad \Delta_n = E_n - E_1. \quad (6.4)$$

We will investigate the different fit models and the associated systematic uncertainties in section 15.2.

For the extraction of the ground state energy, the correlator can be fitted to a single exponential function at large times or to a sum of exponentials, when a range including shorter time distances t is considered.

If we are only interested in the ground state energies and not in the matrix elements $|C_n|^2$, we can also consider the so-called effective mass which approximates the energy E_0 . One possible definition for the effective mass based on the correlation function on two subsequent time slices is determined via

There exist other definitions which differ only by cut-off effects.

$$m_{\text{eff}}(t) = \ln \left(\frac{C(t)}{C(t+a)} \right). \quad (6.5)$$

From the effective mass, an estimate for the energy E_0 can be extracted by a one-parameter fit to a constant at large times t . For fits including relatively small source-sink separations, we have to take the systematic effects due to excited states into account. We can derive these effects by rewriting eq. (6.4), using the Taylor expansion of the natural logarithm and the observation that eq. (6.5) represents a discretized derivative of the logarithm of the correlation function. We arrive at

$$m_{\text{eff}}(t) = -\frac{d}{dt} \ln(C(t)) \quad (6.6)$$

$$= -\frac{d}{dt} \ln \left(|C_1|^2 e^{-E_1 t} \left[1 + \frac{|C_2|^2}{|C_1|^2} e^{-\Delta_2 t} + O(e^{-\Delta_3 t}) \right] \right) \quad (6.7)$$

$$= E_1 + \Delta_2 \frac{|C_2|^2}{|C_1|^2} e^{-\Delta_2 t} + O(e^{-\Delta_3 t}), \quad (6.8)$$

where we just write the first order correction explicitly. We see that the effective mass forms a plateau with exponential corrections at small times.

6.1.2 Correlation functions and open boundary conditions

If periodic boundary conditions in the time direction are used, as it is a common choice for lattice simulations, the spectral decomposition in eq. (6.3) has to be extended by the backward-propagating counterpart $C(T-t)$. Origin of these terms are the so-called wrappers, i.e., exponentially decaying contributions which wrap around the boundaries. If the source-sink separation is of the order of $T/2$, both contributions are of equal size.

T is the temporal extent of the lattice.

The simulations considered in our work feature open or Schrödinger functional boundary conditions in the time direction and therefore no wrappers are present in our data. Nevertheless, the analysis of correlation functions from lattices with these boundary conditions is non-trivial since effects of the boundaries are encountered. Before we extract meson and quark masses on configurations with open boundary conditions, we have to think about the impact of the boundaries on these observables.

The effects of the boundaries do not only manifest themselves as cut-off effects, but also as physical ones: Since the boundaries have the quantum numbers of the vacuum, states with vacuum quantum numbers contribute to correlation functions, if source or sink are close to the boundaries and the contractions are non-vanishing [142, 143]. They have to be considered in spectral quantities as the effective mass in eq. (6.8). Additionally, cut-off effects that scale with $O(a)$ are expected at the boundaries. As explained in section 3.4, these can be cured with the proper counter terms [58].

In contrast to effective masses, quark masses from the PCAC relation do not suffer from exponential corrections. They are local operator identities and valid on every time slice, differing by $O(a)$ from the continuum value [31, 144]. Considering the boundaries introduce additional cut-off effects, we expect significant deviations from a plateau close to the boundaries.

Since translational invariance in the time direction is broken by the boundaries, some thought has to be spent to fix the absolute location of the source of correlation functions on lattices with open boundary conditions. If correlation functions from sources on different time slices are to be combined, e.g., by an average over these sources, special care has to be taken.

Since we use tree-level improvement, we expect effects of order $O(ag_0^2)$.

6.2 THE GENERALIZED EIGENVALUE PROBLEM

If the extraction of energies of higher states in the spectral decomposition (6.3) is desired, a two or three state fit to the correlation function does not lead to satisfying results in most cases. In this case the exponential contributions cannot be safely extracted before the signal gets lost in the noise or energy levels are too close to each other to discriminate between them in the presence of statistical fluctuations.

Variational methods can be used to overcome this problems. We start by introducing the $N \times N$ correlation matrix $C(t)$ which is defined by

$$C_{ij}(t) = \langle O_i(t) O_j^*(0) \rangle, \quad (6.9)$$

and has the spectral decomposition

$$C_{ij}(t) = \sum_{i=1}^{\infty} e^{-E_n t} \psi_{ni} \psi_{nj}^*, \quad i, j = 1, \dots, N, \quad (6.10)$$

$$\psi_{ni} \equiv (\psi_n)_i = \langle 0 | \hat{O} | n \rangle, \quad (6.11)$$

We denote operators in the Hamiltonian formulation, corresponding to fields with a hat. \hat{H} can be defined as the logarithm of the transfer matrix [26].

where the states $|n\rangle$ are eigenstates of the Hamiltonian ,

$$\hat{H}|n\rangle = E_n|n\rangle \quad E_n < E_{n+1}. \quad (6.12)$$

It was shown [145, 146] that the eigenvalues $\lambda_n(t, t_0)$, $\lambda_1 > \lambda_2 > \dots$ of the generalized eigenvalue problem (GEVP)

$$C(t)v_n(t, t_0) = \lambda_n(t, t_0)C(t_0)v_n(t, t_0), \quad n = 1, \dots, N, \quad t > t_0 \quad (6.13)$$

are given by

$$\lambda_n(t, t_0) = e^{-(t-t_0)E_n} [1 + O(e^{-t\Delta_n})], \quad \Delta_n = \min_{m \neq n} |E_n - E_m|. \quad (6.14)$$

The solution of the GEVP therefore allows to extract the energies of higher states from the time dependence of the eigenvalues $\lambda_n(t, t_0)$. The number of states that can be extracted depends on the dimension of the correlation matrix $C(t)$. At most N energy levels can be determined from a $N \times N$ correlation matrix.

The operators which are used to build the correlation matrix, have to have the correct quantum numbers to project on the desired states. At the same time, they have to be linearly independent to ensure the positive definiteness of $C(t_0)$. The interpolators can be constructed explicitly by different lattice definitions of the continuum operator or by an iterative smearing procedure applied to one definition of the lattice interpolator.

It was suggested in [146] to extract the energy levels at fixed $t_0 < t$ by

$$E_n^{\text{eff}} = \ln \left(\frac{\lambda_n(t, t_0)}{\lambda_n(t+1, t_0)} \right), \quad (6.15)$$

which is similar to the definition of the effective mass from correlation functions in eq. (6.5). Since the exponentially decaying contributions of higher states disappear with an increasing source-sink separation, the energies can in principle be extracted at this point. Unfortunately, if different energy levels are close to each other, the decay of excited states is slow and large source-sink separations have to be considered. At the same time, the absolute size of the systematic error due to hidden excited states remains unknown.

An explicit calculation of the exponential corrections to energies and matrix elements constructed from the GEVP [147] revealed that

We will look at different smearing techniques in section 6.5.

these can be further suppressed. If the condition $t_0 \geq t/2$ is satisfied, the eigenvalues are given by

$$\lambda_n(t, t_0) = e^{-(t-t_0)E_n} [1 + O(e^{-t\Delta_{N+1,n}})] , \quad \Delta_{m,n} = E_m - E_n . \quad (6.16)$$

We can see that the exponential corrections now vanish much faster and that the extraction of nearby states does not pose a problem anymore. By fitting the energy levels $E_n, n = 1, \dots, N$ and the energy level E_{N+1} with an enlarged correlation matrix, a determination of the size of the systematic effects due to the excited state contributions is possible. In practice, choosing a sufficiently large value for t_0 might be difficult since large statistical fluctuations can violate the positivity of $C(t_0)$ at late source-sink separations.

This estimation can be stabilized by simultaneously fitting all energies and their corrections as done in [148].

In our work, we choose the normalization of the states such that the relations

$$(v_m(t, t_0), C(t_0)v_n(t, t_0)) = \delta_{mn} \quad (6.17)$$

and

$$(v_n(t, t_0), C(t)v_n(t, t_0)) = \lambda_n(t, t_0) \quad (6.18)$$

hold.

If we define the operators

$$R_n = (v_n(t, t_0), C(t)v_n(t, t_0))^{-1/2} \frac{\lambda_n(t_0 + t/2, t_0)}{\lambda_n(t_0 + t, t_0)} , \quad (6.19)$$

$$\hat{Q}_n^{\text{eff}} = R_n (\hat{O}, v_n(t, t_0)) , \quad (6.20)$$

as it was done in [147], we can calculate matrix elements of the local operators \hat{P} from

$$\begin{aligned} \langle 0 | \hat{Q}_n^{\text{eff}} e^{-\hat{H}t} \hat{P} e^{-\hat{H}t} (\hat{Q}_n^{\text{eff}})^\dagger | 0 \rangle &= \langle Q_n^{\text{eff}}(2t) P (Q_n^{\text{eff}}(0))^* \rangle \quad (6.21) \\ &= \langle n | \hat{P} | n' \rangle + O(e^{-\Delta_{N+1,n}t_0}) , \quad (6.22) \end{aligned}$$

provided $t_0 \geq t/2$ holds. We do not go into more detail concerning the extraction of matrix elements from the GEVP at this point. We will do this with a special focus on HQET, when we define the matrix elements considered in the effective theory in chapter 17.

6.3 FITTING

In many applications we like to determine the parameters of a model, based on our numerical data. The quality of the final results depends on the statistical accuracy and the consideration of all sources of systematic effects. The determination of our best estimators for the model parameters is done via fitting, the minimization of the difference between the model and the data. Pedagogical introductions can be found in [124, 140].

6.3.1 *Fitting correlated data*

Lattice data are not only correlated between subsequent configurations, but also between different lattice sites on single configurations. When correlation functions are fitted to a model function, this correlation has to be taken into account. This is done using the covariance matrix C between the data points.

If we consider a model $\phi(x, A)$ with N_A many parameters A_i and a variable x , e.g., $\phi(x, A) = A_1 + xA_2$ and we measure data \bar{y}_j at N_x many values x_j , then we expect them to be normally distributed according to the distribution [140, 149]

$$P(\bar{y}) = (\det(2\pi C))^{-1/2} \exp\left(-\frac{1}{2}\delta\bar{y}^T C^{-1}\delta\bar{y}\right). \quad (6.23)$$

Here $\delta\bar{y} = \bar{y} - Y$ is the deviation of the measured values from the correct ones and the covariance matrix is defined by $C_{ij} = \langle\delta\bar{y}_i\delta\bar{y}_j\rangle$.

When we perform a fit to measured data, we like to determine the best possible estimators \bar{a} for the parameters A_i , based on the data points \bar{y}_j and their covariance. This is done via a minimization of

$$\chi^2(\alpha) = r^T W r, \quad r_j = \bar{y}_j - \phi_j(\alpha) \quad (6.24)$$

with a weight matrix W . In order to exploit all data and correlations, it seems to be the best way to choose the inverse of the covariance matrix C^{-1} as weight matrix. We will call this procedure *correlated fitting*. Unfortunately, the covariance matrix itself can only be extracted with errors. Fluctuation in the data may lead to a large condition number of the matrix. In addition, if the number of data points is larger than the number of independent configurations, C has eigenvalues which are zero. In the latter case, an inversion is not possible. In the first (more frequent) case, the determination of the smallest eigenvalues which have the largest weight, is imprecise.

There are several studies which indicate that correlated fitting may introduce a bias or give worse results than *uncorrelated fitting* [150–153]. In the case of uncorrelated fitting, it is common to use the inverse of the diagonal of C as weight matrix W . Effectively, each value is weighted with the inverse of its variance and the correlation between the data is neglected.

If we know the underlying model, we expect the best estimators \bar{a} to be found for the minimal χ^2 . If we do not know the model, the situation is more complicated. For uncorrelated data and a model which is linear in the fit parameters, we expect [140]

$$\frac{\chi_{\min}^2}{\text{NDF}} = 1, \quad \text{where } \text{NDF} = N_x - N_A. \quad (6.25)$$

If the ratio is significantly larger than unity, the model is not able to describe the data sufficiently. If it is significantly smaller, we have

The minimization is done iteratively and may be non-trivial for a complex model.

too much freedom in our model, not all parameters are constrained sufficiently by the data.

In the case of correlated data, this rule of thumb is not valid anymore, since the effective number of degrees of freedom is smaller than $N_x - N_A$. Therefore, a model selection may become difficult. This problem is overcome by the introduction of an effective χ^2 [149, 154] which is the expected value of χ_{\min}^2 for data with correlation according to C. The formula for its determination can be derived based on an expansion of the estimated model parameters and the model function

$$\mathbf{a} = \mathbf{A} + \delta\mathbf{a} \quad (6.26)$$

$$\phi(\mathbf{a}) = \phi(\mathbf{A}) + \sum_{i=1}^{N_A} \delta a_i \partial_i \phi + \dots, \quad \partial_i \phi = \frac{\partial \phi(\mathbf{A})}{\partial A_i} \quad (6.27)$$

and its final form reads

$$\chi_{\text{exp}}^2(\bar{\mathbf{a}}) = \text{Tr} \left[(\mathbf{1} - P_\phi) W^{1/2} C W^{1/2} \right]. \quad (6.28)$$

Here, W is the weight matrix which is used in the fit and C is the covariance matrix derived from the data. The projector $(\mathbf{1} - P_\phi)$ is orthogonal to the projector P_ϕ and P_ϕ projects onto the subspace spanned by the $N_x \times N_A$ matrix

$$Q = W^{1/2} V, \quad V_{ij} = \frac{\partial \phi(x_i, \mathbf{A})}{\partial A_j} \quad (6.29)$$

We can determine it via

$$P_\phi = Q(Q^T Q)^{-1} Q^T. \quad (6.30)$$

We now expect

$$\frac{\chi_{\min}^2(\bar{\mathbf{a}})}{\chi_{\text{exp}}^2(\bar{\mathbf{a}})} \simeq 1 \quad (6.31)$$

if the data are described by the model. Using χ_{exp} , we can test a model function in the presence of correlations without the need to perform the potentially problematic inversion of the covariance matrix.

6.4 GRADIENT FLOW OBSERVABLES

Strongly coupled gauge fields show large fluctuations. In lattice QCD, this can lead to undesired effects since ultra violet (UV) modes of the fields can reduce the overlap of lattice interpolators with the ground state or complicate the definition of topological observables. These UV fluctuations can even be unphysical. In this case, they are removed by the ensemble average.

For a long time, smearing techniques have been used to improve the overlap of lattice interpolators with physical wave functions and

More precisely, the effective value for N_x is reduced due to the correlation.

The term $W^{1/2} C W^{1/2}$ reduces to the correlation matrix for uncorrelated fits with $W = (\text{diag}(C))^{-1}$.

$\chi_{\text{exp}}^2(\bar{\mathbf{a}}) = \text{Tr}[\mathbf{1} - P_\phi] = \text{NDF}$ for $W = C^{-1}$

we will come back to these techniques in section 6.5. A rather recent development is the introduction of the Wilson or Gradient flow [155, 156] which solves the problems described above and offers more theoretical control than smearing.

6.4.1 The Wilson flow

The definition of a fermion flow is given in [157]. We will not consider it in this work.

The Yang-Mills gradient flow is defined by the evolution of the gauge fields in a fictitious time $t \geq 0$. It is given by the partial differential equation [157]

$$\frac{dB_\mu(x, t)}{dt} = D_\nu G_{\nu\mu}(x, t) = -\frac{\delta S_{YM}(B)}{\delta B_\mu(x, t)}, \quad B_\mu(x, 0) = A_\mu(x) \quad (6.32)$$

with the covariant derivative

$$D_\mu = \partial_\mu + [B_\mu, \cdot] \quad (6.33)$$

and the field strength

$$G_{\mu\nu} = [D_\mu, D_\nu] = \partial_\mu B_\nu - \partial_\nu B_\mu + [B_\mu, B_\nu]. \quad (6.34)$$

The fields B_μ depend on the four dimensional space-time coordinate x and the flow time t which is not to be confused with the Euclidean time x_0 or the Monte Carlo time. The initial field configurations are given by the gauge fields. The evolution of B_μ according to $D_\nu G_{\mu\nu}$ leads to a smoothing of the gauge fields along the negative gradient of the Yang-Mills action towards the minimum of the action [158]. Effectively, this leads to a smoothing of the fields over a region with radius $\sqrt{8t}$ in all four dimensions [156].

In [157, 159] it was shown to all orders of perturbation theory that gluonic observables built from the gauge fields B_μ are automatically finite after the removal of the cut-off, i.e., they do not need any additional renormalization when the four dimensional theory has been renormalized.

The Wilson flow [156, 160, 161] is a lattice version of the Yang-Mills gradient flow and it is defined by

$$a^2 \frac{d}{dt} V_\mu(x, t) = -g_0^2 [\partial_{x,\mu} S_W(V_\mu)] V_\mu(x, t) \quad V_\mu(x, 0) = U_\mu(x) \quad (6.35)$$

with the Lie-algebra valued derivative of the action, as we have seen it in the HMC algorithm, eq. (5.5). Here, this derivative is taken with respect to the flowed gauge links $V_\mu(x, t)$. If the numerical solution is determined by the Euler method, we can write the evolution of the gauge link for a step size ϵ as

$$V_\mu(x, t + \epsilon) = \exp(-\epsilon g_0^2 \partial_{x,\mu} S_W) V_\mu(x, t) \quad (6.36)$$

and this form allows us to gain insight in the evolution. In practice, we use the Runge-Kutta-Fehlberg method [158, 162–164] to solve the partial differential equation.

6.4.2 The Zeuthen flow

Unfortunately, the Wilson flow has rather large cut-off effects. There exist different approaches to improve the scaling towards the continuum limit, see e.g. [165–167]. As we have seen it for many other discretized quantities, the systematic improvement of the Wilson flow à la Symanzik can be achieved such that no $O(a^2)$ cut-off effects are introduced by the flow itself [168]. The so-called Zeuthen flow is defined by

$$a^2 \frac{d}{dt} V_\mu(x, t) = -g_0^2 \left(1 + \frac{a^2}{12} \nabla_\mu^* \nabla_\mu \right) [\partial_{x,\mu} S_{LW}(V_\mu)] V_\mu(x, t) \quad (6.37)$$

with the same initial condition as the Wilson flow, the lattice forward and backward derivatives and the Lüscher-Weisz action, eq. (2.17). All observables which we will define from the gradient flow, can be based on the Wilson and the Zeuthen flow and the different definitions differ in their cut-off effects.

In the Schrödinger functional, the term proportional to a^2 is set to 0 if $\mu = 0$ and one of the endpoints is at a time boundary at 0 or T [45].

6.4.3 Setting a scale from the flow

It is possible to define a renormalized coupling from the flow and this coupling can be employed to set the scale in lattice simulations [156, 157, 169]. In the continuum, the Yang-Mills action density from flowed fields can be defined from the field strength tensor via

$$E = -\frac{1}{2} \text{Tr} [G_{\mu\nu} G_{\mu\nu}] . \quad (6.38)$$

This definition can be directly translated to a discretized form of the action density by constructing the field strength tensor from clover plaquettes. Another definition is defined analogously to the Wilson action (2.6) by

$$E = 2 \sum_{p \in P_x} \text{ReTr} [1 - V(p, t)] , \quad (6.39)$$

where we take the sum over all unoriented plaquettes. Both definitions differ by their cut-off effects and can be used to define a reference scale t_0 by the implicit equation

$$\{t^2 \langle E \rangle\}_{t=t_0} = 0.3 . \quad (6.40)$$

Of course, other values than 0.3 could be chosen on the right hand side. The alternative scale w_0 which is based on the derivative of $t^2 E$ has been constructed to reduce cut-off effects [165].

The definition for t_0 based on clover plaquettes has been used to set the scale for the CLS ensembles.

In [156], the connection between $t^2\langle E \rangle$ in infinite volume and the renormalized coupling $\bar{g}_{\overline{MS}}(\mu)$ in the \overline{MS} scheme at the scale $\mu = 1/\sqrt{8t}$ was shown to be

$$t^2\langle E(t) \rangle = \frac{3(N^2 - 1)\bar{g}^2}{128\pi^2} \{1 + \bar{c}_1\bar{g}^2 + O(\bar{g}^4)\}, \quad (6.41)$$

for the gauge group $SU(N)$ with the numerical constant \bar{c}_1 . From this relation it is apparent that we can define a renormalized coupling from the flow.

6.4.4 The gradient flow coupling in the Schrödinger functional

We will use the definition of the renormalized gradient flow coupling \bar{g}_{GF} from the flow in our matching procedure in the Schrödinger functional. The coupling for Schrödinger functional boundary conditions has been defined and studied in [158]. It is defined for an arbitrary $SU(N)$ gauge field by

$$\bar{g}_{GF}^2 = N^{-1}t^2\langle E(t) \rangle = \bar{g}_{\overline{MS}}^2 + O(\bar{g}_{\overline{MS}}^4) \quad (6.42)$$

with the normalization constant N . The second equation shows the connection between the gradient flow coupling and a renormalized coupling defined in the \overline{MS} scheme. If we set the scale by the box size L , we can write

$$\mu = \frac{1}{\sqrt{8t}} = \frac{1}{cL}, \quad (6.43)$$

where c is a dimensionless constant that represents the fraction of the smoothing range over the total box size [158]. In order to fully specify the coupling in the Schrödinger functional with boundary fields and broken translational invariance in the time direction, we have to define the value of the boundary fields ρ and the time-slice x_0 , where we evaluate the action density. Explicitly writing out all the dependencies, we arrive at

$$\bar{g}_{GF}^2(L) = N^{-1}(c, \rho, x_0/T)t^2\langle E(t, x_0) \rangle|_{t=c^2L^2/8}, \quad (6.44)$$

as the definition of the coupling. The normalization constant N in the continuum limit was determined in [158] to ensure

$$\bar{g}_{GF}^2 = g_0^2 + O(g_0^4). \quad (6.45)$$

Based on the investigations in [158], a number of choices were made in [45] in order to fix the definition of the coupling. In our work, we stick to these choices and define in the following the gradient flow coupling as we use it in our work.

It was found that the cut-off effects induced by the boundaries in the Schrödinger functional are reduced if only the spatial (or magnetic)

The investigation of the coupling in a finite box with periodic boundary conditions has been done in [170].

Table 6.1: Normalization constant $\mathcal{N}(c, a/L)$ for $c = 0.3$ and $x_0 = T/2$, the tree-level improved Lüscher-Weisz gauge action, zero boundary fields and the gradient flow coupling defined from the Zeuthen flow [171].

L_0/a	$\mathcal{N}(c, a/L)$	L_0/a	$\mathcal{N}(c, a/L)$
4	0.012341170468270	20	0.008575359627103
6	0.010162691462430	24	0.008569387847540
8	0.009031614807931	32	0.008565541650006
10	0.008744966371393	40	0.008564480684962
12	0.008650917856809	48	0.008564098025073
16	0.008591758449508	64	0.008563853943383

components of the action density are used. The statistical precision is not impaired by neglecting the time component. The normalization constant N has been determined in eq. (3.1) of [158]. In the same reference, a normalization constant with explicitly removed cut-off effects has been introduced. The values of this constant $\mathcal{N}(c, a/L)$ for all relevant values of L/a in this work are given in table 6.1. The use of $\mathcal{N}(c, a/L)$ is expected reduce cut-off effects in the coupling. The constant c was fixed to 0.3 in our project on HQET. Following [45], we therefore define

$$\bar{g}_{\text{GF}}^2 \equiv \bar{g}_{0.3}^2(L) \quad (6.46)$$

with

$$\bar{g}_c^2(L) = t^2 \mathcal{N}^{-1}(c, a/L) \frac{\langle E_{\text{mag}}(t, x) \hat{\delta}(Q) \rangle}{\langle \hat{\delta}(Q) \rangle} \Big|_{\sqrt{8t}=cL, x_0=T/2} \quad (6.47)$$

and

$$E_{\text{mag}}(t, x) = \frac{1}{4} [G_{ij}^a G_{ij}^a]^{LW}, \quad i, j = 1, 2, 3. \quad (6.48)$$

If we do not specify it explicitly, the use of the Zeuthen flow is assumed. With the definition of the Lüscher-Weisz gauge action from eq. (2.17), we can express the magnetic component of the action density by

$$E_{\text{mag}} = \sum_x \left[\frac{5}{3} P_{ij}(x, t) - \frac{1}{12} R_{ij}(x, t) \right], \quad (6.49)$$

where we denote the spatial plaquettes with P_{ij} and the spatial 2×1 rectangles with R_{ij} . In the Schrödinger functional with dynamical quarks, the choices made for the boundary fields and the phase angles

In the determination of the improvement coefficients in the CLS coupling region, we chose $c = 0.35$.

θ are of importance for the precise definition of the coupling. The choices made in our work have been specified in chapter 3.

Reweighting is explained in section 5.1.5.

In the definition of the coupling in eq. (6.47) we included a reweighting with the reweighting factor $\hat{\delta}(Q)$ which we did not yet explain. We define the reweighting factor by

$$\hat{\delta}(Q) = \begin{cases} 1, & |Q| < 0.5 \\ 0, & \text{otherwise} \end{cases}, \quad (6.50)$$

where Q denotes the topological charge. This procedure was introduced and motivated in [172]. In the next section, we introduce a discretized definition of the topological charge. Afterwards we motivate the reweighting of eq. (6.50) which effectively projects onto the sector of vanishing topological charge.

6.4.5 The topological charge from the gradient flow

The topological charge is used to distinguish different topological sectors which emerge in the theory. In the continuum, these different topological sectors are disconnected in field space and it is not possible to tunnel between them. Two fields belonging to different homotopy classes cannot be continuously deformed into another. There exist infinitely many such classes which are characterized by their integer topological charge Q . This charge can be defined from the field strength tensor via [173]

$$Q = -\frac{1}{32\pi^2} \int d^4x \epsilon_{\mu\nu\rho\sigma} \text{Tr} [F^{\mu\nu} F^{\rho\sigma}]. \quad (6.51)$$

This statement is only true, if the Dirac operator exhibits exact zero modes.

From the Atiyah-Singer index theorem [174] it is known that the topological charge can also be determined from the left handed n_- and right handed n_+ zero modes of the Euclidean Dirac equation via [11]

$$Q = \frac{1}{N_f} (n_- - n_+). \quad (6.52)$$

The topological charge therefore connects gauge and fermionic properties of the theory.

When we simulate the field theory with Monte Carlo methods, we have to ensure that all topological sectors are sampled according to their weight in the path integral in the continuum limit. In order to monitor the sampling, we have to have a definition of the topological charge on the lattice. Since Wilson's formulation of quarks on the lattice breaks chiral symmetry explicitly, we cannot express Q in terms of the zero modes of the Dirac operators but have to determine the charge from the field strength tensor.

It was shown [156, 175–177] that Q can be defined from the lattice version of the field strength tensor such that it reduces to the topological charge in the continuum limit, if the smoothness condition

$$s_p = \text{ReTr} [1 - U(p)], \quad h = \max_p s_p, \quad h < 0.067 \quad (6.53)$$

is satisfied by the plaquettes p on the lattice. Towards the continuum limit, the gauge configurations with $h \geq 0.067$ have a quickly decreasing weight and therefore, the disconnected sectors emerge naturally.

At the same time, all possible values of h are present in the functional integral.

The gradient flow can be used to smooth the fields at larger values of a and therefore allow the gauge configurations to pass the condition in eq. (6.53). Since the flow transformation is invertible, all physical properties of the theory are conserved and the definition of the topological charge from the gauge fields becomes unambiguous [156].

In our work, we define the topological charge density by [178]

$$q(x, t) = -\frac{1}{32\pi^2} \epsilon_{\mu\nu\rho\sigma} \text{Tr} [G_{\mu\nu}(x, t) G_{\rho\sigma}(x, t)] \quad (6.54)$$

with the field strength tensors from the flowed clover plaquettes. The topological charge is then defined by

$$Q = a^4 \sum_x q(x, c), \quad (6.55)$$

where c is chosen as for the gradient flow coupling and the Zeuthen flow is used. This definition does not lead to integer values, but values which are distributed around integers. Towards the continuum limit, the width of these distributions shrinks and in the continuum limit the integer valued topological charge is restored.

We will inspect this behavior more closely when we look at measurements of Q on our ensembles.

In section 3.4 we motivated the use of open boundary conditions for large volume simulations with the ability to properly sample all topological sectors when going to the continuum limit. This is not possible, when using Schrödinger functional boundary conditions and we have to live with ensembles where Q is stuck in single sectors for large fractions of the affordable Monte Carlo time. Fortunately, the probability of topological sectors with non-zero charge is suppressed in small volumes.

There has been a suggestion to use open-SF boundary conditions for certain applications [179].

In [172], the projection onto the sector of vanishing topological charge according to eq. (6.50) was introduced. Since Q is non-integer on the lattice, this sector is defined by nearest-integer rounding. If observables explicitly depend on the topological sector and it is not possible to sample all sectors on some ensembles, a bias is introduced in the continuum extrapolation. For some observables, this bias can be removed when the quantity is projected onto the sector of vanishing topological charge on all ensembles. For large lattice spacings, this amounts effectively to a reduction of the statistics up to 70%. At the

same time, critical slowing down towards the continuum limit might be avoided.

In the case of the gradient flow coupling, the projection onto the sector of vanishing topological charge is mandatory since it is part of our definition. For other observables, we have to be more careful. If we want to determine physical observables on large volumes, we have to include all topological sectors to sample the path integral. This is different when we look at observables defined from Ward identities, e.g., current quark masses. The operator identities hold within each sector and therefore our results are still valid if we only consider one topological sector. A projection to the sector of vanishing topological charge just changes the cut-off effects [74]. We will indicate the use of this projection in the Schrödinger functional, when applied to other observables than the coupling.

6.5 SMEARING

As indicated in the last section, smearing techniques are widely used to increase the overlap of the discretized interpolators with physical wave functions and to eliminate unphysically high modes of the gauge fields. In general, we can distinguish between gauge smearing, where the gauge fields are altered, and quark smearing which is applied to the quark fields. In this section, we want to introduce the smearing techniques used in our work.

6.5.1 APE smearing

APE smearing [180] is iteratively applied to the spatial gauge links. One step of spatial APE smearing replaces the original gauge link according to

$$U_i(\mathbf{n}) \rightarrow P_{\text{SU}(3)} \left[(1 - \alpha)U_i(\mathbf{n}) + \frac{\alpha}{4} \sum_{j=\pm 1, j \neq i}^{\pm 3} U_j(\mathbf{n})U_i(\mathbf{n} + \hat{\mathbf{a}}_j)U_j(\mathbf{n} + \hat{\mathbf{i}})^{-1} \right]. \quad (6.56)$$

In this replacement, the contributions of three-link paths of neighboring gauge links are added to each spatial link. Thereby, its spatial extent is enlarged. High modes are suppressed if we choose the smearing parameter α in the range $0 \leq \alpha \leq 0.75$ [181]. If APE smearing is performed including the links in the time direction, the normalization of the sum changes to $\alpha/6$.

After the replacement by the sum of gauge links, the new gauge links are not elements of the gauge group $\text{SU}(3)$ anymore. Therefore, we project the components of the gauge links onto the group. In

equation (6.56), this projection is indicated by $P_{\text{SU}(3)}$. We project the complex 3×3 matrices W according to the procedure given in reference [182]. We carry out the replacement

$$W \rightarrow \frac{W}{\sqrt{\frac{1}{3}\text{Tr}[WW^\dagger]}} \quad (6.57)$$

and proceed with four iterations of

$$W \rightarrow X \left(1 - \frac{i}{3}\text{Im}[\det(X)] \right), \quad (6.58)$$

$$\text{where } X = W \left(\frac{3}{2} - \frac{1}{2}W^\dagger W \right).$$

The iterative smearing in eq. (6.56) is applied a number of times and the gauge links get smeared out more and more during this procedure. Operators based on different number of smearing steps can be combined in the variational method, cf. section 6.2.

6.5.2 HYP smearing

The spatial extent of the smearing procedure is limited, when hypercubic blocking (HYP) smearing [183, 184] is applied. In this method, the extent of the smeared links is limited to the hypercubes attached to the original link. The smearing is performed via three smearing steps which we will specify in the following.

The final HYP smeared links $V_\mu(\mathbf{n})$ which are used later on in the extraction of physical quantities, are defined by

$$V_\mu(\mathbf{n}) = P_{\text{SU}(3)} \left[(1 - \alpha_1)U_\mu(\mathbf{n}) + \frac{\alpha_1}{6} \sum_{\pm \nu \neq \mu} \tilde{V}_{\nu;\mu}(\mathbf{n}) \tilde{V}_{\mu;\nu}(\mathbf{n} + \mathbf{a}\hat{\nu}) \tilde{V}_{\nu;\mu}(\mathbf{n} + \hat{\mu})^{-1} \right], \quad (6.59)$$

where we introduced so-called decorated links \tilde{V} . In the notation of [183], $\tilde{V}_{\mu;\nu}(\mathbf{n})$ is a decorated link in direction μ at the lattice point \mathbf{n} which is constructed without staples extending in direction ν . These links are defined by

$$\tilde{V}_{\mu;\nu}(\mathbf{n}) = P_{\text{SU}(3)} \left[(1 - \alpha_2)U_\mu(\mathbf{n}) + \frac{\alpha_2}{4} \sum_{\pm \rho \neq \nu, \mu} \tilde{V}_{\rho;\nu\mu}(\mathbf{n}) \tilde{V}_{\mu;\rho\nu}(\mathbf{n} + \mathbf{a}\hat{\rho}) \tilde{V}_{\rho;\nu\mu}(\mathbf{n} + \hat{\mu})^{-1} \right], \quad (6.60)$$

where the notation $\bar{V}_{\rho,\nu\mu}(n)$ is similar to the one above; the link is extended by staples which do not extend in directions ν or μ . Finally, the links \bar{V} are defined by a modified APE smearing step via

$$\bar{V}_{\mu,\nu\rho}(n) = P_{SU(3)} \left[(1 - \alpha_3) U_\mu(n) + \frac{\alpha_3}{2} \sum_{\pm\eta \neq \rho, \nu, \mu} U_\eta(n) U_\mu(n + a\hat{\eta}) U_\eta(n + \hat{\mu})^{-1} \right]. \quad (6.61)$$

In this step, only two staples in the direction orthogonal to ρ , ν and μ are added.

The procedure defined above achieves a smeared, fat link that consists of unsmeared, thin links inside the hypercube attached to the original link. The parameter set $\alpha \equiv (\alpha_1, \alpha_2, \alpha_3)$ has to be optimized to achieve a good performance for the problem at hand. In our work, we consider two sets of these parameters. We define

$$\text{HYP1:} \quad \alpha = (0.75, 0.6, 0.3), \quad (6.62)$$

which is the set of parameters introduced in [183] and

$$\text{HYP2:} \quad \alpha = (1.0, 1.0, 0.5), \quad (6.63)$$

motivated in [182].

We will use both versions of HYP smearing in the context of HQET in chapter 16. It was found in [185] that an exponentially improved statistical precision can be achieved for observables including static quarks if HYP smearing is used. Since we do not know a priori which version performs better, we will always consider HYP1 and HYP2. In a comparison of the final results, we will investigate which the set of parameters leads to smaller cut-off effects or an enhanced statistical precision.

The smearing itself and the calculation of static quark lines is relatively cheap.

6.5.3 Gaussian smearing

Up to now we have considered modifications of the gauge links. We will now look at smearing techniques which are used to enlarge the spatial extent of quark fields. The construction of extended interpolators which are contracted to form quark propagators, aims at increasing the overlap with physical wave functions.

One possibility to construct such extended sources is Gaussian smearing [186]. Here, the new field is constructed iteratively by the replacement

$$\psi(x) \rightarrow \psi'(x) = (1 + \kappa_G a^2 \Delta) \psi(x), \quad (6.64)$$

The discretized Laplace operator Δ is defined in app. A.

where we use the three dimensional lattice Laplace operator $\Delta \equiv -\nabla_i \nabla_i^*$. The parameter κ_G denotes the coupling strength of the nearest neighbor interactions. The extent of the new field depends on the number of iterations n_k . The relation between n_k and the physical extent is not known a priori, but approximately it is given by

$$r_k \approx 2a\sqrt{\kappa_G n_k}. \quad (6.65)$$

We can also define the square of the smearing radius via

$$\langle r^2 \rangle \equiv \frac{\sum_{\mathbf{x}} r^2 \text{Tr} [\psi^\dagger(\mathbf{x})\psi(\mathbf{x})]}{\sum_{\mathbf{x}} \text{Tr} [\psi^\dagger(\mathbf{x})\psi(\mathbf{x})]}, \quad r^2 = x_1^2 + x_2^2 + x_3^2 \quad (6.66)$$

and determine it on the lattice. Measured values of $\langle r^2 \rangle$ can be used to tune to constant smearing radii across different lattice spacings.

If we express the smearing procedure by the operator S_k defined by

$$S_k = (1 + \kappa_G a^2 \Delta)^{n_k}, \quad (6.67)$$

we can build interpolating fields with gamma structure Γ via

$$O_k^{ij} = a^3 \sum_{\mathbf{x}} \bar{\psi}^i(\mathbf{x}) S_k \Gamma \psi^j(\mathbf{x}). \quad (6.68)$$

When Gaussian smearing is applied, we use APE smeared gauge fields in the Laplace operator. With this construction, unphysical fluctuations of the gauge fields have been smoothed out and do not propagate into the quark fields [187].

6.5.4 Smearing from three-dimensional fermions

In [188] the creation of extended interpolating operators from quark fields and additional three-dimensional spinor fields has been introduced. The implementation requires the inversion of a three-dimensional Dirac operator.

We start our discussion with the definition of smeared sources from spinor fields $\phi(\mathbf{x})$. The corresponding action on a single time slice is given by

$$S_{\text{sp}} = a^3 \sum_{\mathbf{x}} \phi(\mathbf{x}) \left[\frac{1}{2} \{ \gamma_i (\nabla_i + \nabla_i^*) - a \nabla_i^* \nabla_i \} + m_{3D} \right] \phi(\mathbf{x}), \quad (6.69)$$

and no propagation in time is allowed for the spinor fields $\phi(\mathbf{x})$. They are quenched but couple to the spatial parts of the gauge field and therefore receive quantum corrections. The extended quark field is defined from the spinor fields ϕ and the unsmeared quark fields ψ via

$$\psi_{\text{sp}}(\mathbf{x}) = a^3 \sum_{\mathbf{y} \in \Lambda(\mathbf{x}_0)} \phi(\mathbf{x}) \bar{\phi}(\mathbf{y}) \Gamma_{3D} \psi(\mathbf{x}_0, \mathbf{y}), \quad (6.70)$$

We use $\kappa_G = 0.1$

$\Lambda(\mathbf{x}_0)$ denotes the set of spatial sites on time slice \mathbf{x}_0 .

where we sum over all sites on time slice x_0 . The Dirac structure Γ_{3D} is chosen appropriately to obtain the correct quantum numbers.

When two such extended quark fields are contracted, we obtain the propagation of a three-dimensional fermion in space, the standard propagation in space-time to a different time slice and subsequently another three-dimensional propagation. Therefore an extended interpolator has been constructed from the above definition.

Spinor fields have been chosen for the construction of this smearing method since they lower the canonical dimension of the extended quark fields and have nice renormalization properties. They are automatically renormalized if the quark fields ψ are properly renormalized and due to the improved short distance behavior which can be deduced from the operator product expansion we expect an improved overlap with the physical ground state.

Similarly to quark hopping parameters, we can define the hopping parameter for the spinor fields by

$$am_{3D} = \frac{1}{2} \left(\frac{1}{\kappa_{3D}} - 6 \right), \quad (6.71)$$

and the critical hopping parameter can be obtained from the point where the mass of a corresponding three-dimensional pion vanishes.

6.5.5 Smearing from three-dimensional bosons

The construction of extended quark fields from quenched, three-dimensional fields can also be defined from scalar fields [188]. In this case, we use the scalar fields $\phi(x)$ with the lattice action

$$S_{sc} = a^3 \sum_x \left(-\frac{1}{2} \phi(x) \Delta \phi^\dagger(x) + \frac{m_{3D}^2}{2} \phi^\dagger(x) \phi(x) \right) \quad (6.72)$$

on the time slice x_0 with the 3D Laplace operator as it is used in Gaussian smearing. The extended quark field is then defined by

$$\psi_{\alpha,sc}(x) = a^3 \sum_{y \in \Lambda(x_0)} \phi(x) \bar{\phi}(y) \psi_\alpha(x_0, y), \quad (6.73)$$

where we denoted the spinor index with α . In this case, we need four inversions of the scalar operator to determine all four spinor components of the extended quark field.

The 3D hopping parameter is defined as for the spinor fields, eq. (6.71). In contrast to the three-dimensional fermions, the theory with quenched three-dimensional bosons is not renormalizable and the theoretical control over the smearing across different lattice spacings is lost [188]. Nevertheless, we can use the bosons to define a valid smearing procedure if we are not interested in the renormalization properties. The canonical dimension of an interpolator built

from $\psi_{\alpha,sc}$ is still lower than in the unsmearred case and we expect the short distance properties to be improved. The tuning of the hopping parameter has to be performed via explicit measurements of the radius defined in eq. (6.66). This radius can also be used to compare interpolators based on different smearing techniques with similar smearing radii.

HEAVY QUARK EFFECTIVE THEORY

Since we are not able to cover all energy scales that contribute to physical observables including bottom quarks in large-volume simulations, we have to resort to an effective theory to describe QCD observables including one heavy quark. Our theory of choice is Heavy Quark Effective Theory (HQET) which is constructed from an expansion of the QCD Lagrangian in inverse powers of the heavy quark mass m_h . In the static limit, i.e., for one infinitely heavy quark, HQET allows us to deduce underlying features of the Lagrangian. At the same time, the expansion in powers of $1/m_h$ is fast to enough to be able to describe physics at the scale of the b quark to high precision using the next-to-leading terms in the expansion. This makes HQET an extremely useful tool to incorporate b quarks in lattice simulations

In this chapter we will sketch the derivation of the HQET Lagrangian, describe its symmetries and explain the implications on physical observables. Afterwards we will formulate the theory on the lattice and describe our strategy to renormalize the effective theory via a non-perturbative finite-volume matching with QCD.

7.1 CONTINUUM HQET

Sometimes, instead of trying to describe all features of a theory, it makes sense to focus on specific properties and to describe these by an effective theory which is valid in a well defined limit of the full theory. In our case, we are not able to describe the low energy properties of QCD while taking into account the effects of heavy quarks.

It is known that the impact of very heavy particles is small or even irrelevant at low energies. Fermi's theory of the Weak interactions [189] is an example for an effective theory which works very well if we restrict ourselves to the description of low energy processes [190]. The structure of the weak interaction cannot be resolved, if energies at the scale of hadrons are considered.

In the construction of an effective theory for heavy quarks, three steps are necessary [191]. We have to integrate out the heavy degrees of freedom in the generating functional of the full theory which is possible because the heavy particles do not appear as external states at low energies. The result of this procedure is a Lagrangian which is non-local because of the virtual particles of the full theory which are able to travel over a distance $\propto 1/m_h$. Therefore, an expansion of the effective theory in $1/m_h$ is done to disentangle short and long distance effects. In the last step which is called matching, we have

A quark is considered as heavy, if its mass is considerably larger than Λ_{QCD} .

to renormalize the coefficients of the effective theory in order to incorporate the quantum corrections of the full theory.

Before we look at the implementation of HQET and the non-perturbative matching procedure on the lattice, we investigate the effective theory and its properties in the continuum.

7.1.1 Derivation of the HQET Lagrangian

We summarize the basic steps of the derivation of the HQET Lagrange density based on [192–194]. We start with the QCD Lagrangian with a single heavy quark Q and an unspecified number of light quarks. The heavy quark, although based on the same theoretical footing, differs from the light quarks just by its mass which we consider to be significantly larger than Λ_{QCD} . We now want to separate the light from the heavy degree of freedoms and to reflect this separation in the Lagrange density.

We start with the standard QCD Lagrangian and split it into a heavy and a light part to obtain

$$\mathcal{L}_{\text{QCD}} \equiv \mathcal{L}_{\text{heavy}} + \mathcal{L}_{\text{light}} = \bar{Q}(i\not{D} - m_Q)Q + \mathcal{L}_{\text{light}}. \quad (7.1)$$

$\mathcal{L}_{\text{heavy}}$ contains the heavy quark fields Q and \bar{Q} , the mass of the heavy quark m_Q and the standard covariant derivative. From now on, we will not consider $\mathcal{L}_{\text{light}}$ anymore which is just the standard QCD Lagrange density for the light quarks.

We proceed by splitting up the heavy quark field into an upper/large component ϕ and a lower/small component χ via

$$\phi_v = \frac{1}{2}(1 + \not{v})Q = P_+Q, \quad \not{v}\phi_v = \phi_v \quad (7.2)$$

$$\chi_v = \frac{1}{2}(1 - \not{v})Q = P_-Q, \quad \not{v}\chi_v = -\chi_v \quad (7.3)$$

where we introduced the projection operators P_{\pm} with the properties

$$\begin{aligned} P_{\pm} &= \frac{1}{2}(1 \pm \not{v}), & P_{\pm}P_{\pm} &= P_{\pm}, \\ P_+P_- &= P_-P_+ = 0, & P_+ + P_- &= 1. \end{aligned} \quad (7.4)$$

These project onto the hadron's velocity v . The hadron's momentum can be written as

$$p_Q^{\mu} = m_Q v^{\mu} + k^{\mu}, \quad (7.5)$$

where the residual momentum k is much smaller than the heavy quark mass m_Q and we normalize the velocity as

$$v \cdot v = 1. \quad (7.6)$$

In this way, we can consider the heavy quark as almost static in the reference frame of the hadron. All interactions with light quarks are

small compared to the heavy quark mass and the heavy quark is almost on-shell.

We now decompose the covariant derivative into a time (\parallel) and a spatial (\perp) component,

$$\begin{aligned} D^\mu &= D_\parallel^\mu + D_\perp^\mu, & (7.7) \\ D_\parallel^\mu &= v^\mu(v \cdot D), & D_\perp^\mu &= (g^{\mu\nu} - v^\mu v^\nu)D_\nu. \end{aligned}$$

In the next step we need the relations

$$\{\mathcal{D}_\perp, \not{v}\} = 0 \quad (7.8)$$

and

$$P_\pm \gamma_\mu = \pm v_\mu + \gamma_\mu P_\mp \quad (7.9)$$

which can be derived using the definition of the projectors and the velocity of the heavy quark.

We can now rewrite the heavy part of the Lagrangian and separate the small and large components. Since mixed terms vanish, we end up with

$$\begin{aligned} \mathcal{L}_{\text{heavy}} &= (\bar{\Phi}_v + \bar{\chi}_v) \left(i\mathcal{D}_\parallel + i\mathcal{D}_\perp - m_Q \right) (\phi_v + \chi_v) \\ &= \bar{\Phi}_v (i(v \cdot D) - m_Q) \phi_v + \bar{\Phi}_v i\mathcal{D}_\perp \chi_v \\ &\quad - \bar{\chi}_v (i(v \cdot D) + m_Q) \chi_v + \bar{\chi}_v i\mathcal{D}_\perp \phi_v. \end{aligned} \quad (7.10)$$

In order to decompose the momentum of the heavy quark into a large and a residual piece, we multiply the quark fields with a phase factor,

$$\phi_v = e^{-im_Q(v \cdot x)} h_v, \quad \chi_v = e^{-im_Q(v \cdot x)} H_v \quad (7.11)$$

and with

$$i\partial^\mu \phi_v(x) = e^{-im_Q(v \cdot x)} (m_Q v^\mu + i\partial^\mu) h_v \quad (7.12)$$

$$i\partial^\mu \chi_v(x) = e^{-im_Q(v \cdot x)} (m_Q v^\mu + i\partial^\mu) H_v \quad (7.13)$$

and

$$\bar{\Phi}_v i\mathcal{D}_\perp \chi_v = \bar{h}_v i\mathcal{D}_\perp H_v \quad (7.14)$$

$$\bar{\chi}_v i\mathcal{D}_\perp \phi_v = \bar{H}_v i\mathcal{D}_\perp h_v \quad (7.15)$$

we arrive at

$$\begin{aligned} \mathcal{L}_{\text{heavy}} &= \bar{h}_v i(v \cdot D) h_v - \bar{H}_v (i(v \cdot D) + 2m_Q) H_v \\ &\quad + \bar{h}_v i\mathcal{D}_\perp H_v + \bar{H}_v i\mathcal{D}_\perp h_v. \end{aligned} \quad (7.16)$$

From this representation of the Lagrange density we can see that the field H_v corresponds to the mass $2m_Q$ but h_v is massless. The mixed

terms couple the heavy and the light degrees of freedom and describe the pair creation or annihilation of heavy quarks and antiquarks.

In the next step, we like to integrate out the heavy degrees of freedom. In this sketch of the derivation we do this to tree-level, i.e., we solve the equations of motion in order to express H_v in terms of h_v . From the field equation

$$(i(v \cdot D) + 2m_Q)H_v = i\mathcal{D}_\perp h_v \quad (7.17)$$

we derive

$$H_v = \frac{1}{(i(v \cdot D) + 2m_Q)} i\mathcal{D}_\perp h_v \quad (7.18)$$

and use this relation to obtain

$$\mathcal{L}_{\text{heavy}} = \bar{h}_v \left[i(v \cdot D) + i\mathcal{D}_\perp \frac{1}{(i(v \cdot D) + 2m_Q)} i\mathcal{D}_\perp \right] h_v. \quad (7.19)$$

As we can see from eq. (7.19), a non-local term is present in the Lagrangian as the result of integrating out the heavy degrees of freedom. Since $(v \cdot D) \ll 2m_Q$ holds, we can use the expansion

$$\frac{1}{(i(v \cdot D) + 2m_Q)} = \frac{1}{2m_Q} \sum_{n=0}^{\infty} \left(\frac{-i(v \cdot D)}{2m_Q} \right)^n \quad (7.20)$$

in the Lagrangian in eq. (7.19) and write

$$\mathcal{L}_{\text{heavy}} = \bar{h}_v i(v \cdot D) h_v - \frac{1}{2m_Q} \bar{h}_v \mathcal{D}_\perp \sum_{n=0}^{\infty} \left(\frac{-i(v \cdot D)}{2m_Q} \right)^n \mathcal{D}_\perp h_v. \quad (7.21)$$

We call the first term of this Lagrangian the static part and the accuracy of the effective theory increases with the number of terms considered in the expansion in the second term. In our work, we will consider the static theory, as well as the theory including the leading term of the expansion in $1/(2m_Q)$. Therefore we explicitly write out the first two terms,

$$\mathcal{L}_{\text{heavy}} = \mathcal{L}^{\text{stat}} + \mathcal{L}^{(1)} + \mathcal{O}\left(\frac{\Lambda_{\text{QCD}}}{m_Q^2}\right) \quad (7.22)$$

$$= \bar{h}_v i(v \cdot D) h_v - \frac{1}{2m_Q} \bar{h}_v \mathcal{D}_\perp \mathcal{D}_\perp h_v + \mathcal{O}\left(\frac{\Lambda_{\text{QCD}}}{m_Q^2}\right), \quad (7.23)$$

and with a few lines of algebra, we can express the leading correction $\mathcal{L}^{(1)}$ by [195],

$$\mathcal{L}^{(1)} = -\frac{1}{2m_Q} \bar{h}_v (iD_\perp)^2 h_v - \frac{1}{2m_Q} \bar{h}_v \left(\frac{g}{2} \sigma_{\mu\nu} F^{\mu\nu} \right) h_v \quad (7.24)$$

$$\equiv -\omega_{\text{kin}} \mathcal{O}_{\text{kin}} - \omega_{\text{spin}} \mathcal{O}_{\text{spin}}. \quad (7.25)$$

We used
 $[D^\mu, D^\mu] =$
 $-igF^{\mu\nu}$ and $\sigma_{\mu\nu}$
 defined in app. A.

At the classical level, the pre-factors are given by $\omega_{\text{kin}} = \omega_{\text{spin}} = 1/(2m_Q)$ but at the quantum level, both coefficients receive corrections. In Euclidean space and in the rest frame of the heavy quark, we write

$$\mathcal{L}_{\text{heavy}} = \bar{h}_v D_0 h_v - \omega_{\text{kin}} \mathcal{O}_{\text{kin}} - \omega_{\text{spin}} \mathcal{O}_{\text{spin}}, \quad (7.26)$$

with the kinetic and spin operators

$$\mathcal{O}_{\text{kin}} = \bar{h}_v \mathbf{D}^2 h_v \quad \mathcal{O}_{\text{spin}} = \bar{h}_v g \boldsymbol{\sigma} \cdot \mathbf{B} h_v, \quad \boldsymbol{\sigma} \cdot \mathbf{B} = \frac{-i}{2} \sigma_{ij} F^{ij}. \quad (7.27)$$

The kinetic operator describes the kinetic energy from the off-shell residual motion of the heavy quark. The spin term describes the interaction of the gluon field with the spin of the heavy quark. It contains the non-Abelian analogue of the Pauli term. Based on this Lagrangian, we can now discuss the properties of the effective theory, before we formulate it on the lattice.

7.1.2 Heavy Quark Symmetry

The static Lagrangian has more symmetries as the Lagrange density of QCD and they are subsumed under the term heavy quark symmetry [191]. We see that no Dirac matrices are present in $\mathcal{L}^{\text{stat}}$ and therefore interactions of the heavy quark with gluons leave its spin unchanged [196]. The Lagrangian is invariant under the $SU(2)$ symmetry group of spin rotations. Since the heavy quark mass is not present in the static Lagrangian, it is also invariant under rotations in the flavor space of N_h heavy quark flavors [197]. Therefore, the symmetry group of this combined spin-flavor symmetry is $SU(2N_h)$. In the limit of infinitely heavy quarks, the strong interaction becomes independent of the quark's mass and spin. It is important to note that these symmetries are symmetries of the effective theory and not of the QCD Lagrangian, even when the limit of heavy quarks is considered. The terms in $\mathcal{L}^{(1)}$ break heavy quark symmetry explicitly. Therefore, we expect symmetry violations to increase when the heavy quark mass decreases.

This situation is somewhat different to chiral perturbation theory.

Heavy quark symmetry has implications on the spectroscopy of states containing heavy quarks. We can expand the mass of a meson with one heavy quark in the inverse heavy quark mass and obtain [198]

$$m_H = m_Q + \bar{\Lambda} + \frac{\Delta m^2}{2m_Q} + \mathcal{O}(1/m_Q^2), \quad (7.28)$$

where all contributions without dependence on the heavy quark are collected in $\bar{\Lambda}$ [199]. From the factor of $1/m_Q$, we see that all contributions stemming from the next-to-leading order terms are collected in Δm^2 . We can express this term by [199, 200]

$\bar{\Lambda}$ depends on the flavor of the light quark.

$$\Delta m^2 = -\lambda_1 + 2 \left[J(J+1) - \frac{3}{2} \right] \lambda_2, \quad (7.29)$$

where J is the total spin of the meson. The term λ_1 parametrizes the kinetic energy of the heavy quark inside the meson and the second term describes the interaction of the gluon field with the heavy quark spin [194]. The values of $\bar{\Lambda}$, λ_1 and λ_2 are properties of the light constituents of the meson and therefore independent of the heavy quark mass. We see that the spin interaction induces a hyperfine splitting of states with same light quark spin j but different total angular momentum $J = j \pm \frac{1}{2}$ which are degenerate in the limit of infinitely heavy quarks.

The symmetries of HQET allow a so-called residual mass term δm for the heavy quark which is of order $O(\Lambda_{\text{QCD}})$. If it is not included in the static Lagrangian, it will be induced by quantum corrections [199]. If $\delta m = 0$ is chosen the heavy quark mass coincides with the pole mass in perturbation theory. From this choice, the binding energy

$$\bar{\Lambda} = \lim_{m_Q \rightarrow \infty} (m_H - m_Q) \quad (7.30)$$

can be defined. If $\delta m \neq 0$ is chosen, the binding energy changes such that the physical observables remain unchanged under this change.

7.2 HQET ON THE LATTICE

Having discussed the derivation of the continuum Lagrangian of HQET and the symmetry properties of the effective theory, we now turn to the regularization on the lattice. The two light quarks and the strange quark are unaffected by this discussion and they are still described by the actions defined in chapter 2. Pedagogical introductions concerning the implementation of HQET on the lattice can be found in [41, 201, 202] and we have used them as basis for the discussion in this section.

7.2.1 Lattice action

If we choose the rest frame of the B-meson as our reference frame, the static part of the Lagrangian is given by the first term in eq. (7.26). If we name the quark fields by $\psi_h, \bar{\psi}_h$ to match the notation of the literature and replace the covariant derivative by the lattice version of the backward covariant derivative given by

$$\nabla_0^* \psi(x) = \frac{1}{a} \left[\psi(x) - U_0^\dagger(x - a\hat{0}) \psi(x - a\hat{0}) \right], \quad (7.31)$$

we arrive at the so-called Eichten-Hill action [195, 203, 204]. In this action, no doublers are present and the leading discretization effects are

of order $O(\alpha^2)$ [205]. We can express the lattice action corresponding to eq. (7.26) via [206]

$$S_{\text{HQET}} = \alpha^4 \sum_{\mathbf{x}} \left\{ \mathcal{L}_{\text{stat}}(\mathbf{x}) + \sum_{\nu=1}^n \mathcal{L}^{(\nu)}(\mathbf{x}) \right\}, \quad \text{with} \quad (7.32)$$

$$\mathcal{L}_{\text{stat}} = \bar{\psi}_h(\mathbf{x}) [\nabla_0^* + \delta m] \psi_h, \quad (7.33)$$

$$\mathcal{L}^{(\nu)} = \sum_{\mathbf{i}} \omega_{\mathbf{i}}^{(\nu)} L_{\mathbf{i}}^{(\nu)}(\mathbf{x}), \quad (7.34)$$

where we have included a mass counter-term δm which has mass-dimension one. This counter-term is needed to cancel a linear divergence in the self-energy of the static quark [204]. In our calculations we will set it to zero in the lattice action and absorb its effect in the redefinition of the b-quark mass. The $1/m$ corrections $\mathcal{L}^{(\nu)}$ are given by the second and third term of eq. (7.26) with discretized versions of the products $\sigma \cdot \mathbf{B}$ and \mathbf{D}^2 .

In exchange, the b-quark mass becomes linearly divergent.

In practice, we will replace the naive lattice derivative by the more general version

$$\nabla_W^* \psi(\mathbf{x}) = \frac{1}{\alpha} \left[\psi(\mathbf{x}) - W_0^\dagger(\mathbf{x} - \alpha \hat{0}) \psi(\mathbf{x} - \alpha \hat{0}) \right]. \quad (7.35)$$

Here W_0 is a time-like transporter which differs from U_0 since it is based on a HYP smeared gauge field. In [185] it has been shown that an exponentially improved statistical precision can be gained by this choice of the derivative. In our work, we will consider HYP1 and HYP2 smearing as introduced in section 6.5.2.

The heavy quark propagator to static order can be obtained from the solution of [202]

$$D_0 G_h(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x}, \mathbf{y}) P_+. \quad (7.36)$$

Since the covariant derivative is given in terms of gauge fields, no inversion of the Dirac matrix is necessary and we can compute the propagator analytically via

$$G_h(\mathbf{x}, \mathbf{y}) = \theta(x_0 - y_0) \delta^{(3)}(\mathbf{x} - \mathbf{y}) \mathcal{P}(\mathbf{y}, \mathbf{x})^\dagger P_+. \quad (7.37)$$

In this equation we have used the Wilson line from \mathbf{x} to \mathbf{y} which is defined by

$$\begin{aligned} \mathcal{P}(\mathbf{x}, \mathbf{x}) &= 1, \\ \mathcal{P}(\mathbf{x}, \mathbf{x} + N\hat{\mu}) &= W_\mu(\mathbf{x}) W_\mu(\mathbf{x} + \alpha\hat{\mu}) \cdots W_\mu(\mathbf{x} + N\alpha\hat{\mu}). \end{aligned} \quad (7.38)$$

We can see from this definition that the static heavy quark propagates only forward in time. This is expected since we choose the rest frame of the B-meson as reference frame. Accordingly, a heavy anti-quark propagates backward in time.

7.2.2 Expectation values at next-to-leading order

We cannot define the path integral from the HQET action at order $1/m$ since the action contains the dimension five operators \mathcal{O}_{kin} and $\mathcal{O}_{\text{spin}}$. The path integral is not renormalizable when these operators are included in the action. To circumvent this issue, we take an approach that is similar to Symanzik's idea of an effective lattice theory that we have used in the context of improved lattice actions.

We can expand the Boltzmann factor in the path integral in powers of the inverse heavy quark mass $1/m_h$,

$$\begin{aligned} e^{-(S_{\text{light}}+S_{\text{HQET}})} &= \exp \left\{ - \left(S_{\text{light}} + a^4 \sum_{\mathbf{x}} \mathcal{L}_{\text{stat}}(\mathbf{x}) \right) \right\} \\ &\times \left[1 - a^4 \sum_{\mathbf{x}} \mathcal{L}^{(1)}(\mathbf{x}) + \frac{1}{2} a^4 \left(\sum_{\mathbf{x}} \mathcal{L}^{(1)}(\mathbf{x}) \right)^2 \right. \\ &\left. - a^4 \sum_{\mathbf{x}} \mathcal{L}^{(2)}(\mathbf{x}) + \dots \right] \end{aligned} \quad (7.39)$$

and see that only one term in the expansion contributes at order $O(1/m_h)$. Based on this expansion we can determine the expectation value of an operator \mathcal{O} via

$$\begin{aligned} \langle \mathcal{O} \rangle &\approx \frac{1}{Z} \int \mathcal{D}\phi \mathcal{O} \left(1 + a^4 \sum_{\mathbf{x}} [\omega_{\text{kin}} \mathcal{O}_{\text{kin}}(\mathbf{x}) \right. \\ &\quad \left. + \omega_{\text{spin}} \mathcal{O}_{\text{spin}}(\mathbf{x})] \right) e^{-(S_{\text{light}}+S_{\text{HQET}}^{\text{stat}})} \\ &= \langle \mathcal{O} \rangle_{\text{stat}} + \omega_{\text{kin}} a^4 \sum_{\mathbf{x}} \langle \mathcal{O} \mathcal{O}_{\text{kin}} \rangle_{\text{stat}} + \omega_{\text{spin}} a^4 \sum_{\mathbf{x}} \langle \mathcal{O} \mathcal{O}_{\text{spin}} \rangle_{\text{stat}}. \end{aligned} \quad (7.40)$$

The expectation value $\langle \cdot \rangle_{\text{stat}}$ is determined from the path integral including the light action and the static part of the heavy Lagrangian. As shorthand notation for eq. (7.40) we will write

$$\langle \mathcal{O} \rangle \equiv \langle \mathcal{O} \rangle_{\text{stat}} + \omega_{\text{kin}} \langle \mathcal{O} \rangle_{\text{kin}} + \omega_{\text{spin}} \langle \mathcal{O} \rangle_{\text{spin}}. \quad (7.41)$$

At this point we like to note that the $O(a)$ improvement of the action, as we have done it in chapter 2 for the light part of the action, is crucial for a determination of observables at next-to-leading order of HQET. Since terms of $O(a)$ and terms of $O(1/m_h)$ are both of mass dimension minus one they will mix on the lattice. If operators of $O(a)$ are present, we cannot disentangle their contributions from those of the terms belonging to $O(1/m_h)$. With the improved relativistic action as we have introduced it in chapter 2 and the static part of the HQET Lagrangian which has leading discretization effects of $O(a^2)$, we are in the position to include next-to-leading terms of HQET. The leading corrections present in our theory are then of order $O(a^2, a/m_h)$.

7.3 NON-PERTURBATIVE MATCHING WITH QCD

We have introduced the matching parameters ω_{kin} and ω_{spin} in the classical derivation of the HQET Lagrangian. At the quantum level, these parameters receive corrections and we have to properly renormalize them to take into account all loop effects of the dynamical theory. On the lattice, the HQET parameters have an explicit dependence on the lattice spacing and the heavy quark mass.

A perturbative determination of the matching coefficients can only be done up to some order g_0^{2l} and the mixing of operators differing in dimensions by p translates into coefficients diverging as a^{-p} . After the perturbative determination of a coefficient c_k we are left with the remainder [206]

$$\Delta c_k \sim g_0^{2(l+1)} a^{-p} \sim a^{-p} [\ln(a\Lambda)]^{-(l+1)} \xrightarrow{a \rightarrow 0} \infty. \quad (7.42)$$

Therefore, the continuum limit of the discretized theory does not exist if the matching coefficients are determined perturbatively.

7.3.1 Matching in finite volume

We will renormalize the HQET parameters by the requirement that the effective theory describes the same physics as QCD including the relativistic heavy quark. For each free parameter of the effective theory we have to define a so-called matching condition. In principle we could determine physical observables and match them with experimental input. With this approach the predictive power of the effective theory is quickly lost, even at leading order, since we do not only have to fix the parameters of the Lagrangian, but also the renormalization parameters of composite operators. Therefore, we have to take a different path.

We like to fix the HQET parameters by results gained from simulations of full QCD on the lattice. If we tried to do this in large volume, we had the same problems which led to the consideration of the effective theory: We are not able to compute large volume quantities involving bottom quarks since we are not able to choose lattice spacings that are small enough to sufficiently suppress discretization effects.

As pioneered in [206], this unpleasant situation can be resolved if we consider observables defined in small volumes, provided that both QCD and HQET are well-defined in this volume and the parameters in the Lagrangians are volume-independent.

We define the matching equations to fix HQET parameters Φ_i^{HQET} in finite volume to be

$$\Phi_i^{\text{HQET}}(L, M, a) = \Phi_i^{\text{QCD}}(L, M, 0) \quad (7.43)$$

with

$$\Phi_i^{\text{QCD}}(L, M, 0) = \lim_{a \rightarrow 0} \Phi_i^{\text{QCD}}(L, M, a), \quad (7.44)$$

i.e., we fix the parameters of HQET at finite lattice spacing and in finite volume by the continuum limit of the corresponding QCD observables in the same volume. We have made the dependence of the matching parameters on the matching volume L , the renormalized heavy quark mass M and the lattice spacing a explicit.

The simulations to perform this matching have to satisfy some constraints. The lattice spacing of the simulations which are used to determine the continuum limit in eq. (7.44) has to be small enough to ensure $a \gg m_b$ to control all discretization effects. At the same time, the spatial volume L^3 has to be large enough to allow for a reliable expansion in $1/m$. This leads to the constraint $1/L \ll m_b$. Due to the limitation in computing resources, we are limited to lattices with $L/a = O(10)$. The combination of these constraints leads to the choice of a volume of about $L \approx 0.5 \text{ fm}$ and lattice spacings of the QCD ensembles in the range $0.02 \text{ fm} > a > 0.008 \text{ fm}$.

In contrast to the matching observables, physical observables have to be computed in large volume and the main objective of the matching procedure is to determine the HQET parameters in the region of large volume and lattice spacings in the region $0.1 \text{ fm} > a > 0.04 \text{ fm}$. Therefore, we will use a step scaling of the parameters from the matching region to the region of large volume simulations.

Before we detail the process of the matching in finite volume and the step scaling procedure we like to mention how to constrain the light parameters of the theory. Since the action of the light quarks remains the same when switching from QCD to HQET, we can trivially fulfill the matching condition of eq. (7.43) for the light degrees of freedom. We will fix the volume of the lattices and the light quark masses from the conditions

$$\Phi_1^{\text{HQET}}(L) = \bar{g}^2(L) = \text{const}, \quad \text{and} \quad \Phi_2^{\text{HQET}}(L) = L m_l = 0. \quad (7.45)$$

These parameters are independent of the heavy quark mass and by fixing them we obtain parameter triples $(L/a, \beta, \kappa_l)$ which allow us to take the continuum limit.

7.3.2 Determination of the matching parameters

We now explain the full matching procedure along the lines of [207] and figure 7.1 which is taken from this reference. We denote the spatial extent of the matching volume by L_1 and obtain the matching equation

$$\Phi_i^{\text{HQET}}(L_1, M, a) = \Phi_i^{\text{QCD}}(L_1, M, 0). \quad (7.46)$$

Remember that we work with three mass-degenerate quark flavors in the sea.

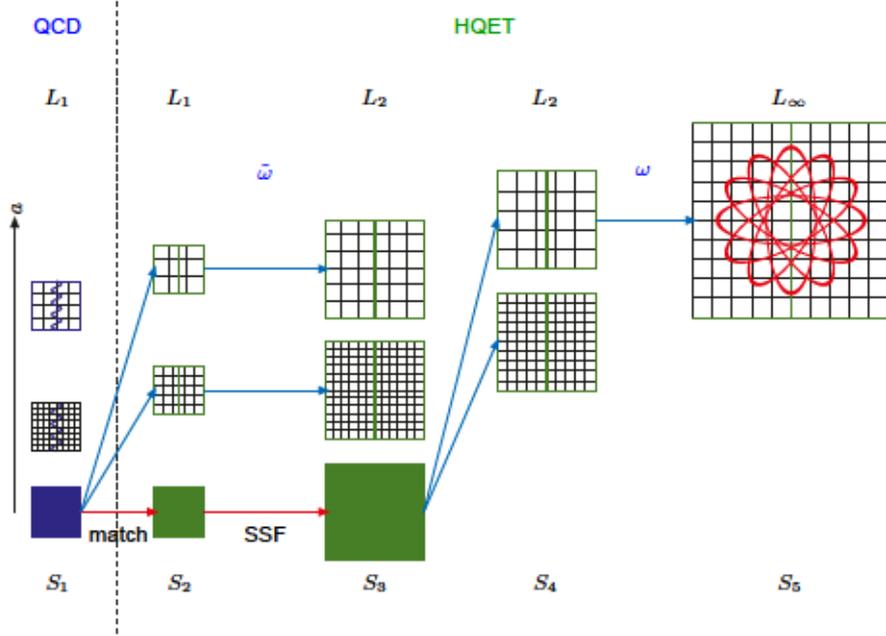


Figure 7.1: ALPHA's strategy for B-physics via a non-perturbative finite-volume matching of QCD and HQET. Taken from [207].

As a first step, we have to set up ensembles with small lattice spacings to obtain the continuum limit of the matching observables with relativistic heavy quarks $\Phi_i^{\text{QCD}}(L_1, M, 0)$, as depicted as S_1 in fig. 7.1.

In the next step, we determine the corresponding observables in the effective theory on a set of ensembles with the same spatial extent L_1 but coarse lattice spacings. We formally split the HQET observables into two parts

$$\Phi_i^{\text{HQET}}(L, M, a) = \eta(L, a) + \phi(L, a)\omega(M, a) + O(1/m_h^2) \quad (7.47)$$

such that the dependence on the heavy quark mass M is contained in the factor $\omega(M, a)$ and we include effects of order $O(1/m_h)$. Based on this representation, we can solve eq. (7.46) to obtain the matching parameters in the matching volume L_1 which we call $\tilde{\omega}$ from

$$\tilde{\omega}(M, a) \equiv \phi_{S_2}^{-1}(L_1, a) [\Phi^{\text{QCD}}(L_1, M, 0) - \eta_{S_2}(L_1, a)]. \quad (7.48)$$

With this procedure, we have combined steps S_1 and S_2 from 7.1 to determine the matching parameters in finite volume.

In the next step S_3 , we determine the matching parameters on lattices with a spatial extent $L_2 = 2L_1$ via a step scaling procedure. We therefore have to generate another set of ensembles. These ensembles have the same bare parameters as the HQET ensembles in the matching volume but twice the number of lattice points in each direction. Since the lattice spacing remains the same, the physical extent is scaled by a factor of two.

We suppress the index i and indicate on which set of ensembles the bare observables have been determined, instead.

The HQET parameters in L_2 are determined via

$$\Phi_{S_3}^{\text{HQET}}(L_2, M, 0) = \lim_{a \rightarrow 0} \{\eta_{S_3}(L_2, a) + \phi_{S_3}(L_2, a)\tilde{\omega}(M, a)\} \quad (7.49)$$

using $\eta_{S_3}(L_2, a)$ and $\phi_{S_3}(L_2, a)$ determined on the ensembles with $L = L_2$ and $\tilde{\omega}(M, a)$ from the matching volume. Since the range of lattice spacings in this step has only a small overlap with the lattice spacings of the large-volume CLS simulations, we have to include another step in the procedure.

The final matching parameters ω_i for the renormalization in large volume (S_5) are determined using additional simulations in the volume L_2^3 but with coarser lattice spacings than in S_3 . This set of ensembles is called S_4 in figure 7.1. The resolutions of the ensembles in this set cover the range of resolutions of the large-volume simulations.

The final matching parameters for the renormalization of the effective theory in large volume are then determined via

$$\omega(M, a) \equiv \phi_{S_4}^{-1}(L_2, a) \left[\Phi_{S_3}^{\text{HQET}}(L_2, M, 0) - \eta_{S_4}(L_2, a) \right]. \quad (7.50)$$

We indicate explicitly by the subscript S_4 that $\phi_{S_4}^{-1}(L_2, a)$ and $\eta_{S_4}(L_2, a)$ have been determined at the resolutions of S_4 which are different from the ones in eq. (7.49).

Depending on the chosen set of matching observables, the matching equations can be defined in a different way to improve the precision of the calculation. Therefore, the strategy described above only gives an introduction in the general setup of the matching. A full description of the strategy as it is used in the two flavor theory is given in [208].

The setup of the line of constant physics and the generation of the four groups of ensembles is independent of the specific choice of the matching observables. Since the tuning towards the LCP and the generation of the $O(25)$ ensembles is the main part of this work, we do not go into more detail concerning the matching conditions.

7.3.3 Quark mass dependence

As indicated in the matching condition in eq. (7.43), the parameters of the effective theory have an explicit dependence on the mass of the heavy quark in the QCD simulations. To map out this dependence and to be able to interpolate the matching parameters to the point where the heavy quark mass has the physical value of the bottom quark mass, we scan a range of renormalized quark masses around the bottom quark mass.

We will parametrize the dependence in terms of the dimensionless quantity $z = L_1 M$, where L_1 is spatial extent of the matching volume and M is the Renormalization Group Invariant (RGI) mass. In order to fix the hopping parameters for the determination of the matching

We will comment on the determination of RGI quark masses in section 8.3.

parameters at fixed renormalized quark mass across all ensembles in the matching volume, we will use eq. (4.17) in a situation with heavy valence and massless sea quarks. Therefore, we have to determine the coefficients Z_m and b_m in the coupling region of the matching. This task will be done in chapter 12.

We can employ the value of $M_b = 6.874(66)$ GeV from [209] based on [210] for $N_f = 2 + 1$ flavor QCD to get an estimate for z in the matching region. With the conversion factor 197.326980 MeV fm from [2] we can convert the RGI value to $M_b = 34.8(34)$ fm $^{-1}$ and with $L_1 \approx 0.5$ fm we determine $z \approx 17.4$ for the RGI bottom quark mass in the matching volume.

Based on this argument, we will determine the matching parameters for a range of values of z around 17.4 and parametrize them in z . A dedicated measurement in large volume will allow us to determine the bottom quark mass on the CLS ensembles based on our non-perturbative matching a posteriori.

7.4 HQET OBSERVABLES

Whereas the matching is done in finite volume, the physical observables which we want to determine in the effective theory, will be evaluated on large-volume ensembles. In the context of this work, we began to measure some of the correlation functions to determine basic observables as the mass of the bottom quark and the B-meson decay constant. Both observables can only be determined properly when the measurements in the effective theory are combined with the HQET parameters obtained in the matching procedure. Since both observables are based on the axial current, we introduce its representation in the effective theory.

7.4.1 The axial current in HQET

As stated in the last section, renormalization, improvement and the inclusion of next-to-leading terms in the HQET expansion are closely intertwined. The time component of the renormalized and improved axial current in HQET is given by [205, 211]

$$A_0^{\text{HQET}} = Z_A^{\text{HQET}} (1 + b_A^{\text{stat}} a m_q) \left(A_0^{\text{stat}} + a c_A^{(1)} \delta A_0^{(1)} + a c_A^{(2)} \delta A_0^{(2)} \right), \quad (7.51)$$

where the coefficients Z_A^{HQET} , b_A^{stat} , $c_A^{(1)}$ and $c_A^{(2)}$ depend on the coupling but not on the quark mass. The currents are given by

$$A_0^{\text{stat}}(x) = \bar{\psi}_q(x) \gamma_0 \gamma_5 \psi_h(x) \quad (7.52)$$

and

$$\delta A_0^{(1)}(x) = \bar{\psi}_q(x) \frac{1}{2} \gamma_5 \gamma_i (\nabla_i^S - \overleftarrow{\nabla}_i^S) \psi_h(x) \quad (7.53)$$

$$\delta A_0^{(2)}(x) = \bar{\psi}_q(x) \frac{1}{2} \gamma_5 \gamma_i (\nabla_i^S + \overleftarrow{\nabla}_i^S) \psi_h(x) \quad (7.54)$$

where the light quark fields are labeled by the subscript q . The symmetrized spatial derivatives are given by

$$\nabla_i^S = \frac{1}{2} (\nabla_i + \nabla_i^*) \quad \overleftarrow{\nabla}_i^S = \frac{1}{2} (\overleftarrow{\nabla}_i + \overleftarrow{\nabla}_i^*). \quad (7.55)$$

The current $\delta A_0^{(2)}$ vanishes for correlation functions defined at vanishing momentum. Therefore we do not have to consider it in the determination of the b quark mass and decay constants.

7.4.2 The B meson mass

The correlation function corresponding to the axial current is now given by [201]

$$\begin{aligned} C_{AA} &= e^{-m_{\text{bare}} x_0} \alpha^3 \left(Z_A^{\text{HQET}} \right)^2 (1 + b_A^{\text{stat}} \alpha m_q)^2 \\ &\times \left[C_{AA}^{\text{stat}}(x_0) + c_A^{(1)} C_{\delta AA}^{\text{stat}}(x_0) + \omega_{\text{kin}} C_{AA}^{\text{kin}}(x_0) + \omega_{\text{spin}} C_{AA}^{\text{spin}}(x_0) \right] \end{aligned} \quad (7.56)$$

with the correlation functions

$$C_{AA}^{\text{stat}}(x_0) = \alpha^3 \sum_{\mathbf{x}} \langle A_0^{\text{stat}}(x) (A_0^{\text{stat}}(0))^\dagger \rangle_{\text{stat}} \quad (7.57)$$

$$C_{AA}^{\text{kin}}(x_0) = \alpha^3 \sum_{\mathbf{x}} \langle A_0^{\text{stat}}(x) (A_0^{\text{stat}}(0))^\dagger \rangle_{\text{kin}} \quad (7.58)$$

$$C_{AA}^{\text{spin}}(x_0) = \alpha^3 \sum_{\mathbf{x}} \langle A_0^{\text{stat}}(x) (A_0^{\text{stat}}(0))^\dagger \rangle_{\text{spin}} \quad (7.59)$$

$$\begin{aligned} C_{\delta AA}^{\text{stat}}(x_0) &= \alpha^3 \sum_{\mathbf{x}} \langle A_0^{\text{stat}}(x) (A_0^{(1)}(0))^\dagger \rangle_{\text{stat}} \\ &+ \alpha^3 \sum_{\mathbf{x}} \langle A_0^{(1)}(x) (A_0^{\text{stat}}(0))^\dagger \rangle_{\text{stat}}. \end{aligned} \quad (7.60)$$

The exponential factor is included to compensate for a power divergence. From this correlation function, we can determine the B -meson mass via

$$m_B = - \lim_{t \rightarrow \infty} \log C_{AA}(t) = m_{\text{bare}} + E^{\text{stat}} + \omega_{\text{kin}} E^{\text{kin}} + \omega_{\text{spin}} E^{\text{spin}} \quad (7.61)$$

with the HQET matrix elements

$$E^{\text{stat}} = - \lim_{t \rightarrow \infty} \frac{\partial_0 + \partial_0^*}{2} \log [C_{AA}^{\text{stat}}(t)] \quad (7.62)$$

$$E^{1/m} = - \lim_{t \rightarrow \infty} \frac{\partial_0 + \partial_0^*}{2} \frac{C_{AA}^{1/m}(t)}{C_{AA}^{\text{stat}}(t)}, \quad 1/m \in [\text{spin}, \text{kin}]. \quad (7.63)$$

The contribution of $C_{\delta AA}^{\text{stat}}$ vanishes for the meson mass and the bare mass

$$m_{\text{bare}} = m_b + \delta m \quad (7.64)$$

is given by the b quark mass and a counter term. To determine the B-meson mass, we therefore have to calculate m_{bare} as one of the matching parameters.

7.4.3 The GEVP in HQET

As already indicated in section 6.2, we can use the GEVP method to suppress excited state contributions in context of large-volume HQET observables. The method to include terms of next-to-leading order has been developed in [147]. In the derivation, the $1/m$ corrections are considered as small perturbations of the static correlation functions.

Section 4 of [147] explains the method using static-light B_s mesons as example.

Instead of simple correlation functions, we now use matrices of correlation functions, based on N different interpolating operators, as given in eq. (6.9), and solve the GEVP (6.13) for the static correlation matrix

$$C^{\text{stat}}(t)v_n^{\text{stat}}(t, t_0) = \lambda_n^{\text{stat}}(t, t_0)C^{\text{stat}}(t_0)v_n^{\text{stat}}(t, t_0). \quad (7.65)$$

The generalized eigenvalues and eigenvectors from this solution can then be used to compute the energy levels via

$$E_n^{\text{eff,stat}}(t, t_0) = -\frac{1}{a} (\log [\lambda_n^{\text{stat}}(t + a, t_0)] - \log [\lambda_n^{\text{stat}}(t, t_0)]) \quad (7.66)$$

$$E_n^{\text{eff},1/m}(t, t_0) = \frac{\lambda_n^{1/m}(t, t_0)}{\lambda_n^{\text{stat}}(t, t_0)} - \frac{\lambda_n^{1/m}(t + a, t_0)}{\lambda_n^{\text{stat}}(t + a, t_0)} \quad (7.67)$$

for $1/m \in [\text{spin}, \text{kin}]$ and the ratio of eigenvalues is given by

$$\frac{\lambda_n^{1/m}(t, t_0)}{\lambda_n^{\text{stat}}(t, t_0)} = \left(v_n^{\text{stat}}(t, t_0), \left[[\lambda_n^{\text{stat}}(t, t_0)]^{-1} C^{1/m}(t) - C^{1/m}(t_0) \right] v_n^{\text{stat}}(t, t_0) \right). \quad (7.68)$$

Based on these formulas for the effective energies, we are able to compute the B meson mass with reduced systematic errors compared to eqs. (7.62–7.63).

Part II

THE LINE OF CONSTANT PHYSICS FOR A NON-PERTURBATIVE MATCHING OF HQET AND QCD

SIMULATIONS ALONG A LINE OF CONSTANT PHYSICS

In section 7.3 we have described our matching strategy and specified how we translate the finite-volume results to parameters of the effective theory in large volume, where we like to compute physical observables.

In this chapter, we explain how we set up the line of constant physics and how the tuning of the bare parameters of the theory has been performed. We will also discuss results concerning the step scaling functions for the renormalized coupling and the renormalized quark mass and compare them with the results from [45, 64]. We will focus on the physics in this chapter and discuss the technical and algorithmic challenges in the ensemble generation in chapter 9.

We perform the finite-volume simulations in the Schrödinger functional and use the tree-level Symanzik improved gauge action. Relativistic sea and valence quarks are represented via clover improved Wilson quarks. All details concerning the setup of simulations in the Schrödinger functional and the discretized action have been given in the chapters 2 and 3. In all simulations connected to the matching we set c_t and \tilde{c}_t to their one-loop values according to eqs. (3.18–3.25) and the boundary θ angles to 0.5. Therefore, our setup matches the one given in [45].

8.1 THE LINE OF CONSTANT PHYSICS: QCD

The line of constant physics in finite volume is defined by the volume, the masses of the light quarks and the mass of the heavy quark. Since heavy quarks are only included as valence quarks, we do not have to consider this part of the LCP definition at the level of the ensemble generation.

To simulate relativistic heavy quarks with our computing resources, we aim at a lattice extent of ≈ 0.5 fm. In line with our strategy of mass-independent renormalization schemes, we will include three massless quark flavors in our simulations. Therefore, the definition of the light part of the line of constant physics is given by

$$\bar{g}_{\text{GF}}^2(L_1/2) = g_*^2, \quad m_l = 0 \quad (8.1)$$

where the gradient flow coupling used to fix the spatial extent is defined in eq. (6.46) and m_l denotes the mass of the three mass-degenerate quark flavors.

As it can be seen from eq. (8.1), we do not define our line of constant physics in the matching volume L_1 but in the volume $L_0 \equiv L_1/2$. This

is done for technical reasons. We will fix the exact value of g_*^2 by the simulation with the finest resolution, i.e., the lattice with the largest value of L/a . For all coarser lattice spacings, we will have to do a series of simulations in order to tune the bare parameters such that eq. (8.1) is fulfilled on every ensemble.

If the tuning was done in the matching volume, we would have to simulate multiple ensembles with up to 48 lattice points in all four dimensions. This is not feasible with our computing resources. Therefore, we fix the line of constant physics in the volume L_0 . In the matching volume, the LCP will be violated by cut-off effects of order $O(a^2)$, but we will recover a correct result in the continuum limit which is taken before the matching is done, cf. eq. (7.43).

In the tuning volume L_0 , we generate ensembles at five different resolutions corresponding to $L_0/a \in \{12, 16, 20, 24, 32\}$. We will choose $T = L$ for all ensembles such that the volume is given by $V = L_0^4$. With this setup we are able to perform a reliable continuum extrapolation on five lattices in the volume $L_1 = 2L_0$. We will fix g_*^2 based on the measured value on the ensemble with $L_0/a = 32$ to save the cost for any tuning at this most expensive ensemble in L_0 .

8.1.1 The running of the coupling

We like to achieve $L_0 \approx 0.25$ fm, but the exact value is irrelevant as long as the physical extent remains constant on all ensembles. An initial guess for the bare coupling g_0^2 for our simulation with $L_0 = 32a$ can be inferred from the investigations in [45, 212]. There, the evolution of the renormalized coupling in the range $L \in [0.05 \text{ fm}, 1.0 \text{ fm}]$ has been determined using a large set of Schrödinger functional ensembles with the same setup as we have described it in chapter 3.

The scale at the upper end of this range is denoted by $1/L_{\text{had}}$ and since $L_{\text{had}} \approx 1$ fm, we know that we can reach $L \approx 0.25$ fm by two steps in a step scaling, where the lattice extent is divided by two. The corresponding step scaling function $\sigma(u)$ [213] is given by

$$\sigma(u) = \bar{g}^2(2L) \Big|_{\bar{g}^2(L)=u} \quad (8.2)$$

and can be considered as discrete version of the renormalization group β function [45]. From this β function $\sigma(u)$ is implicitly defined via

$$\log(2) = - \int_{\sqrt{u}}^{\sqrt{\sigma(u)}} \frac{dx}{\beta(x)} \quad (8.3)$$

with

$$\beta(\bar{g}) = -L \frac{\partial \bar{g}(L)}{\partial L} \Big|_{\bar{g} \rightarrow 0} \sim -b_0 \bar{g}^3 - b_1 \bar{g}^5 + \dots \quad (8.4)$$

where the two leading coefficients are independent of the renormalization scheme and given by

$$b_0 = \frac{11 - \frac{2}{3}N_f}{(4\pi)^2}, \quad b_1 = \frac{102 - \frac{38}{3}N_f}{(4\pi)^2}. \quad (8.5)$$

This expansion of the β function in the coupling is only applicable in the perturbative regime of QCD. Using the non-perturbative results from [45], the running can be done in the strongly-coupled region of our simulations.

When $\sigma(u)$ is known, we are able to perform the inverse step scaling from L_{had} down to L_0 . In [45] the non-perturbative β function is parametrized by a polynomial via

$$\beta(g) = -\frac{g^3}{P(g^2)} \quad P(g^2) = p_0 + p_1 g^2 + p_2 g^4 + \dots \quad (8.6)$$

and the corresponding form of the step scaling function reads

$$\begin{aligned} \log(2) &= -\int_{\sqrt{u}}^{\sqrt{\sigma(u)}} \frac{dx}{\beta(x)} = \int_{\sqrt{u}}^{\sqrt{\sigma(u)}} dx \frac{P(x^2)}{x^3} \\ &= -\frac{p_0}{2} \left[\frac{1}{\sigma(u)} - \frac{1}{u} \right] + \frac{p_1}{2} \log \left[\frac{\sigma(u)}{u} \right] \\ &\quad + \sum_{n=1}^{n_{\text{max}}} \frac{p_{n+1}}{2n} [\sigma^n(u) - u^n]. \end{aligned} \quad (8.7)$$

The parameters p_0 , p_1 and p_2 have been determined in [45] and their values and covariances are given in eqs. (4.15–4.16) of that reference.

Starting from $\bar{g}_{\text{GF}}^2(L_{\text{had}}) = 11.31$, we can perform two steps of the inverse step scaling function to arrive at

$$\sigma^{-1}(\sigma^{-1}(\bar{g}_{\text{GF}}^2(L_{\text{had}}))) = 3.95(18), \quad (8.8)$$

for the approximate target of the coupling in L_0 .

8.1.2 Interpolation of the coupling at fixed L/a

In tables 1 and 8 of [45] the gradient flow couplings which have been determined from a large set of ensembles that was used for the calculation of the step scaling function are listed. We can employ these results to perform interpolations at fixed L/a to determine the bare coupling g_0^2 which results in a fixed renormalized coupling \bar{g}_{GF}^2 . The available data can be extended by the numbers in table IV of [214].

A similar procedure has been performed in [45]. As described in A.2 of that reference, a Padé ansatz has been used to fit

$$v(\beta) = \frac{1}{\bar{g}_{\text{GF}}^2(L, \beta)}, \quad (8.9)$$

The final fits in [45] are polynomials of degree three and four.

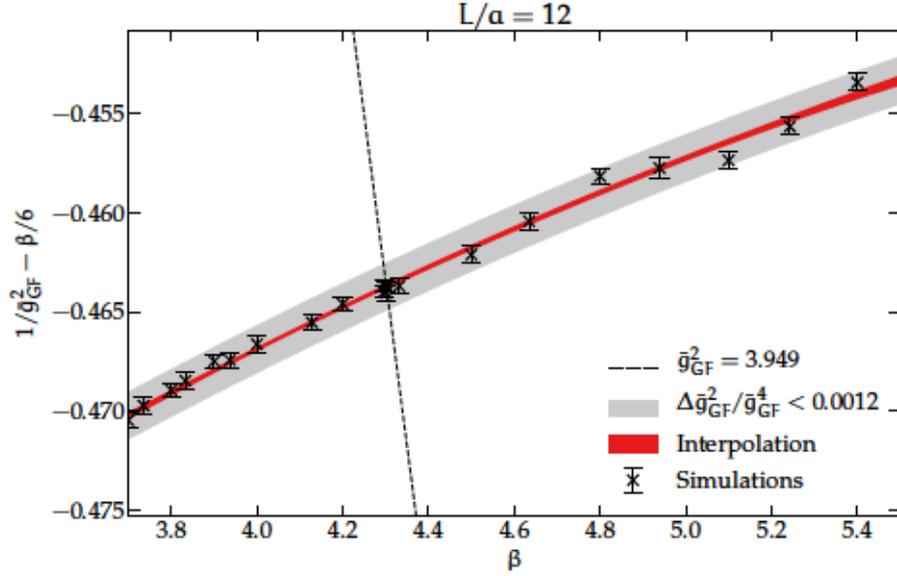


Figure 8.1: Interpolation of \bar{g}_{GF}^2 in β for $L_0/a = 12$ similar to fig. 11 of [45]. Since we performed a number of simulations around $\beta = 4.3$, single data points overlap in this region. We show the interpolating curve by the red band and our goal on the precision by the gray curve. The dashed line denotes g_*^2 .

and we can use the same approach to determine a value of β for the tentative goal from eq. (8.8). In our interpolations, we did not see a benefit in using a Padé ansatz versus a quadratic ansatz in a region localized around the target value. Since the error estimation from a linear fit is more reliable and the fit is more likely to converge, we use a quadratic interpolation. Using the available data, we arrive at $\beta \approx 4.9$ for our simulation at $L_0/a = 32$.

The value of g_*^2 is determined from the measured gradient flow coupling on the ensemble with $L_0/a = 32$ and $\beta = 4.9$. Based on approximately 5000 configurations which have been separated by 20 molecular dynamics units, we determine the target value

$$\bar{g}_{\text{GF}}^2(L_0/a = 32) = 3.949(11), \quad \rightarrow \quad g_*^2 = 3.949. \quad (8.10)$$

Based on this target, we generate the ensembles with smaller lattices.

For $L_0/a \in \{12, 16\}$ we employ the interpolation of $v(\beta)$ to determine a first guess for the bare coupling g_0^2 which is needed to achieve $\bar{g}_{\text{GF}}^2(L_0/a) = g_*^2$. The data points for the interpolation and the curve from a quadratic fit ansatz for $L_0/a = 12$ are shown in figure 8.1.

In both cases, we are not able to match g_*^2 to the required precision with the first ensemble. After we have generated enough statistics to safely determine the gradient flow coupling, we perform a new interpolation including the newly determined data point.

We tabulate all ensembles generated in the tuning procedure in table B.1. This table includes the bare parameters of the action, the

We will shortly comment on the required precision.

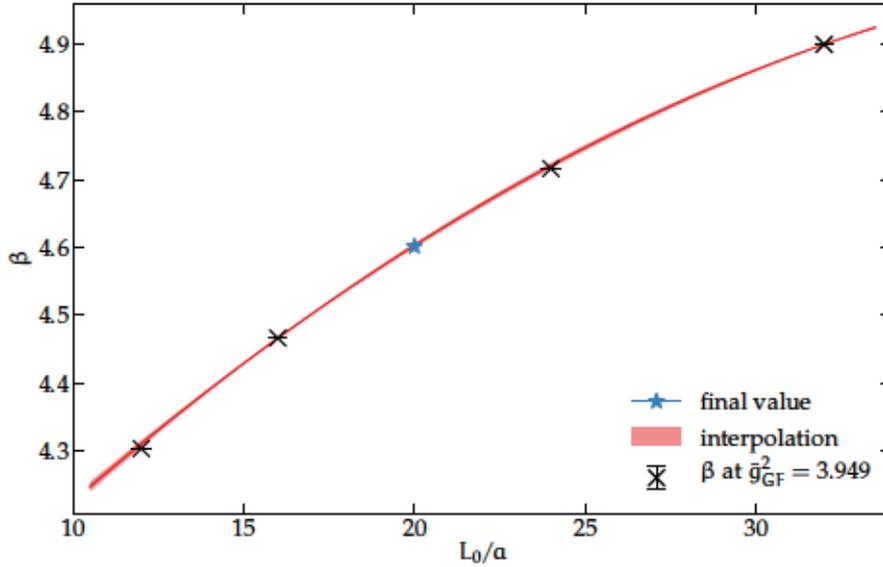


Figure 8.2: Interpolation of β for $\bar{g}_{\text{GF}}^2 = 3.949$ as determined from the interpolations at fixed L_0/a to $L_0/a = 20$. The blue star depicts the result for β at $L_0/a = 20$ as determined from the tuning procedure. It is not included interpolation.

number of replica, the separation of the ensembles in Monte Carlo time, the gradient flow coupling and the mass of the sea quarks. In some cases, the mass of the sea quark has not been determined, since it had been clear from the coupling that another simulation at the same lattice spacing had to be performed.

For the tuning at $L_0/a = 24$, we use the information provided in [45]. As it is the case for $L_0/a = 32$, the ensembles in the interpolation differ from the LCP by cut-off effects, since they have been generated as step scaling ensembles from ensembles with $L/a = 12$. Nevertheless, using the existing data together with our own tuning runs enables us to perform a reliable interpolation.

For $L_0/a = 20$, no data points for an interpolation have been available in the target region. Therefore, we follow a two-step procedure in this case. First, we perform the interpolations in $L_0/a \in \{12, 16, 24\}$ and determine the corresponding values of β at these lattice spacings. In the second step we combine the values of β with the pair $(32, 4.9)$ and perform a quadratic interpolation of β to $L_0/a = 20$. After the generation of two ensembles with $L_0/a = 20$ in the target region, the measured values of \bar{g}_{GF}^2 have been used to perform local interpolations. We show the interpolation in L_0/a together with the value for β obtained on the final ensemble in fig. 8.2.

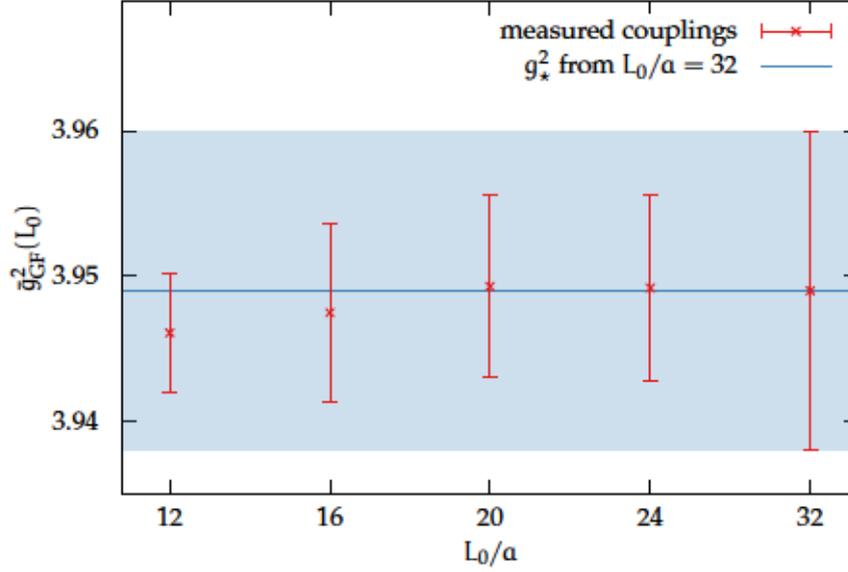


Figure 8.3: Gradient flow couplings at the five different lattice spacings in L_0 after the tuning procedure together with the target g_*^2 .

8.1.3 Target precision

As it can be seen from table B.1, we have generated four to six ensembles for every lattice spacing. The number of runs has been dictated by the target precision of the gradient flow coupling which we will derive in this section.

Based on the work in the two-flavor theory [215], we like to achieve a tuning that is more precise by about a factor of three. This amounts to a relative precision of

$$\frac{\Delta L_0}{L_0} = 0.01. \quad (8.11)$$

Since the tuning is done via the gradient flow coupling, we have to translate this target into a relative precision of the coupling. If we expand $\bar{g}^2(L + \Delta L)$ and use eq. (8.4) for the derivative of \bar{g} with respect to L , we can derive the relation

$$\frac{\Delta L}{L} = -\frac{\Delta \bar{g}^2}{2\bar{g}\beta(\bar{g})} = \left[\frac{-\bar{g}}{2\beta(\bar{g})} \right] \frac{\Delta \bar{g}^2}{\bar{g}^2} \quad (8.12)$$

which connects the relative error of the coupling with the relative error of the lattice extent. Using the non-perturbative determination of the β function from [45], we can determine the target precision of the coupling to be

$$\frac{\Delta \bar{g}^2}{\bar{g}^2} = 2 \frac{\bar{g}^2}{P(\bar{g}^2)} \frac{\Delta L}{L} \quad (8.13)$$

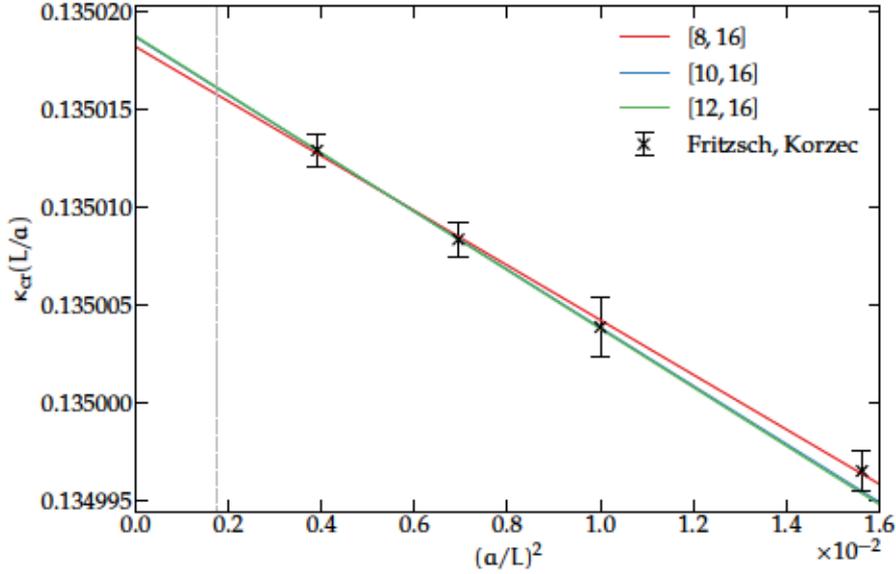


Figure 8.4: Extrapolation of the critical hopping parameters obtained from [216] for $\beta = 4.7165$ based on different fit intervals. The dashed line denotes $L/a = 24$.

and at the target coupling g_*^2 we arrive at

$$\frac{\Delta\bar{g}^2}{\bar{g}^2} = 0.0047, \quad \frac{\Delta\bar{g}^2}{\bar{g}^4} = 0.0012, \quad \Delta\bar{g}^2 = 0.0187. \quad (8.14)$$

From the values in table B.1 we can see that we have surpassed this goal on the relative precision for the simulation at $L_0/a = 32$, where we arrive at a relative error of $\Delta L/L \approx 0.6\%$. For all other lattice spacings we have reached an even higher precision. In figure 8.3, we show the couplings determined on all five ensembles. It is apparent that they have an excellent overlap with each other and g_*^2 .

8.1.4 Tuning towards the critical hopping parameter

The line of constant physics in eq. (8.1) consists of two requirements and to fulfill both, we have to tune the bare parameters of the Lagrange density. Whereas the tuning of β has been described in the former section, we will now describe how we have tuned the sea quark hopping parameter κ_1 such that the three sea quarks are massless. We note in passing that both tuning procedures have to be performed simultaneously since there is a dependence of the gradient flow coupling on the sea quark masses and the critical hopping parameter κ_{cr} depends on β .

The determination of the critical hopping parameter in dependence of g_0^2 and L/a has been performed in [216] for $L/a \in \{8, 10, 12, 16\}$ and the resulting interpolation formulae have been provided in [45]. The parametrization given by eq. (A.3) and table 7 of [45] guarantee the

We use the quark mass from the PCAC relation for the tuning towards κ_{cr} .

condition $Lm_1 < 0.005$ to hold. We will use this criterion as boundary for the tuning, i.e., we will consider quarks as massless if they fulfill $Lm_1 < 0.005$. This bound is more restrictive by an order of magnitude compared to the two-flavor simulations in [215].

Concerning the interpolation formulas of [45] we have to keep in mind one important point. In [45, 216] the quark mass has been defined as average over all configurations, regardless of the topological sector. In our work, we will consider the quark masses after the projection to the sector of vanishing topological charge, as defined in eq. (6.50). Since the topological charge is zero for all our configurations in L_0 , this detail is not important for the discussion in this section but it will become important for larger physical volumes.

For our simulations with $L/a \in \{12, 16\}$ we are able to employ the interpolation formula from [45] to determine $\kappa_1 = \kappa_{\text{cr}}$. For the other ensembles no information on κ_{cr} is available. To determine a first guess for κ_{cr} at $L/a \in \{20, 24, 32\}$, we use the information from the interpolation formulas for smaller values of L/a at the corresponding value of β . As shown in figure 8.4, we use these values of the critical hopping parameter for smaller values of L/a to perform a linear extrapolation in $(a/L)^2$. To estimate systematic effects, we consider three fits with four, three and two points and use the extrapolation based on two points as final result. In the exemplary determination of the critical hopping parameter for $(L_0/a = 24, \beta = 4.7165)$, the extrapolations using the two and three leftmost points give the same result.

In our tuning procedure, we have made the experience that the value determined from the extrapolation is determined well enough to achieve a sea quark mass close to the bound of $Lm_1 < 0.005$. In many cases, a shift of $O(2 \cdot 10^{-6})$ in the hopping parameter leads to a sea quark mass that is zero within errors. We have also tried extrapolations in $(a/L)^3$ which is the leading behavior in perturbation theory [62]. In many cases, this extrapolation leads to even larger quark masses.

In general, the extrapolation leads to a good first guess for the hopping parameter. Since the autocorrelation time of the sea quark mass is rather short, it is possible to monitor the quark mass at the beginning of the simulations and, if necessary, shift the hopping parameter such that the absolute value of the quark mass is reduced. If new ensembles had to be generated because the gradient flow coupling had not been matched g_*^2 to the required precision, we have used our experiences to adjust κ_1 from the extrapolation by small shifts. The determination of current quark masses on our ensembles is described in chapter 12.

We have been able to generate ensembles with vanishing quark masses for nearly all resolutions. The ensemble with $L_0/a = 32$ which has been used to set g_*^2 has a quark mass which is slightly larger than

It has to be ensured that the ensemble is thermalized, before a Monte Carlo average is taken.

our bound. Since the violation is small, we believe that this does not pose a problem for the matching procedure.

8.1.5 Step scaling to the matching volume

We have gathered the gradient flow couplings and sea quark masses of all ensembles which have been generated in the tuning procedure in L_0 , in table B.1. We have written the coupling β in a bold font for those ensembles where the LCP condition (8.1) is satisfied.

These five ensemble are the basis for the ensemble generation in the matching volume L_1 . We generate five ensembles with the same bare parameters of the Lagrangian, β and κ_1 , but double the number of lattice sites in every direction. These ensembles with the physical volume $L_1 = 2L_0$ are listed in table 8.1 together with the measured gradient flow couplings and quark masses.

Table 8.1: QCD ensembles for the matching with HQET in L_1 . N_{rep} is the number of replica runs, while N_{ms} gives the number of configurations, separated by τ_{ms} molecular dynamics units. The trajectory length is $\tau = 2$ throughout.

L_1/a	β	κ	N_{rep}	$\frac{\tau_{\text{ms}}}{\text{MD}}$	N_{ms}	\bar{g}_{GF}^2	$L_1 m_1$
24	4.3030	0.1359947	2	10	12000	5.592(11)	+0.00860(19)
32	4.4662	0.1355985	3	10	13500	5.689(14)	+0.00761(13)
40	4.6017	0.1352848	10	10	8825	5.755(22)	+0.00802(13)
48	4.7165	0.1350181	6	30	3190	5.770(26)	+0.00610(16)
64	4.90	0.1345991	4	40	2035	5.848(34)	+0.01486(15)

We have for the first time generated an ensemble with Schrödinger functional boundary conditions and 64 lattice points all four dimensions. The tuning of the algorithmic parameters to optimize the performance for this demanding ensemble is described in chapter 9. For every ensemble we have aimed at statistics of about 100 000 MDU to be able to map out the matching parameters at order $1/m_h$.

From experiences in the two-flavor theory [215] we know that the matching observables show a noticeable autocorrelation that increases towards the continuum limit. Therefore, we have increased the distance in Monte Carlo time between two configurations for the two finest ensembles. With this procedure we are able to reduce the memory consumption of our project and at the same time we do not lower the effective statistics.

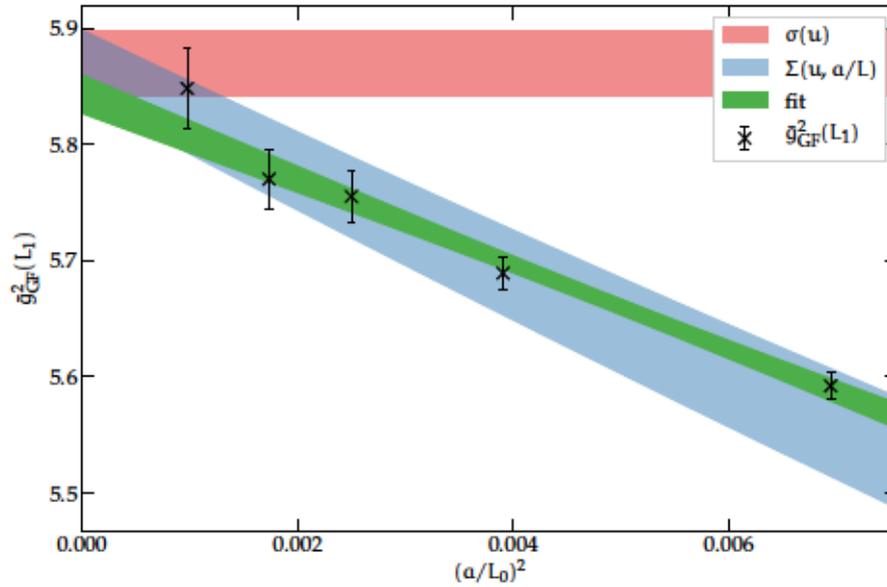


Figure 8.5: Gradient flow couplings $g_{GF}^2(L_1/a)$ in $(a/L_0)^2$ together with a linear extrapolation to the continuum limit. We show the result from $\sigma(u)$ and the parametrization of $\Sigma(u, a/L)$ from [45, 217] for comparison.

From table 8.1, we can see that the sea quark masses are significantly larger than on the ensembles in L_0 . This can be explained by cut-off effects. The difference

$$L_0[m_{11}(2L_0) - m_{11}(L_0)] \tag{8.15}$$

is expected to vanish with $O(a^2)$ in the continuum limit. The validation of this behavior fails and we rather see a scaling proportional to a/L . Since $O(a)$ improvement of PCAC masses is obtained from the tuning of c_A , we suspect that the value of c_A from the interpolation formula provided in [67] that has been used for the calculation of PCAC masses in this study might be incorrect in our coupling region. It has been extracted for couplings in the range $\beta \in [3.3, 3.81]$ and the interpolation formula has been constrained to the perturbative one-loop behavior for small couplings. In the region of the matching ensembles, the interpolation formula of [67] is not constrained by data. Therefore, we have to expect remnant $O(a)$ effects for observables that include c_A .

The cut-off effects in the gradient flow coupling can be compared with predictions of [45] where the step scaling function of the coupling $\sigma(u)$ has been determined. In this study, ensembles with $L/a \in \{8, 12, 16\}$ and the corresponding ensembles with the same bare parameters and $2L/a$ lattice points have been used to determine the discretized step scaling function $\Sigma(u, a/L)$. Afterwards, the function $\sigma(u)$ has been obtained from the continuum limit of $\Sigma(u, a/L)$ in a global fit in u and a/L . The continuum limit has been published in

A dedicated determination of c_A in the matching region might be useful.

We have introduced $\sigma(u)$ in eq. (8.7)

[45] and the parametrization of the cut-off effects is documented in [217].

We show our data together with the continuum step scaling function $\sigma(u)$ and the parametrization of $\Sigma(u, a/L)$ for $u = 3.949$ in figure 8.5. It can be seen that our data points fall exactly onto the parametrization of the cut-off effects. Since the autocorrelation time of the gradient flow coupling increases towards the continuum limit, the statistical precision is reduced. Nevertheless, our data provides an excellent check for the results of [45]. From these results, we compute

$$\sigma(3.949) = 5.867(29). \quad (8.16)$$

for the continuum limit of the discretized step scaling function. From a linear fit to the five data points obtained from the ensembles in L_0 and $2L_0$, we extract the continuum limit

$$\sigma(3.949) = 5.843(16) \quad (8.17)$$

and we show the corresponding curve in figure 8.5. If we exclude the rightmost data point, we obtain

$$\sigma(3.949) = 5.871(22). \quad (8.18)$$

Since the integrated autocorrelation time of the gradient flow coupling increases towards the continuum limit, our continuum extrapolation is not tightly constrained by the finest ensembles. The systematic effect due to the variations of the fit range is of the same size as the statistical error. Therefore, we cannot increase the relative precision of L from the use of our extrapolation.

However, our results provide a valuable test of the continuum extrapolation performed in [45] which was the basis for the determination of the strong coupling constant α_s in [214]. An improvement of the precision of $\sigma(u)$ could be reached, if our results were included in the global continuum extrapolation of [45].

The set of ensembles listed in table 8.1 will be used to determine the matching observables in the full theory, i.e., QCD with three flavors of massless quarks and one relativistic heavy quark with a mass around the bottom quark mass. In the next steps, we will specify the light line of constant physics for the effective theory.

8.1.6 The matching volume

Before we proceed to the generation of the HQET ensembles in the matching volume, we like to determine the physical size of the matching volume and its relative error. We perform the conversion from lattice units to physical units using the information from [142, 214]. There, a reference scale has been defined on the CLS ensembles [60] at

$$\mu_{\text{ref}}^* = 478(7) \text{ MeV} \quad (8.19)$$

We will determine the physical value of L_1 in the next section.

We have used this
hadronic scale in
eq. (8.8).

and this reference scale is connected to the hadronic scale via

$$\frac{\mu_{\text{ref}}^*}{\mu_{\text{had},1}} = 2.428(18). \quad (8.20)$$

We now turn around the arguments of section 8.1.1 and determine the scale factor $s = L_{\text{had}}/L_0$ based on the measured coupling g_*^2 . Using eq. (8.7), we can determine this scale factor from two couplings g_1 and g_2 via

$$\begin{aligned} \ln [s (g_1^2, g_2^2)] &= \int_{g_1}^{g_2} dx \frac{P(x^2)}{x^3} \\ &= \frac{P_0}{2g_1^2} - \frac{P_0}{2g_2^2} + p_1 \ln \left(\frac{g_2}{g_1} \right) + \frac{P_2}{2} (g_2^2 - g_1^2). \end{aligned} \quad (8.21)$$

With

$$g_2^2 = \bar{g}_{\text{GF}}^2(L_{\text{had}}), \quad g_1^2 = \sigma(g_*^2) \quad (8.22)$$

and

$$L_{\text{had}} = \frac{1}{\mu_{\text{had}}} = \frac{2.428}{\mu_{\text{ref}}^*} \quad (8.23)$$

we determine the physical extent of the tuning volume to

$$L_1 = 0.4998(35) \text{ fm}, \quad (8.24)$$

where we have used $\sigma(g_*^2)$ of eq. (8.16). We have thus been able to generate our ensembles such that the matching volume has an extent of 0.5 fm and we have fixed this physical extent with a relative error of $\Delta L_1/L_1 = 0.7\%$. If we take into account the error on $\bar{g}_{\text{GF}}^2(L_0/a = 32)$ in the determination of the physical size, we obtain a relative error of 0.9% which is still below our initial target of 1%.

Based on eq. (8.24), we are able to determine the range of lattice spacings on the QCD side of the matching. It is given by

$$a \in [0.0078 \text{ fm}, 0.021 \text{ fm}] \approx [(9.5 \text{ GeV})^{-1}, (25.3 \text{ GeV})^{-1}]. \quad (8.25)$$

Based on these values, we are confident that we are able to perform a controlled continuum limit of observables at the scale of the bottom quark mass.

8.2 THE LINE OF CONSTANT PHYSICS: HQET

The light part of the action remains unchanged, when we switch from QCD to HQET. The only difference, in terms of the ensemble generation, is that we will generate ensembles with much coarser lattice spacings. Since we will match the continuum limit of the QCD observables with the HQET observables at finite lattice spacing, the

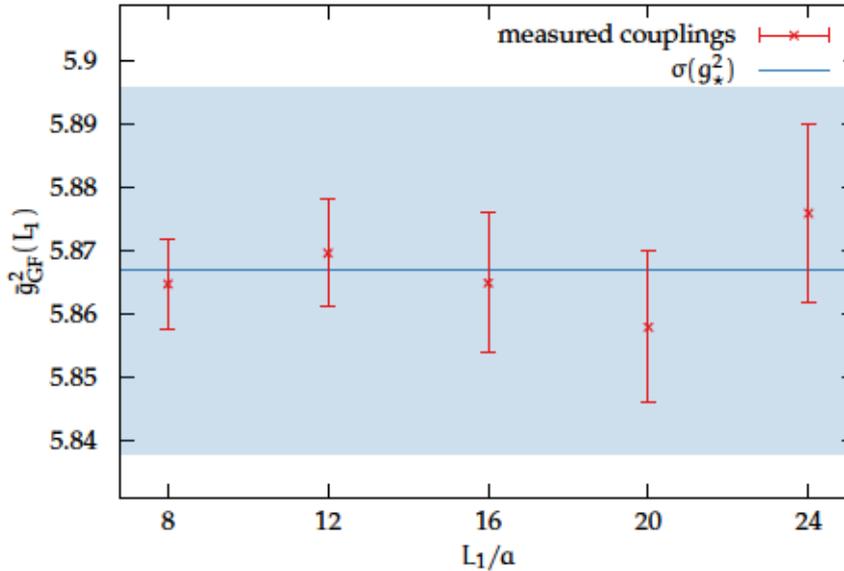


Figure 8.6: Gradient flow couplings at the five different lattice spacings in L_1 after the tuning procedure together with the target $\sigma(g_*^2)$.

line of constant physics for the HQET ensembles in the volume L_1 is defined by

$$\bar{g}_{\text{GF}}^2(L_1) = \sigma(g_*^2), \quad m_1 = 0. \quad (8.26)$$

We use the continuum step scaling function of [45] to define the LCP. This is done for practical reasons. The generation of the QCD ensembles in L_1 is computationally demanding and is a task on the time scale of many months. The continuum limit of the measured gradient flow couplings can only be obtained as soon as all ensembles have been generated. To speed up the ensemble generation in the HQET sector, we rely on the step scaling function of [45]. In hindsight, the agreement of eqs.(8.16) and (8.17) justifies this approach.

8.2.1 Tuning in the matching volume

On the HQET side of the matching procedure, we generate five ensembles with $L_1/a \in \{8, 12, 16, 20, 24\}$. From the determination of the physical lattice extent, we deduce that we perform our simulations in the range

$$a \in [0.021 \text{ fm}, 0.062 \text{ fm}] \approx [(3.2 \text{ GeV})^{-1}, (9.5 \text{ GeV})^{-1}], \quad (8.27)$$

which would be clearly inappropriate for relativistic bottom quarks. Since we consider the heavy quarks in the effective theory, we do not face significant discretization effects and we will be able to perform a reliable continuum extrapolation of the matching parameters.

The tuning to the line of constant physics of eq. (8.26) is performed as it was done for the QCD ensembles and we list all ensembles

On the large-volume CLS ensembles, the lattice spacings are in the range $a \in [0.039 \text{ fm}, 0.087 \text{ fm}]$

generated in the tuning procedure in table B.2. The ensembles where the coupling is written in a bold font are the ones to be used in the matching procedure. Due to the small volume, all configurations are in the sector of vanishing topological charge.

Based on investigations in the two-flavor theory [215] we know that heavy-light observables in the effective theory are weakly correlated in Monte Carlo time. In exchange, we face an exponential decrease of the signal-to-noise ratio with increasing source-sink separation. Therefore, we reduce the distance between two configurations to 4 MDU. Aiming at statistics of 60 000 MDU per ensemble, we will have to perform measurements on $O(15\,000)$ configurations per ensemble.

Since the length of the Monte Carlo runs is reduced compared to the QCD ensembles, we observe larger uncertainties in the gradient flow couplings. Nevertheless, we are able to tune all ensembles to a good precision and we show the measured couplings together with the goal from $\sigma(g_*^2)$ in figure 8.6.

8.2.2 Step scaling towards large volume simulations

Table 8.2: Step scaling ensembles in $2L_1$. Boundary conditions and action are chosen exactly as in [45] and we use $T = L$. N_{rep} is the number of replica runs, while N_{ms} gives the number of configurations, separated by τ_{ms} molecular dynamics units. $N_{\text{ms}}^{(0)}$ gives the number of configurations in the sector of vanishing topological charge. The trajectory length is $\tau = 2$ throughout. Values marked by a star have been obtained on short Monte Carlo histories, compared to their integrated autocorrelation time.

$2L_1/a$	β	κ	N_{rep}	$\frac{\tau_{\text{ms}}}{\text{MD}}$	N_{ms}	$N_{\text{ms}}^{(0)}$	\bar{g}_{GF}^2	$2L_1 m_1$
16	3.6537	0.1370722	4	8	15000	12213	9.377(27)	0.0446(6)
24	3.8349	0.1369654	2	16	15000	12423	10.485(39)	0.0192(3)
32	4.0018	0.1366803	3	4	15000	13663	10.810(93)	0.0144(3)
40	4.1394	0.1363857	6	4	1457	1455	10.69(23)*	0.0124(6)
48	4.2530	0.1361224	4	4	5247	5244	10.84(16)*	0.0130(3)

To reach a scale that is applicable for large-volume simulations, we have to perform another step scaling procedure. This time, we scale from L_1 to $2L_1$. We list the five ensembles generated in this volume in table 8.2. As apparent from the table, we have not yet reached our goal for the statistics of the two ensembles with $2L_1/a = 40, 48$. Since $\tau_{\text{int}}(\bar{g}_{\text{GF}}^2)$ is of $O(150\text{ MDU})$, the determination of the gradient flow coupling might be biased due to the comparably short Monte Carlo histories.

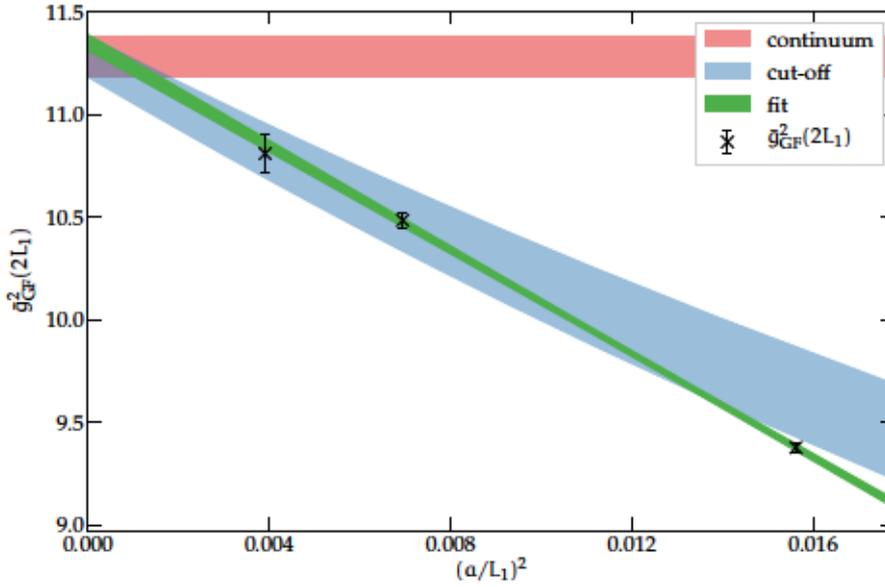


Figure 8.7: Gradient flow couplings $\bar{g}_{\text{GF}}^2(2L_1/a)$ in $(a/L_1)^2$ together with a linear extrapolation to the continuum limit. We show the result for $\sigma(\sigma(u))$ and the parametrization of $\Sigma(u, a/L)$ from [45, 217] for comparison.

In larger volumes the topological charge is not frozen to the sector $Q = 0$ anymore. At the same time, we do not observe a freely fluctuating charge but longer periods where the charge stays within single sectors. We show the number of configurations in the sector of vanishing topological charge in table 8.2.

In the two-flavor project [215], the contact to large volume simulations has been obtained using the continuum limit of the observables obtained from the ensembles in the volume $2L_1$. In our work, the overlap of the lattice spacings in this step with the ones of the large-volume CLS simulations is not large enough to allow for a safe interpolation of the matching parameters. Therefore, we have to simulate another set of ensembles with the same physical lattice extent but different resolutions.

To tune this last set of lattices to the correct physical volume, we look a second time at the step scaling function of [45] and compare their results with the results obtained from the ensembles in tab. 8.2. From two steps of the continuum step scaling function we obtain

$$\sigma(\sigma(g_*^2)) = 11.27(10) \quad (8.28)$$

which, for the same reasons as before, will be taken as goal for the tuning procedure in the volume $L_2 = 2L_1$. For the reasons outlined above, we take only the three coarsest ensembles into account for the continuum extrapolation of our data. From a linear fit we obtain

$$\sigma(\bar{g}_{\text{GF}}^2(L_1)) = 11.35(4). \quad (8.29)$$

Or, equivalently, one step of stepsize $s = 4$

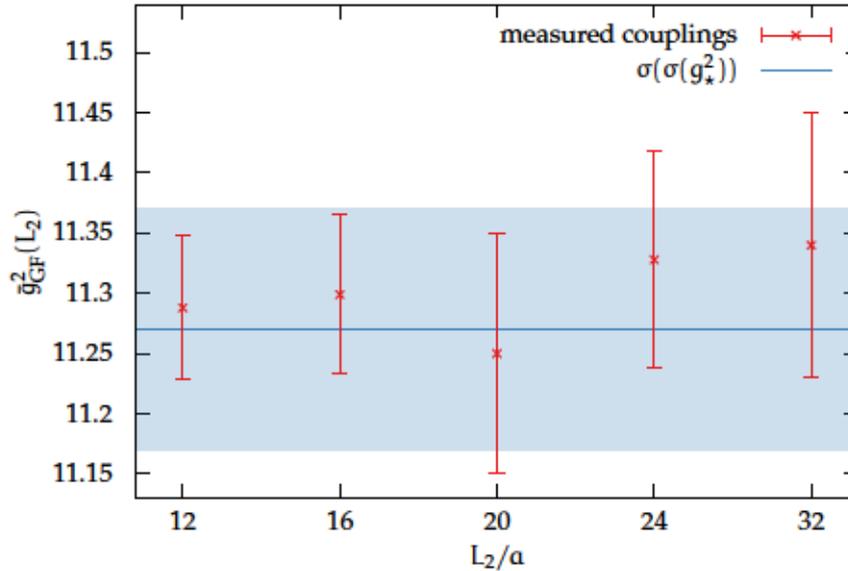


Figure 8.8: Gradient flow couplings at the five different lattice spacings in L_2 after the tuning procedure together with the target $\sigma(\sigma(g_*^2))$. The measured values are listed in tab. B.3 and the value for $L/a = 32$ is taken from [214].

This result is compatible with the one from [45] and due to the large amount of statistics, the extrapolated result is very precise. The inclusion of the two data points at finer lattice spacing will help to reduce the systematic uncertainty due to the coarse lattice spacings in the current extrapolation.

8.2.3 Contact to the large-volume simulations

To connect the finite-volume matching parameters with the measurements in the effective theory at large volume, we have to be able to interpolate the parameters in the range of inverse couplings β given by the CLS simulations. Therefore, we generate another set of ensembles with $L_2/a \in \{12, 16, 20, 24, 32\}$ which translates into

$$a \in [0.031 \text{ fm}, 0.083 \text{ fm}] \approx [(2.4 \text{ GeV})^{-1}, (6.4 \text{ GeV})^{-1}], \quad (8.30)$$

and gives us an excellent overlap with the CLS ensembles.

We perform a third tuning procedure to achieve the generation of all ensembles along the chosen line of constant physics. Since the topological charge is fluctuating strongly on the coarser ensembles, a significant fraction of the configuration is in a sector where the topological charge is non-zero. Therefore, the interpolation formulae of [45] for κ_σ become less accurate and we have to devote an increased effort on the tuning of the hopping parameter. We list all ensembles generated at this step in table B.3.

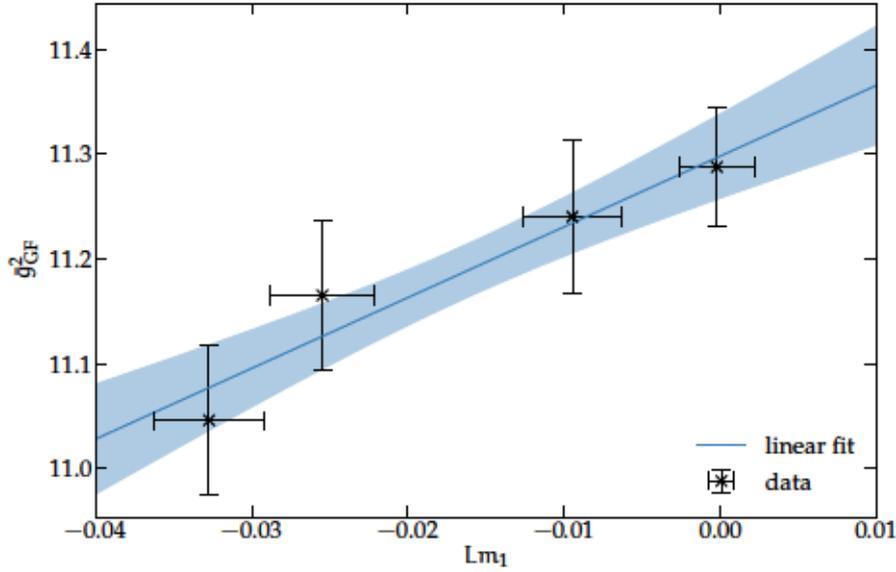


Figure 8.9: Gradient flow coupling in dependence of the sea quark mass at $L_2/a = 12$ and $\beta = 3.4014$.

The generation of several ensembles with the same bare coupling β and different hopping parameters κ_1 allows us to study the correlation of the gradient flow coupling with the sea quark mass. For $L_2/a = 12$ and $\beta = 3.4014$ we have generated four ensembles. In fig. 8.9 we show the gradient flow couplings obtained on these ensembles depending on the sea quark mass. Since all ensembles have been generated during the tuning procedure, the spread of the sea quark masses is rather small. Nevertheless, it is possible to perform a linear fit to obtain

$$\bar{g}_{GF}^2(Lm_1 = 0) = 11.298(40) \quad \frac{d\bar{g}_{GF}^2}{dLm_1} = 0.56(16) \quad (8.31)$$

for the chiral limit of the gradient flow coupling and its derivative with respect to the sea quark mass at this point in the parameter space. The shift of the gradient flow coupling due to a residual quark mass $Lm_1 < 0.005$ is therefore completely negligible.

For our measurements at $L_2/a = 32$, we do not have to generate a new ensemble since we are able to use an already existing one that has been generated in [214]. This ensemble matches the line of constant physics condition to a good precision and we are able to save the cost of a tuning at the largest value of L_2/a . We illustrate the result of this tuning in figure 8.8, where we show the target coupling and the measured couplings obtained on the five ensembles.

Since we do not have to perform a continuum limit at this stage of the matching procedure, we estimate that 30 000 MDU are enough to obtain precise results.

A similar analysis in the region $\bar{g}_{GF}^2 \approx 4$ revealed a negligible dependence on the mass [217].

8.3 THE RUNNING OF THE MASS

So far, we have described how we generated our ensembles such that all of them are on the line of constant physics, defined by eq. (8.1). What remains to be done is to fix the renormalized heavy quark mass on the QCD side, to allow for a parametrization of the matching coefficients in the heavy quark mass. As pointed out in section 7.3.3, we like to fix the renormalization group invariant (RGI) quark mass and we have not yet specified, how the running from the scale of finite-volume simulations to RGI values is performed.

8.3.1 The step scaling function of the mass

Results for the three-flavor theory have been obtained in [64] based on the methods of [48, 218] and these references will be the basis of our discussion in this section. We consider the scale evolution of quark masses, as we have done it for the coupling, and use the renormalization group equation

$$\mu \frac{d}{d\mu} m_{R,i} = \tau(\bar{g}(\mu)) m_{R,i}, \quad i = 1 \dots, N_f, \quad \mu = \frac{1}{L} \quad (8.32)$$

for the N_f quark flavors in our theory. The renormalization group function τ has the expansion

$$\tau(g) \underset{\bar{g} \rightarrow 0}{\sim} -d_0 \bar{g}^3 - d_1 \bar{g}^5 - d_2 \bar{g}^7 + \dots \quad (8.33)$$

with the universal coefficient

$$d_0 = \frac{8}{(4\pi)^2}. \quad (8.34)$$

The integration of eq. (8.33), including the explicit subtraction of the leading term, leads to the RGI mass

$$M_i = m_{R,i} [2b_0 \bar{g}^2(\mu)]^{-\frac{d_0}{2b_0}} \exp \left\{ - \int_0^{\bar{g}(\mu)} dg \left[\frac{\tau(g)}{\beta(g)} - \frac{d_0}{b_0 g} \right] \right\}, \quad (8.35)$$

which can be interpreted as integration constant of the renormalization group equation (8.32). Since τ is independent of the quark flavor i , the same holds for the ratio $M_i/m_{R,i}(\mu)$ and one can show that ratios $m_{R,i}(\mu)/m_{R,j}(\mu)$ are scale-independent. The scheme independent value M_i is non-perturbatively defined if τ is evaluated non-perturbatively.

The step scaling function of the mass σ_P is defined from the step scaling function of the coupling in eq. (8.3) via

$$\sigma_P(u) = \exp \left\{ - \int_{\sqrt{u}}^{\sqrt{\sigma(u)}} dg \frac{\tau(g)}{\beta(g)} \right\}, \quad (8.36)$$

The derivation is similar to the definition of the β function in eq. (8.4)

In contrast to the coefficient b_1 of the running coupling, the coefficient d_1 depends on the renormalization scheme.

and since we defined $\sigma(u)$ by two scales that differ by a factor of two, we can write

$$\sigma_P(u) = \frac{m_{R,i}(\mu)}{m_{R,i}(\mu/2)} \Big|_{u=\bar{g}^2(\mu)}. \quad (8.37)$$

On the lattice, the continuum step scaling function σ_P is obtained from the continuum limit of the discretized function Σ_P ,

$$\sigma_P(u) = \lim_{a \rightarrow 0} \Sigma_P(g_0^2, a\mu) \Big|_{u=\bar{g}^2(\mu)}, \quad (8.38)$$

where Σ_P is defined from the renormalization constant of the pseudoscalar density Z_P via

$$\Sigma_P(g_0^2, a\mu) \equiv \frac{Z_P(g_0^2, a\mu/2)}{Z_P(g_0^2, a\mu)}. \quad (8.39)$$

8.3.2 Determination of RGI quark masses

The work of [64] allows us to determine RGI quark masses from renormalized quark masses computed at the matching scales $\mu_{\text{mat},0} = 1/L_0$ and $\mu_{\text{mat},1} = 1/L_1$. The scaling from the low to the high-energy regime is split into two parts according to

$$M_i = \frac{M_i}{m_{R,i}(\mu_0/2)} \frac{m_{R,i}(\mu_0/2)}{m_{R,i}(\mu_{\text{mat}})} m_{R,i}(\mu_{\text{mat}}), \quad (8.40)$$

where the scale μ_0 has been introduced in [64]. The first term is given by

$$\frac{M_i}{m_{R,i}(\mu_0/2)} = 1.7505(89). \quad (8.41)$$

The running in the low energy region is parametrized for a range of couplings via

$$\tau(\bar{g}) = -\bar{g}^2 \frac{\sum_{n=0}^3 f_n \bar{g}^{2n}}{\sum_{k=0}^2 p_k \bar{g}^{2k}} \quad (8.42)$$

where the parameters p_k are the same as in eq. (8.21) and the parameters f_n together with the covariances $\text{cov}(f_i, f_j)$ and $\text{cov}(f_i, p_j)$ are provided in eqs.(4.6–4.8) of [64]. For the running to the volume L_0 , we determine based on these information

$$\frac{m_{R,i}(\mu_0/2)}{m_{R,i}(\mu_{\text{mat},0})} = 0.8423(26), \quad (8.43)$$

and derive the total running factor from renormalized quark masses in the tuning volume L_0 to RGI quark masses to be

$$h(L_0) \equiv \frac{M_i}{m_{R,i}(\mu_{\text{mat},0})} = 1.4744(87). \quad (8.44)$$

The relative uncertainty of this factor is about 0.6% and we will see that this is by far the dominant contribution to the total error of RGI quark masses determined in the matching volume.

8.3.3 *Determination of the step scaling function*

With the ensembles described in this section, we are able to determine $\sigma_P(\mathbf{u})$ for two values of \mathbf{u} , just as we have done it for $\sigma(\mathbf{u})$. In principle, the data obtained from these determinations can be used together with the data of [64] to improve the precision of $h(L_0)$.

To determine $\sigma_P(3.949)$ and $\sigma_P(5.867)$, we have to determine Z_P on the ensembles in the volume L_0 and $2L_0$ on the QCD side and L_1 and $2L_1$ on the HQET side of the matching. We will explain how to determine Z_P in section 12.4.1, where the renormalization constant is calculated for the determination of renormalized quark masses.

In [64], ensembles with $L/a = 8, 12, 16$ and the corresponding $2L$ ensembles have been generated to determine Σ_P for seven values of the gradient flow coupling. The results for Z_P and Σ_P per ensemble are given in table 8 of that reference.

We list the values of Z_P and the corresponding result for Σ_P in table 8.3 and show the continuum extrapolation of the results together with the results from [64] in fig. 8.10. We are able to cover a region much closer to the continuum limit than it has been done in [64]. At the same time, we see significant autocorrelation for Z_P and $\tau_{\text{int}}(Z_P)$ increases with decreasing lattice spacing. Therefore, the effective statistics decreases towards the continuum limit and we are not able to tightly constrain the continuum extrapolation with our data points.

Table 8.3: Results for Z_P and Σ_P at the couplings g_*^2 and $\sigma(g_*^2)$.

\bar{g}_{GF}^2	L/a	β	κ	$Z_P(g_0^2, L/a)$	$Z_P(g_0^2, 2L/a)$	$\Sigma_P(g_0^2, L/a)$
3.9461(41)	12	4.3030	0.1359947	0.57835(32)	0.48080(49)	0.8313(10)
3.9475(61)	16	4.4662	0.1355985	0.56972(45)	0.47313(79)	0.8305(16)
3.9493(63)	20	4.6017	0.1352848	0.56503(54)	0.46634(105)	0.8253(21)
3.9492(64)	24	4.7165	0.1350181	0.56003(48)	0.46311(126)	0.8269(24)
3.9490(110)	32	4.9000	0.1345991	0.55390(70)	0.45724(199)	0.8255(38)
5.8648(70)	8	3.6537	0.1370722	0.49386(32)	0.34707(111)	0.7028(23)
5.8697(85)	12	3.8349	0.1369654	0.47643(44)	0.31675(181)	0.6648(39)
5.8650(110)	16	4.0018	0.1366803	0.47349(49)	0.30742(387)	0.6493(82)
5.8760(140)	24	4.2530	0.1361224	0.46807(82)	0.30610(921)	0.6540(197)

The continuum limit of $\sigma_P(g_*^2)$ shows an excellent agreement with the data from [64]. For the three ensembles with the finest lattice spacings we cannot see any cut-off effects within the rather large statistical errors. The overlap of the extrapolated results at the coupling $\sigma(g_*^2)$ is not as good but still both determinations agree within error.

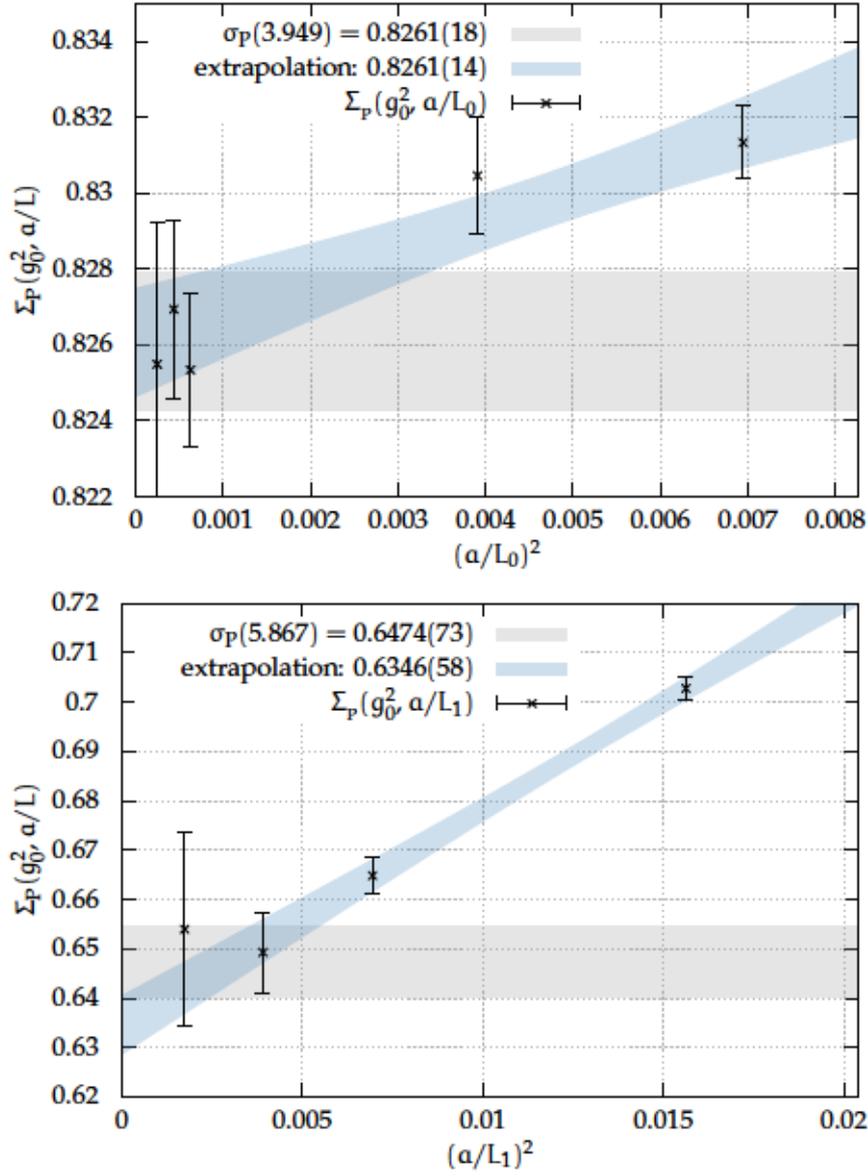


Figure 8.10: Continuum limit of the step scaling functions Σ_P at the couplings g_*^2 and $\sigma(g_*^2)$. The blue uncertainty band shows the fit to our data points whereas the gray area shows the values from [64].

However, the result for $2L_1/a = 48$ shows significant statistical errors and we have not included the result for $2L_1/a = 40$ since the Monte Carlo chain is too short, compared to the autocorrelation time of Z_P . An increase in statistics for the finest lattice spacings will improve the quality of our extrapolation. Even without the data from the last missing ensemble, it is safe to conclude that our continuum extrapolation of Σ_P provides a successful test of the results of [64].

8.4 CONCLUSIONS

After an extensive tuning of the physical and algorithmic parameters we have simulated all ensembles that are needed to perform the measurements for the non-perturbative finite-volume matching of QCD and HQET and to scale the matching parameters to the coupling region of typical large-volume simulations.

Some time will be needed for the last two ensembles.

To a good precision, we have performed our simulations with massless sea quarks and fixed the volume of our finite-volume QCD simulations to a relative precision of 0.7% which is better by more than a factor of three compared to the computations in the two-flavor theory. We have performed our simulations at five different resolutions to achieve reliable continuum extrapolations. To allow for the propagation of relativistic bottom quarks in boxes of spatial extent of 0.5 fm we have generated computationally very demanding ensembles with 48^4 and 64^4 lattice sites.

Together with the findings of chapter 12, where the heavy part of the line of constant physics is fixed, we are now in the position to perform the measurements for the determination of the matching parameters at all stages of the strategy outlined in figure 7.1. As soon as these parameters are determined, we are able to compute the bottom quark mass and the B and B_s meson decay constants at next-to-leading order of HQET based on the measurements described in chapter 16. There, all correlation functions that are necessary for the computation at next-to-leading order have been determined on eight ensembles at three different lattice spacings.

Our results for the step scaling functions $\sigma(u)$ and $\sigma_P(u)$ can be incorporated in the global fits performed in [45] and [64] respectively to increase the precision of the interpolation formulas. High precision of σ in the coupling region of our matching ensembles is needed in the project on renormalization by decoupling [219]. An increased precision of σ_P reduces the uncertainty of the running factor $M/m_{R,\text{had}}$ that is needed to evolve renormalized quark masses from the scale of large-volume simulations to their RGI values. The relative error on this running factor is the dominant contribution to the total error of our result for the RGI charm quark mass determined in chapter 14.

In the last chapter we have described the ensembles that we have simulated for the matching procedure. In total we have generated $O(60)$ different ensembles with Schrödinger functional boundary conditions in three different physical volumes with 8^4 to 64^4 lattice sites. To optimize the performance and to reduce the total cost of the simulations, we have performed an extensive tuning of the algorithmic parameters. The technical difficulties and strategies to circumvent them differed, depending on the region in the physical parameter space.

Some technical aspects of the simulation of three-flavor QCD in finite volume such as the scaling of the spectral range of the Dirac operator and the scaling of autocorrelation times have already been discussed in appendix A of [45]. We have used the information provided in this study to set up our Monte Carlo runs and to obtain initial parameters for the algorithmic setup. In contrast to [45], we have been able to use the newest solvers of the openQCD package which feature preconditioning via the Schwarz alternating procedure and inexact deflation. Therefore, we will use this section to investigate the performance of these solvers in finite volume. Additional information concerning the setup of large-volume simulations with openQCD is available in table 2 of [60] and table 9 of [220].

9.1 SOLVERS

The use of the modern solvers described in section 5.3 enabled us to simulate and measure on large lattices. The complexity of solver algorithms and with it the number of parameters that have to be tuned increase with the performance gain. For the optimization of the deflated solver of the openQCD package, $O(15)$ different parameters have to be tuned.

We use the solvers in two different applications, namely the inversion of the Dirac matrix in the computation of quark propagators and the inversion in the (R)HMC algorithm. In the latter case, the complexity of the tuning procedure is intertwined with the tuning of the algorithmic parameters of the Monte Carlo algorithm. In this case, the stability of the solver is of paramount importance. We do not adjust algorithmic parameters within one Monte Carlo run, since changes in the algorithmic parameters lead to different autocorrelations in the data.

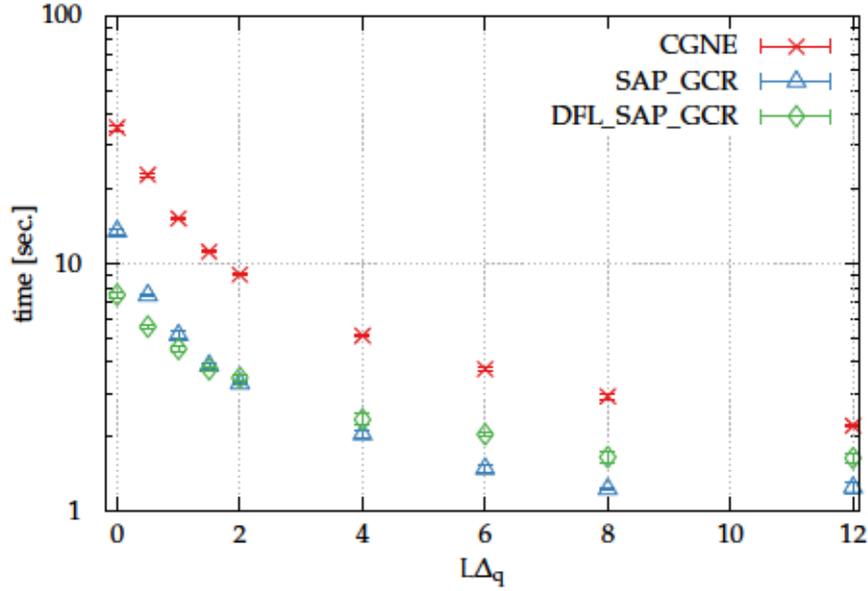


Figure 9.1: Comparison of three different solvers for a range of valence quark masses on a lattice with $L_1/a = 24$.

9.1.1 Quark mass dependence

We start our investigation of different solvers by comparing their performance in the determination of quark propagators. We look at the solvers that have been introduced in section 5.3. On a lattice with $L_1/a = 24$ we determine quark propagators in a range of valence quark masses

$$L\Delta_q \equiv \frac{L}{2} \left(\frac{1}{\kappa} - \frac{1}{\kappa_1} \right) \in [0, 12], \quad (9.1)$$

where κ_1 is the light quark hopping parameter which is close to the critical one. Since we perform our calculation in the Schrödinger functional, we are able to evaluate the propagators at vanishing quark masses. A similar analysis for physical quark masses in large volume has been done in [80].

For the three different solvers, the CGNE solver, the SAP_GCR solver and the DFL_SAP_GCR solver, the implementations of the openQCD package have been used. In figure 9.1 we show the timing for the determination of the forward and the backward propagator, since this is the use case scenario. We also list the timings and the inversion counts for the different solvers in table 9.1.

All three solvers are quark mass dependent such that the inversion is more expensive for smaller quark masses. We see that the CGNE solver performs worse than the two preconditioned solvers and in the chiral limit, the deflated solver is better by a factor of five than the CGNE. As we would expect it, deflation leads to a less steep increase towards

This ensemble will be used for the matching.

We use local lattices with $24 \times 12 \times 4 \times 4$ and a residuum of 10^{-11} . This amounts to 12 inversions of the Dirac equation.

the chiral limit, since the lowest eigenvalues of the Dirac matrix have been projected away and accordingly the condition number is smaller.

For quark masses above $L\Delta_q = 2$, which is above the charm quark mass, the SAP preconditioned solver performs better than the deflated solver. In table 9.1, we can see that the iteration count is constant for both solvers. Therefore we can conclude that the overhead of the deflated solver becomes dominant at large quark masses, where the solve itself is fast.

In the case under investigation, the generation of the deflation subspace takes 5.1(2) seconds. If the two propagators are determined for a single massless quark, the overhead from the generation of the deflation subspace is already smaller than the difference between the DFL_SAP_GCR and the SAP_GCR solver. If a larger number of inversions in the light quark sector is performed, the time that is needed for the generation of the deflation subspace becomes quickly irrelevant.

In table 9.2 we list the results of the same analysis on the ensemble with $L_1/a = 64$. We do not compare the absolute timings between both tables since the number of CPU cores has not been scaled with the same factor as the number of lattice points. Therefore we focus on the comparison of the solvers within one set of configurations. The general statements of our investigation for $L_1/a = 24$ hold for $L_1/a = 64$ as well.

The gain from the use of the deflation subspace is much larger for $L_1/a = 64$. Whereas the ratio between the timings for the CGNE solver and the SAP_GCR solver stays roughly constant when switching from $L_1/a = 24$ to $L_1/a = 64$, the ratio between the timings for the CGNE solver and the DFL_SAP_GCR solver increases from five to 15 and the ratio between the timings for the SAP_GCR solver and the DFL_SAP_GCR solver scales accordingly.

Since the smallest eigenvalue of the Dirac operator in the Schrödinger functional scales with $1/T^2$ [39], the deflation of the Dirac matrix becomes more important for larger lattices. For both lattices, the generation of the deflation subspace takes about the same time as the calculation of the propagator for a light quark.

9.1.2 Tuning of the deflated solver

In the last section we have not specified the parameters of the three different solvers. Since we use the deflated solver whenever it is possible, we will investigate some of the parameters that can be tuned to achieve an optimal solver performance. During our simulations, the starting point for the chosen parameters has been set by experiences made in the simulation of the Schrödinger functional [45] and in large volume [60]. Starting from an initial set of parameters, we have gradually adjusted some of the parameters during the thermalization process.

The iteration counts in the case of the block preconditioned solvers are the average number of Krylov vectors, that are required for the solution.

We use local lattices with 64×8^3 and a residuum of 10^{-11} .

Table 9.1: Comparison of the averaged inversion times and iteration counts between CGNE, DFL_SAP_GCR and SAP_GCR solver for $L_1/\alpha = 24$. The generation of the deflation subspace takes 5.1(2) s.

$L\Delta_q$	time [s]			iterations		
	CGNE	DFL_SAP_GCR	SAP_GCR	CGNE	DFL_SAP_GCR	SAP_GCR
0	34.5(10)	7.04(24)	13.27(37)	924	20	50
0.5	22.2(5)	5.25(22)	7.37(22)	595	15	28
1	14.8(3)	4.26(15)	5.02(6)	394	12	19
1.5	10.9(2)	3.53(12)	3.81(10)	289	10	14
2	8.80(6)	3.23(7)	3.25(8)	230	9	12
4	5.01(10)	2.18(3)	2.03(12)	123	6	7
6	3.67(8)	1.90(3)	1.46(2)	86	5	5
8	2.85(7)	1.52(3)	1.21(1)	66	4	4
10	2.44(2)	1.53(8)	1.22(2)	55	4	4
12	2.17(2)	1.52(7)	1.22(2)	48	4	4

Table 9.2: Comparison of the averaged inversion times and iteration counts between CGNE, DFL_SAP_GCR and SAP_GCR solver for $L_1/\alpha = 64$. The generation of the deflation subspace takes 62.5(9) s.

$L\Delta_q$	time [s]			iterations		
	CGNE	DFL_SAP_GCR	SAP_GCR	CGNE	DFL_SAP_GCR	SAP_GCR
0	964.3(31)	65.5(5)	328.8(88)	2493	19	137
0.5	223.7(5)	41.0(1)	62.8(2)	575	13	28
1	120.4(3)	31.3(1)	35.4(2)	303	10	16
1.5	83.6(2)	25.1(1)	24.7(1)	209	8	11
2	64.7(1)	22.7(1)	20.4(1)	158	7	9
4	36.9(1)	16.7(1)	11.9(1)	84	5	5
6	26.4(1)	13.4(1)	10.1(1)	59	4	4
8	21.6(2)	13.2(1)	10.2(1)	46	4	4
10	19.5(2)	10.5(1)	7.9(1)	39	3	3
12	19.0(1)	10.6(1)	7.8(1)	34	3	3

For larger lattices, we have performed explicit scans for a range of values for different parameters. In this section, we investigate the performance of the determination of light and heavy propagators on four uncorrelated configurations of the ensemble with $L_1/a = 24$ for $O(100)$ different combinations of parameters of the deflated solver.

In the following, we will present the results of our parameter scans for parameters where a clear dependence of the execution time on the value of the parameter is visible. In some cases, fluctuations in the execution time due to technical reasons seem to have a larger impact than the choice of the parameter and we are not able to draw a conclusion.

When varying one parameter, we keep all others constant. We show the basis for the variation, i.e., the parameters without any variation, in listing 9.1. The full documentation of all parameters can be found in [221] and we give an introduction in section 5.3.5. In the shown part of the input file, the parameters are split into sections for the solver, the SAP block setup, the setup of the deflation subspace, the parameters for the subspace generation and the parameters for the deflation projection.

Listing 9.1: Entries of the openQCD input file for the deflated solver. The choices shown here are the basis for all variations in this section.

```
[Solver 0]
solver      DFL_SAP_GCR
nkV         24
isolv       1
nmr         4
ncy         5
nmX         2048
res         1.00e-11

[SAP]
bs          4 4 4 4

[Deflation subspace]
bs          4 4 4 4
Ns          20

[Deflation subspace generation]
kappa       0.1359947
mu          0.0
ninv        10
nmr         4
ncy         5

[Deflation projection]
nkV         18
nmX         2048
res         0.01
```

We have performed our analysis for a range of valence quark masses and in most cases we do not see any qualitative difference in the determination of the optimal parameters for different valence quark masses. Since the inversion of the massless Dirac operator is the most expensive operation and since this inversion is the one needed in the HMC algorithm, we focus on the massless case in this study.

We start our investigation with the solver parameters, which are part of the setup of the deflated DFL_SAP_GCR solver and the SAP_GCR solver. We show scans for two parameters in figure 9.2. n_{mr} and n_{cy} denote the numbers of SAP block solver iterations and SAP cycles. We can see a clear dependence of the solver performance on these two parameters. The overhead from the block preconditioning seems to increase quickly. Only a few applications seem to be enough to achieve a fast solve.

We continue our investigations with the parameters that define how the deflation subspace is generated and show the results of the scans for three parameters in figure 9.3. Some of the parameters, as the number of sources N_s that is shown in the first plot, have an influence on the effort that is needed to generate the deflation subspace. We see, that the quality of the deflation subspace is improved when the number of sources is increased. At some point, the time that is needed for the solution reaches a plateau. At the same time, the cost for the generation of the deflation subspace scales linear with N_s . For the case under investigation, the time for the deflation subspace rises from 4s at $N_s = 10$ to 17s at $N_s = 30$. Therefore, it is only beneficial to use a large number of sources if the Dirac equation has to be solved a number of times.

An increase in the number of inverse iteration steps n_{inv} for the construction of low-precision solutions in the generation of the deflation subspace also results in an increased effort for the generation of the deflation subspace. On the other hand, we see that the quality of the deflation subspace is improved up to $n_{inv} = 10$. A further increase of n_{inv} does not decrease the inversion time, but increases the cost for the generation of the deflation subspace.

The parameters μ and κ can be used to adjust the Dirac operator in the generation of the deflation subspace to the physical parameters of the simulation. In Monte Carlo runs, a small value of μ can be used to add a twist to the Dirac operator to stabilize the subspace generation. In figure 9.3 we can see that increasing the twist beyond some small value leads to a gradually decreasing performance of the deflation subspace. Apparently, the low modes of the Dirac operator cannot be efficiently projected out any more. Interestingly, the quality of the deflation subspace does not show a clear dependence on the value of κ . In our tests, we scanned a range of

$$\frac{L}{2} \left(\frac{1}{\kappa^{diff}} - \frac{1}{\kappa_1} \right) \in [0, 1], \quad (9.2)$$

The argument is the same as in the case of twisted mass reweighting in CLS simulations. The trend shown in fig. 9.3 continues for larger values of μ .

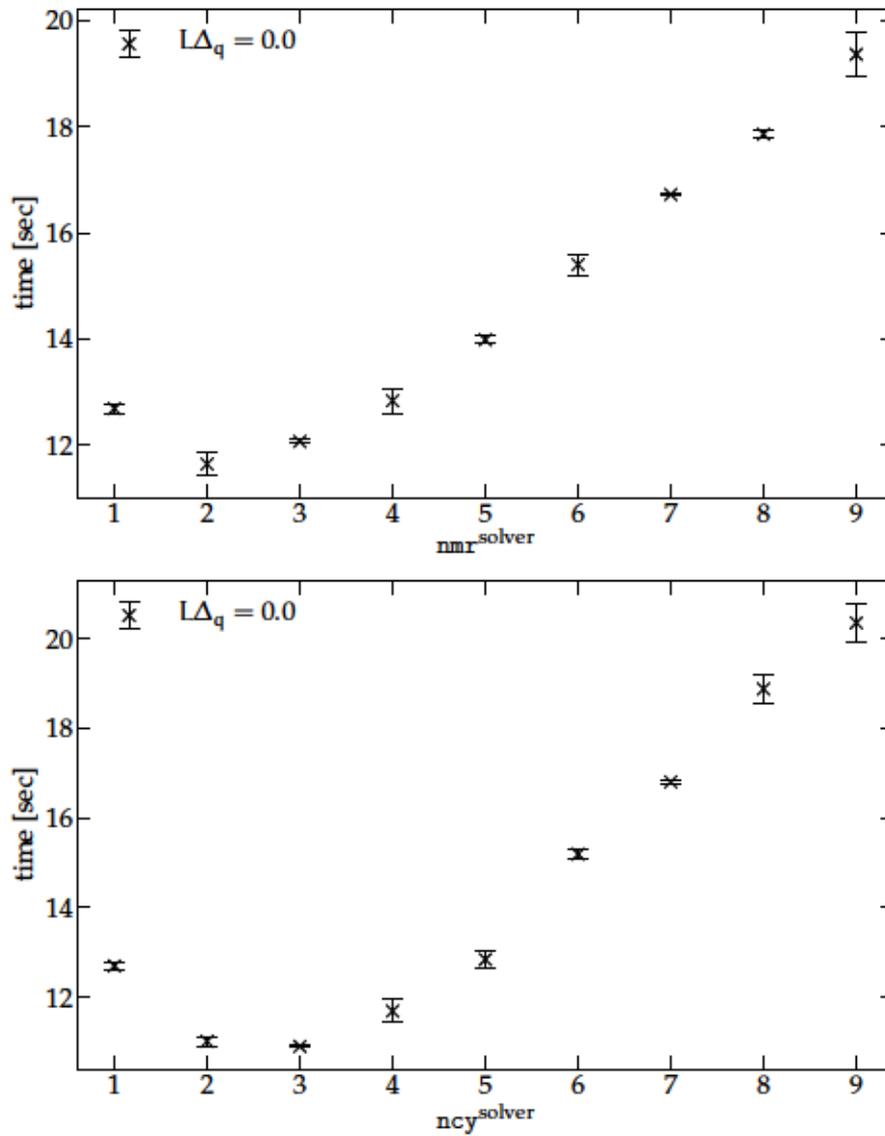


Figure 9.2: Scans for the parameters nmr and ncy of the deflated solver in the openQCD package. Timings are for the computation of two light propagators.

and did not see a negative impact on the performance when decreasing the value of κ^{df} . In our simulations and measurements, we set κ^{df} close to the value of the sea quark hopping parameter.

We do not show an explicit scan of the SAP and deflation block sizes. Based on the experiences in [80, 81, 128, 129], we set these blocks to the smallest possible values of 4^4 for most lattices. A slight increase of the single block sizes does not have a significant influence on the performance. For lattices with $L/a = 20, 40$, where due to the implementation in openQCD only blocks of size 10^4 can be used, the performance is significantly worse than for smaller block sizes on other geometries.

We conclude this section with the observation that the local minima in the parameter scans shift when we switch to other geometries or different regions in the physical parameter space. Especially when increasing the number of lattice sites, a scan of the performance around the conventional values may be necessary to achieve the optimal performance. For example, it is beneficial to increase the number of sources, when larger volumes are considered. An additional complication is introduced by the fact that many parameters are not independent of each other and the optimal value of one parameter can shift if another one is adjusted.

The above results can also be used for Monte Carlo runs using the deflated solver. In this case, the stability of the solver along the Monte Carlo run has to be considered as important criterion for the tuning of the solver parameters. For example, it might be beneficial to choose a small value of μ to stabilize the inversions.

9.2 RHMC

The complexity of the setup of the HMC and the RHMC algorithm is increased when hierarchical integration schemes and frequency splitting are used to optimize the performance. Whereas the hierarchy given in section 5.2.1 for the Hasenbusch frequency splitting provides a simple rule to get a close-to-optimal performance, the parameter space is more complex when we split the rational determinant

The number of poles in the Zolotarev approximation increases with the volume of the lattice, if the bound on the deviation from the true value is held constant. For our most demanding simulations with $L/a = 48, 64$ we use more than ten poles and we can split up the rational determinant to integrate small residues on large time scales using the deflated solver. The multi-shift conjugate gradient solver is used for the first term in eq. (5.34).

In this section we will look into the performance of the RHMC algorithm depending on the number of pseudofermion fields including single poles. We do not perform a large systematic test, but rather like to investigate the general behavior. For this test we use an en-

In most cases, the block size is tightly constrained by the size of the local lattice.

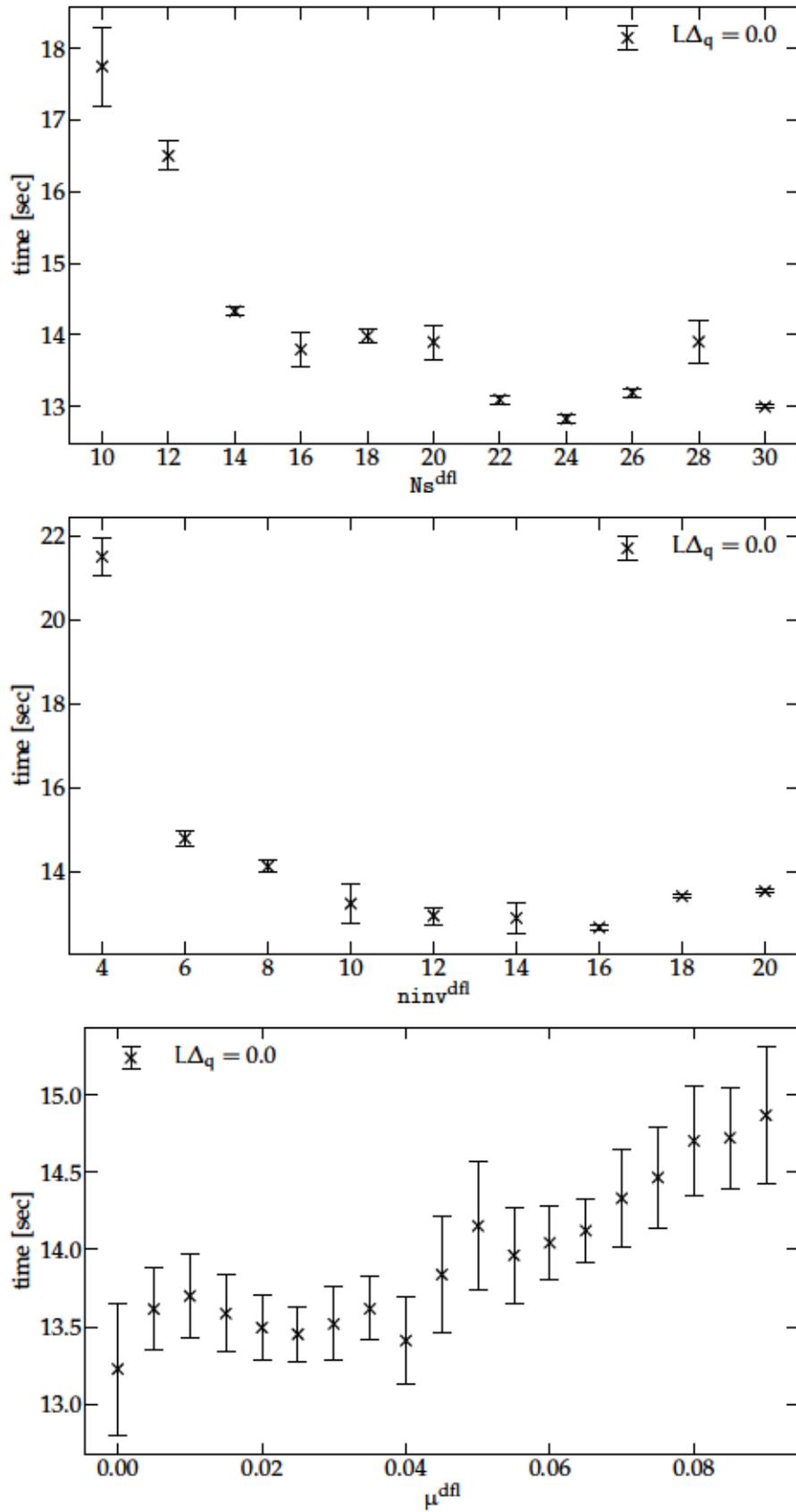


Figure 9.3: Scans for the parameters N_s , n_{inv} and μ of the deflation subspace in the openQCD package. Timings are for the computation of two light propagators.

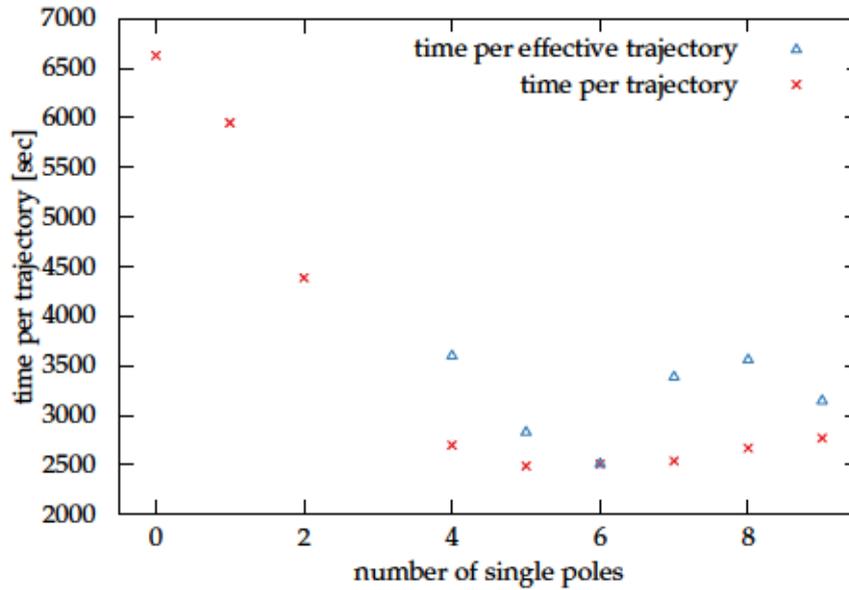


Figure 9.4: Time for one Monte Carlo trajectory of length 2 MDU for $L/a = 48$ with 13 poles in the Zolotarev approximation. The rational determinant is split into $1 + n$ terms, where n is the number of terms including one single pole.

semble with $L/a = 48$ lattice points and 13 poles in the Zolotarev approximation. We set up short Monte Carlo runs to judge on the performance by looking at the time per trajectory of length 2 MDU and the acceptance rate.

Since the runs include $O(10)$ trajectories, the acceptance rate might be insufficiently determined.

We split up the rational determinant according to (5.34) with n terms including one single pole of the Zolotarev approximation each and one term including $(13 - n)$ poles. We show the times per trajectory for n in the range $[0, 9]$ in figure 9.4. The same initial state has been used for all runs and we checked that the time per trajectory remained constant within each run. Apart from the setup of the splitting of the determinant, all algorithmic parameters remain the same for all runs.

We see that the cost for one trajectory can be drastically reduced when we split the determinant in several terms. We also show the time per trajectory divided by the average acceptance probability $\langle P_{\text{acc}} \rangle$. In this way, setups that have a negative impact on the acceptance rate are penalized. For the first three data points, $\langle P_{\text{acc}} \rangle < 0.5$. In this case, we would have to reduce the step size of the integration to obtain proper acceptance rates. This would further increase the cost. The same is true for the run with three single poles, where the solver failed for the chosen step size.

Starting from four poles, the time per trajectory remains approximately constant and changes in the acceptance rates due to an improvement of the algorithmic setup cannot be distinguished from fluctuations due to the small statistics. As soon as the smallest contributions are separated from the other poles, integrated on large time

scales and treated with the deflated solver, the performance does not change significantly. In the parameter tuning for our run in $L_1/a = 48$, the setup using six single poles has been used to perform further optimizations.

As soon as the rational determinant is split, we are able to integrate the highest poles on larger time scales as the other ones. In our simulations on large ensembles we used a hierarchical integration scheme with three levels where the innermost level has been used to integrate the gauge force. The outermost level included single poles of the rational determinant and one part of the HMC determinant which was split via Hasenbusch frequency splitting. The tuning of the acceptance rate has been done via the step size of the outermost level.

9.3 SCALING OF AUTOCORRELATION TIMES

The scaling of autocorrelation times has already been investigated for a range of gradient flow couplings and lattices with size $L/a = 16, 24, 32$ in [45]. There, the expected scaling of the autocorrelation time with a^{-2} [58] could be confirmed by an investigation of the integrated autocorrelation time τ_{int} of the gradient flow coupling.

We show the autocorrelation time of the gradient flow couplings of the ensembles that will be used for the matching procedure in figure 9.5. We see the scaling of τ_{int} with a^{-2} and observe that the absolute value of τ_{int} is significantly larger for the two sets of ensembles in the larger volume L_2 than for the two sets of ensembles in the matching volume. This is in line with the findings in figure 10 of [45]. Since we determine HQET observables which show only small autocorrelation on the L_2 lattices we expect that the precision of the calculation will not be impaired by the increase in τ_{int} .

9.3.1 *The topological charge in finite volume*

Since we work in finite volume and with small values of the lattice spacing, we do not see a free fluctuation of the topological charge. For the ensembles in the volumes L_0 and L_1 the topological charge is frozen and all ensembles are in the sector of vanishing topological charge. For the ensembles in L_2 we see topological activity which is suppressed when the lattice spacing is reduced.

In our renormalization scheme, we define the gradient flow coupling and the quark masses from ensemble averages in the zero topological sector, i.e., we employ the projection

$$\langle \mathcal{O} \rangle = \frac{\langle \mathcal{O} \hat{\delta}(Q) \rangle}{\langle \hat{\delta}(Q) \rangle} \quad (9.3)$$

We started the thermalization process with a cold start to ensure $Q = 0$ at the beginning of the MC run.

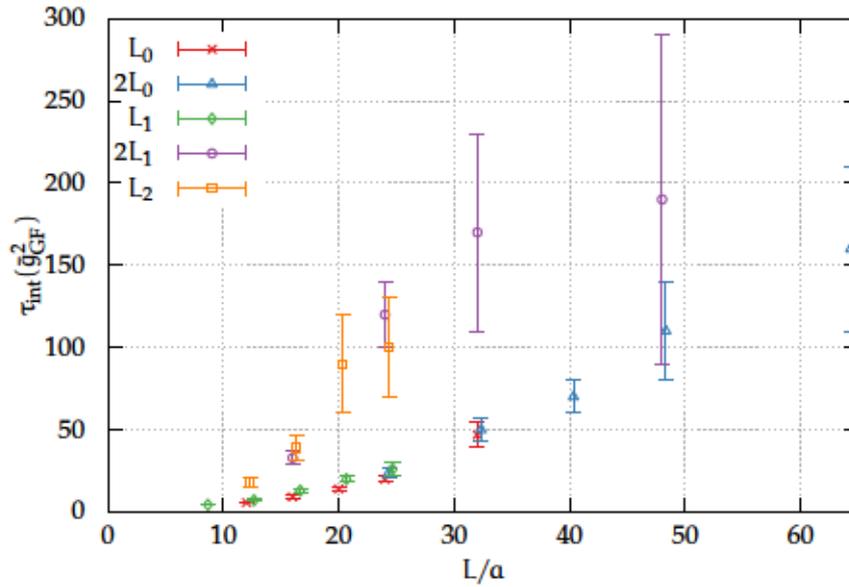


Figure 9.5: Integrated autocorrelation time τ_{int} of the gradient flow coupling g_{GF}^2 in units of MDU for the different sets of ensembles generated for the matching. The points are slightly displaced to improve visibility.

using eq. (6.50). Therefore, we have to monitor the topological charge on all configurations. We determine Q from flowed gauge fields as defined in eq. (6.55) with the Zeuthen flow and $c = 0.3$.

In the upper part of figure 9.6 we show the Monte Carlo history of one replica of the ensemble with $2L_1/a = 16$. We can see that the topological charge is frozen to the sector zero for most of the time. In some cases, the instantons forming on the lattice remain stable for a longer time and the charge is frozen in sectors with $Q \neq 0$. We see that the absolute size of Q outside sector zero is shifted from integer values towards zero. This is based on cut-off effects and integer values are restored in the continuum limit.

In the middle of figure 9.6 we show the dependence of the topological charge on the gradient flow time t for 200 configurations. During the evolution of the gauge fields in the flow time we have evaluated the observable up to $c = 0.4$ in steps of $c = 0.01$. We see that the smoothing procedure reduces the spread of the values of Q from the whole range $[-3, 3]$ to four distinct regions. The choice $c = 0.3$ seems to be just enough to be in a region where Q remains constant

When we compare the determination of Q from the Zeuthen flow Q^{ZN} with the determination using the Wilson flow Q^{W} , we note that more spikes, i.e., deviations from $Q = 0$ are present in the case of the Zeuthen flow. In [222] it is argued, that this might be based on small vacuum fluctuations that are promoted to instanton-like objects by the gradient flow. In the bottom plot of figure 9.6 we show the evolution of Q with t for the Zeuthen and the Wilson flow for three exemplary

We investigate this more closely in chapter 11.

$$c = \sqrt{8t}/L$$

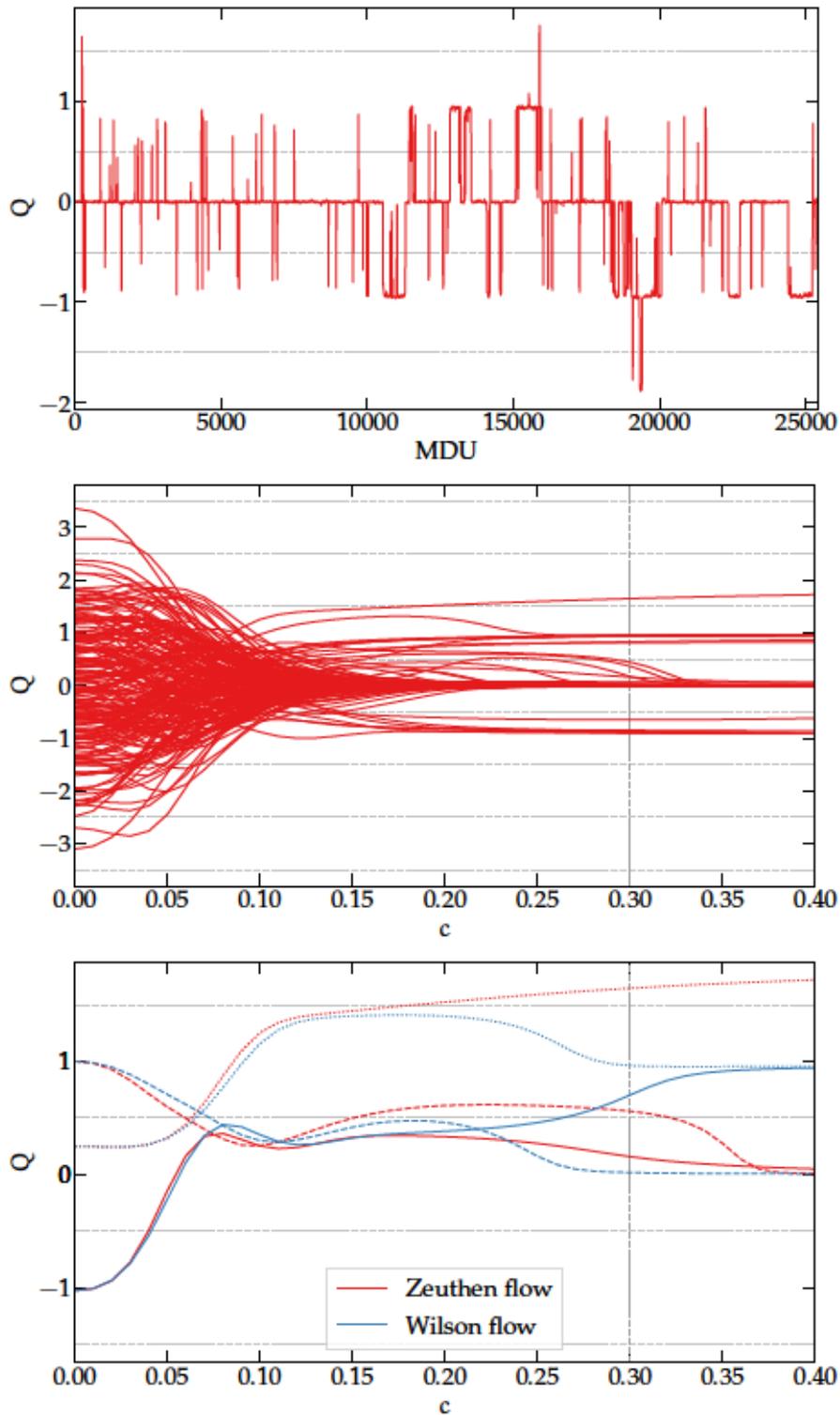


Figure 9.6: Topological charge Q on the ensemble with $2L_1/a = 16$. Sectors of integer charge are visualized by horizontal dotted gray lines. We define Q at $c = 0.3$. *Top*: Part of the Monte Carlo history. *Middle*: Dependence of Q on the flow time c . *Bottom*: Comparison of Q determined from Zeuthen and Wilson flow for three exemplary configurations.

configurations where the integer topological charges determined from both versions of the gradient flow differ at $c = 0.3$. We show Q from the Zeuthen flow in red and from the Wilson flow in blue and depict different configurations by the linetype.

Per definition, Q^{ZN} and Q^W coincide at $t = 0$. For the solid and the dotted lines, the curves deviate from each other starting at $c \approx 0.2$ and end up in two distinct topological sectors. In the case of the dashed lines, Q^{ZN} coincides with Q^W at $c = 0.4$ but not at $c = 0.3$. In general, the number of cases where both definitions of the flow lead to different results is small compared to the number of gauge configurations and the effect is suppressed in the ensemble average. In any case, we are not able to judge which of the two definitions leads to a more reliable result, since the determination of Q from the chiral modes of the Dirac operator is not accessible using Wilson fermions.

Part III

**RENORMALIZATION AND IMPROVEMENT
FROM FINITE-VOLUME SIMULATIONS**

STRATEGY

When we discussed the improvement and renormalization of quark masses in lattice simulations with Wilson quarks in section 4.3, we have seen that a large number of improvement coefficients and renormalization coefficients is needed to determine renormalized quark masses without discretization effects of $O(a)$. In our work, we determined the renormalization constant Z , introduced in eq. (4.28), the improvement coefficient b_m and the combination $b_A - b_P$ in a setup with three massless sea quarks and Schrödinger functional boundary conditions.

The methods we use to determine these observables have their origin in the work of [223], where the same observables have been determined non-perturbatively in the quenched approximation and it was extended in [224–226]. In all of these works (anti-)periodic boundary conditions have been used. The use of Schrödinger functional boundary conditions was first done by the ALPHA collaboration in [227] and [228] in the quenched approximation. This work has been extended to the theory with two massless quarks in [229].

In our work, we refined the strategy which has been applied in [229]. This refined strategy has been published together with the results for one of the two investigated parameter regions in [74] and we will explain it thoroughly in this chapter. We will start with the setup, as it has been used in the quenched and the two flavor case, introduce a parametrization of the current quark masses and derive the new strategy.

10.1 ESTIMATORS FROM TIME SLICE AVERAGES

The determination of estimators R_X for Z , b_m and $b_A - b_P$ is based on two of the definitions for renormalized quark masses in section 4.3. We use the definition from bare subtracted quark masses and from bare current quark masses. We start by introducing two quark flavors 1 and 2 with $m_{q,1} < m_{q,2}$ and a third flavor with the hopping parameter chosen such that

$$m_{q,3} = \frac{1}{2}(m_{q,1} + m_{q,2}) \quad (10.1)$$

holds. If we look at the definition of m_{ij} in eq. (4.27), we can convince ourselves that we can define the estimators based on following ratios

$$\begin{aligned} R_{AP} &\equiv \frac{2(2m_{12} - m_{11} - m_{22})}{(m_{11} - m_{22})(am_{q,1} - am_{q,2})} \\ &= (b_A - b_P) \{1 + O(am_{q,12}; a\text{Tr}[M_q])\} \end{aligned} \quad (10.2)$$

$$\begin{aligned} R_m &\equiv \frac{4(m_{12} - m_{33})}{(m_{11} - m_{22})(am_{q,1} - am_{q,2})} \\ &= b_m \{1 + O(am_{q,12}; a\text{Tr}[M_q])\} \end{aligned} \quad (10.3)$$

$$\begin{aligned} R_Z &\equiv \frac{m_{11} - m_{22}}{m_{q,1} - m_{q,2}} + (R_{AP} - R_m)(am_{11} + am_{22}) \\ &= Z \{1 + O(a^2; a\text{Tr}[M_q])\}. \end{aligned} \quad (10.4)$$

Here, the three quark flavors 1, 2, 3 are defined as specified above and as in eq. (4.35), we consider quark masses m_{ii} as defined from two mass degenerate but distinct flavors.

We have indicated the expected cut-off effects which are present when the estimators are defined as above. All estimators suffer from $O(a\text{Tr}[M_q])$ effects which disappear when we work with massless sea quarks. If a setup with finite quark masses is chosen, a chiral extrapolation to the point of vanishing quark masses removes these effects.

R_{AP} and R_m further suffer from $O(am_{q,12})$ effects. Since these coefficients multiply factors of a , these effects become $O(a^2)$ in the final observables. In the definition of R_Z only effects of $O(a^2)$ are present which we do not have to worry about since this is beyond the order which is considered in the improvement. A definition of R_Z without the term proportional to $(R_{AP} - R_m)$ would lead to $O(a)$ effects. These effects would also be present in every observable which is renormalized by Z .

Different definitions of the estimators based on a replacement of the quark mass difference $(m_{11} - m_{22})$ by other combinations of current quark differences from the three flavors 1, 2, 3 are possible and a priori we do not know, how the size of the cut-off effects is influenced by these choices. In [230], five different possibilities have been investigated and no significant difference has been found. Therefore, the above definitions which have been used in the former studies in the Schrödinger functional have been chosen in our work.

So-far, we did not further specify how the current quark masses are defined. In the derivation of current quark masses we stressed that the PCAC relation is valid on every single time slice of the lattice. Therefore, we can define time-dependent estimators $R_X(x_0)$, $X = Z, m, AP$ based on local PCAC masses $m_{ij}(x_0)$ on every time slice of the lattice. We expect estimators from different time slices to be strongly correlated and to differ only by cut-off and boundary effects. Since

A finite distance to the boundaries is assumed.

all definitions based on different x_0 are equally valid, an average over time slices in a region $[t_i, t_f]$ in the center of the lattice

$$R_X = \frac{a}{t_f - t_i + a} \sum_{x_0=t_i}^{t_f} R_x(x_0) \quad (10.5)$$

is chosen to eliminate fluctuations between different time slices.

10.2 QUARK MASS PARAMETRIZATION

Due to possible numerical instabilities in the computation of the estimators close to the unitary point and further theoretical advantages we refine the three estimators using a different definition of the quark masses that enter eqs. (10.2–10.4). The basis of this redefinition is a parametrization of current quark masses and we will explain it closely along the lines of our publication [74].

10.2.1 Redefinition of the estimators

We start with the observation that the current quark masses m_{ij} are symmetric functions under the exchange of i and j . Therefore, we can express them in the series

$$am_{ij}(am_{q,ij}, a\Delta_{q,ij}) = \sum_{n,k=0}^{\infty} C_{nk}(a\Delta_{q,ij})^{2n}(am_{q,ij})^k, \quad (10.6)$$

where we have introduced the dimensionless coefficients C_{nk} and the mass splitting

$$\Delta_{q,ij} \equiv \frac{1}{2}(m_{q,i} - m_{q,j}) = \frac{1}{4a} \left(\frac{1}{\kappa_i} - \frac{1}{\kappa_j} \right). \quad (10.7)$$

If we compare this parametrization with the definition of the renormalized and improved quark mass in eq. (4.27), we can identify the coefficients C_{00}, C_{01}, C_{02} and C_{10} with combinations of improvement coefficients and renormalization constants. In the limit of chiral sea quark masses, we would get $C_{00} = 0$ and $C_{01} = Z$. Similar observations have been made in [223] and [227].

We can use the series expansion to parameterize degenerate and non-degenerate quark masses in terms of mass-splittings. To adapt this to our goal of the determination of the estimators R_X , we make some specific choices. We choose $m_{q,1}$ to be equal to the mass of the three degenerate sea quarks, i.e., $m_{q,1} = \text{Tr}[M_q]/N_f$. Due to our definition of $m_{q,3}$ in eq. (10.1), we can now employ the mass splitting

$$\Delta \equiv \Delta_{q,21} = \frac{1}{2}(m_{q,2} - m_{q,1}) = m_{q,3} - m_{q,1} \quad (10.8)$$

The exact identifications are not important here, but have been given in (2.13) of [74].

to parameterize the current quark masses needed in the determination of the R_X via

$$am_{11} = \sum_{k=0}^{\infty} C_{0k} (am_{q,1})^k \quad (10.9)$$

$$am_{22}(\Delta) = \sum_{k=0}^{\infty} C_{0k} (am_{q,1} + 2a\Delta)^k \quad (10.10)$$

$$am_{12}(\Delta) = \sum_{k=0}^{\infty} C_{nk} (a\Delta)^{2n} (am_{q,1} + a\Delta)^k. \quad (10.11)$$

In the unitary limit, $\Delta \rightarrow 0$, all three quark masses are equal. Additionally we can see from the formulae that the first derivatives are related via

$$\frac{1}{2} \frac{\partial m_{22}}{\partial \Delta} \Big|_{\Delta=0} = \frac{\partial m_{12}}{\partial \Delta} \Big|_{\Delta=0} = \frac{\partial m_{11}}{\partial m_{q,1}}. \quad (10.12)$$

We can use these observations to parameterize the quark masses in terms of Δ close to the unitary point by

$$am_{12}(\Delta) = am_{11} + N_1 a\Delta + N_2 (a\Delta)^2 + O(a\Delta)^3 \quad (10.13)$$

$$am_{22}(\Delta) = am_{11} + 2N_1 a\Delta + 4D_2 (a\Delta)^2 + O(a\Delta)^3, \quad (10.14)$$

where we introduce the the coefficients N_i and D_i for the non-degenerate and degenerate quark masses. By comparison with eqs. (10.9–10.11) we can see that these are related to the C_{nk} and the sea quark mass $am_{q,1}$. Since we have chosen $m_{q,3} = m_{q,12}$, we can use the identity $m_{33}(\Delta) = m_{22}(\Delta/2)$ to write

$$am_{33}(\Delta) = am_{11} + N_1 a\Delta + D_2 (a\Delta)^2 + O(a\Delta)^3. \quad (10.15)$$

The term $O(a\Delta)^3$ indicates terms of higher orders. These can be included in the parametrization of the masses, leading to additional coefficients. The definitions of the current quark masses m_{22} , m_{12} and

m_{33} can now be used in the definitions of the R_X , eqs (10.2–10.4), to obtain continuous parametrizations in terms of $(a\Delta)$. We can write

$$\begin{aligned} R_{AP} &\equiv \frac{\sum_{k=0}^{\infty} (N_{k+2} - 2^{k+1} D_{k+2}) (a\Delta)^k}{N_1 + \sum_{k=0}^{\infty} 2^{k+1} D_{k+1} (a\Delta)^k} & (10.16) \\ &= \frac{N_2 - 2D_2 + O(a\Delta)}{N_1 + O(a\Delta)} \xrightarrow{\Delta, m_{q1} \rightarrow 0} (b_A - b_P), \end{aligned}$$

$$\begin{aligned} R_m &\equiv \frac{\sum_{k=0}^{\infty} (N_{k+2} - D_{k+2}) (a\Delta)^k}{N_1 + \sum_{k=0}^{\infty} 2^{k+1} D_{k+1} (a\Delta)^k} & (10.17) \\ &= \frac{N_2 - D_2 + O(a\Delta)}{N_1 + O(a\Delta)} \xrightarrow{\Delta, m_{q1} \rightarrow 0} b_m, \end{aligned}$$

$$\begin{aligned} R_Z &\equiv \left(N_1 + \sum_{k=1}^{\infty} 2^k D_{k+1} (a\Delta)^k \right) \\ &\quad + 2 \left(2am_{11} + N_1(a\Delta) + \sum_{k=2}^{\infty} 2^k D_k (a\Delta)^k \right) \\ &\quad \times \frac{\sum_{k=0}^{\infty} (1 - 2^{k+1}) D_{k+2} (a\Delta)^k}{N_1 + \sum_{k=0}^{\infty} 2^{k+1} D_{k+1} (a\Delta)^k} & (10.18) \\ &= N_1 - \frac{2D_2 am_{11}}{N_1} + O(a\Delta)^2 \xrightarrow{\Delta, m_{q1} \rightarrow 0} Z. \end{aligned}$$

In the first step, we have included all higher orders in the sums. In practice, we have to cut at some order of $(a\Delta)$.

From the above equations, we can see that we are able to define the three estimators at the unitary point, i.e., at the point where all sea and valence quark masses are equal. Since the definitions in eqs. (10.2–10.4) are based on finite differences of quark masses, this has not been obvious. If we go even further and drive all masses to the chiral point where sea and valence quark masses vanish, all mass-dependent cut-off effects are removed from the definitions. At this point, the coefficients m_{11} , N_1 , N_2 and D_2 are sufficient to fully describe the three estimators.

*Per definition,
 $m_{11} = 0$ at the
chiral point.*

If we evaluate the R_X at a finite value of Δ , all higher orders of the quark mass parametrizations in eqs. (10.13–10.15) contribute to the estimators. This is also the case, if they are determined from finite differences of quark masses and time slice averages, as it has been done in the former works [227–229].

10.2.2 Polynomial interpolations

The determination of the estimators R_{AP} , R_m and R_Z , consists of two steps. In the first step, we have to extract the coefficients m_{11} , D_i and N_i from current quark masses. In the second step, we can employ these coefficients to determine the R_X .

The extraction of the coefficients can be achieved via a polynomial fit of fixed degree to measured degenerate and non-degenerate PCAC masses. Hence, quark masses m_{12} and m_{22} have to be determined for different choices of Δ . Quark masses m_{33} are indistinguishable from m_{22} and lead to the same parametrization. Therefore, we do not have to treat them separately. For the description of the masses by a polynomial with fixed degree to be valid, we have to restrict ourselves to a region around $a\Delta = 0$.

The degree of the chosen polynomials depends on the range of $a\Delta$ considered in the measurements. If we consider a region close to the origin, we are only able to resolve linear and quadratic effects in $a\Delta$. If we choose a region which extends to heavier quark masses, we have to add more degrees of freedom to properly describe the data. The number of fit coefficients which are needed to describe the data also depends on the statistical errors of the data points. In general, if the data points are more precise, we need have to consider polynomials of higher degree.

The coefficients of the quadratic terms are necessary to determine the estimators at the unitary point where all higher order coefficients vanish. If the R_χ are extracted at non-zero values of Δ , it has to be granted that all coefficients which significantly contribute to the estimators are taken into account. This can be checked by the quality of the polynomial fits in terms of the χ^2 and the deviation of the current quark masses from the interpolations as well as by an explicit comparison of the results from eqs. (10.16–10.18) with the results gained from eqs. (10.2–10.4).

A simultaneous fit of the current quark masses to the forms in eq. (10.13, 10.14) where the parameters m_{11} and N_1 are shared among both parametrizations is performed to extract the coefficients of the polynomials from the data. The constant term $a m_{11}$ can be fixed to the value of the sea quark mass. Since the statistical error is not taken into account in this case, it might constrain the fit too much.

The quark masses based on different valence quark propagators are highly correlated. Therefore we would like to determine the coefficients from correlated fits. Unfortunately, as anticipated in section 6.3.1, this turns out to be highly unstable. We therefore use the statistical errors as weights in an uncorrelated fit. The coefficients of the polynomial fit can be determined with standard methods, see e.g. [124].

10.3 THE LINE OF CONSTANT PHYSICS

Our aim is to determine the renormalization constant Z and the improvement coefficients $b_A - b_P$ and b_m for a range of couplings for the use in the continuum extrapolation of physical observables. Therefore, we like to achieve a smooth variation of the estimators, when the

coupling g_0^2 is changed. To achieve this behavior, we determine the estimators along a line of constant physics.

This means that we keep all kinematic variables and all renormalized quantities fixed in terms of L and we also fix L , when we vary g_0^2 , thereby changing the lattice spacing. With this procedure, the estimators become smooth functions in a/L and g_0^2 [227].

When we define the line of constant physics, we have to make certain decisions in fixing the parameters. Every variation in one of our choices leads to a different behavior of the estimators R_X in the coupling. As long as one set of choices is kept fixed for all couplings, the smooth convergence to the continuum limit is preserved. Although we will carefully motivate all of our choices, it is not clear a priori which choice leads to the best continuum extrapolation of physical observables. This does not pose a fundamental problem, since the concept of $O(a)$ improvement is asymptotic in the sense that we do not try to make statements about the absolute size of higher order effects but try to change the rate towards the continuum limit.

Based on these thoughts, we do not consider the spread of the variations in the estimators based on different choices as systematic errors, but rather define the estimators based on a fixed set of choices. Any deviation from this definition results in an ambiguity that is expected to vanish smoothly in the continuum limit. This approach has been motivated in [51, 227].

In our determination of the R_X , we will motivate our choices and check that ambiguities vanish with the expected rate which is $O(a)$ for the improvement coefficients and $O(a^2)$ for the renormalization constant.

In our calculations, we first have to fix all details of the lattice simulations such as the improvement coefficients c_X and the boundary angles in the Schrödinger functional. For many of these choices, we already stated the exact definitions in the introductory part of this work. All parameters that influence the generation of our ensembles can not be changed and the inspection of the impact of changing these parameters requires a large amount of computing time.

The next set of choices is fixed on the level of the analysis of the lattice data, e.g., the definition of the discretized derivatives, the plateau range for the current quark masses or the degree of the polynomial in the mass parametrizations. Since these choices may be changed on the level of the analysis, we can check that ambiguities arising from different definitions vanish smoothly in the continuum limit.

The last set of parameters concerns physical observables that enter the definitions of the R_X . Most importantly we have to fix L in physical units, since we use this to fix all other quantities. For our calculation, the choice of the masses m_{11} and m_{22} is of importance, as well. While we fix m_{11} to be the current quark mass of the sea quark and tune

these to be massless, we have a freedom in the choice of the valence quark mass m_{22} .

By working on ensembles with massless sea quarks, we are able to remove all $O(a\text{Tr}[M_q])$ effects that would be present in our estimators. In the same spirit, we can choose

$$L\Delta_{22} \equiv L(m_{22}(\Delta) - m_{11}) = 0, \quad (10.19)$$

as part of the definition of the line of constant physics for the partially quenched valence quarks. This definition together with $m_{11} = 0$ leads to the removal of all quark mass dependent cut-off effects in eqs. (10.2–10.4).

In [229], two valence LCP have been defined by the choices $Lm_{22} \approx 0.5$ and $Lm_{22} \approx 2.5$. Since the estimators have been determined from finite differences of quark masses, the massless definition of eq. (10.19) has not been possible in this case. The scaling of the effective energy

$$L\Gamma_{\text{PS}}(z, g_0) = -L\delta_0 \ln [f_A(x_0, M)]|_{x_0=L/2}, \quad z = LM \quad (10.20)$$

determined on a different set of ensembles than used in the computation for the R_X towards the continuum limit has been investigated using both valence LCP. The RGI quark masses z where chosen to be at the scale of the bottom quark mass. It was found that the estimators based on the heavy LCP, i.e., $Lm_{22} \approx 2.5$ led to a longer regime where a scaling proportional to $(a/L)^2$ could be observed, compared to the light LCP. The inclusion of mass-dependent cut-off effects in the definition of the estimators led to a better removal of cut-off effects for quantities based on heavy quarks. A similar observation has been made in [228].

Based on these observations, we come to the conclusion that we can define valence LCP by requiring $L\Delta_{22}$ to be fixed to a constant, finite value. If this value is chosen close to to a characteristic heavy quark scale, where the improvement coefficients are used to compute physical quantities, an improved scaling may be expected. Effectively, this definition of the estimators at finite valence quark mass corresponds to a resummation of all higher order terms in $a\Delta$ for the chosen line of constant physics [74].

Since the connection between the current quark mass and the bare subtracted quark masses is not known a priori, a tuning of the hopping parameters of the valence quarks has been necessary to achieve approximately constant quark masses Lm_{22} in the former studies. Since the parametrization of the current quark masses in terms of Δ is the core of our strategy, we are able to evaluate the estimators R_X at any fixed choice of $L\Delta_{22}$ at the level of the analysis and after all measurements have been performed. Since we have a continuous description of the quark masses in a range of values of $L\Delta_{22}$, we are also able to compute the estimators as smooth functions of Δ_{22} .

M is the RGI mass of the heavy quark.

We perform a similar test in chapter 12.

We like to end this discussion with the observation that fixing $L\Delta_{22}$ might be insufficient, if the data reaches a certain precision. In order to correctly fix the renormalized quark masses which are used to define the valence line of constant physics, the combination

$$\frac{Z_A(g_0^2)}{Z_P(g_0^2, 1/L)} L\Delta_{22} \quad (10.21)$$

has to be held fixed. Since the factor Z_A/Z_P only slowly changes in the considered ranges of g_0^2 , this additional factor might be left out in the computation, as it has been done in [227–229].

QUARK MASS IMPROVEMENT IN THE STRONGLY COUPLED REGIME OF THREE-FLAVOR LATTICE QCD

We have determined the renormalization constant Z and the improvement coefficients $b_A - b_P$ and b_m in two different regimes of three-flavor QCD. In this chapter we describe the calculation in the strongly coupled regime of QCD. Here, we cover the coupling regime that is used in the large-volume CLS simulations. The results presented in this chapter have been published in [74].

11.1 GAUGE ENSEMBLES

For our study, we employ a set of ensembles that has already been used for the non-perturbative determination of c_A [67] and Z_A [65], but has been extended for the present study. We use the tree-level Symanzik improved gauge action and clover improved Wilson fermions as they have been defined in chapter 2 in the Schrödinger functional with all choices as described in chapter 3. Based on these discussions, we only have to specify the boundary angles which we choose to be $\theta = 0$ and the boundary improvement terms which we set to tree-level, i.e., $c_t = \tilde{c}_t = 1$.

The lattice extent has been set to $L \approx 1.2 \text{ fm}$. Since the line of constant physics for these ensembles has been fixed before a scale setting had been done, the evolution of the lattice spacing with the bare coupling has been estimated from the universal two-loop β -function, starting from the pair $(L/a = 12, \beta = 3.3)$. To allow for large plateaus at moderate cost, $T = 3L/2 - a$ has been chosen. More details on the generation of these ensembles can be found in [65, 67] as well as [231].

Although the physical lattice extent varies by about 10% between the different ensembles, it has been checked explicitly that these small deviations do not influence the results obtained from Ward identities in the bulk at the order we are interested in [232].

An overview of the ensembles is given in table 11.1. With the five different choices of the coupling β we cover a range of lattice spacings from $a \approx 0.09 \text{ fm}$ to $a \approx 0.045 \text{ fm}$. When the line of constant physics was set up, it was intended to cover the range of lattice spacings used by the CLS effort. When the CLS ensembles were generated large cut-off effects have been found at $\beta = 3.3$. Therefore, only smaller lattice spacings have been considered and the finest ensemble has an inverse coupling of $\beta = 3.85$ which is slightly outside the range covered by the Schrödinger functional ensembles. Nevertheless,

In an old version of openQCD, it was not possible to simulate even time extents in the Schrödinger functional.

An overview of the CLS ensembles used in our work can be found in fig. 13.1

we are confident that interpolation formulae for the whole range of CLS ensembles can be provided from our work in the Schrödinger functional.

As apparent from tab. 11.1, we have generated several ensembles per gauge coupling, each with a different sea quark hopping parameter κ_1 . This is done to allow us to extrapolate the results at fixed lattice spacing to the chiral point and therefore cancel all $\text{Tr}[M_q]$ effects. Since the sea quark masses in all ensembles are fairly small, we expect linear effects in the sea quark mass and a small slope, compared to the errors of our estimators. Since an ensemble with a small negative quark mass exists for every lattice spacing, we can perform an interpolation rather than an extrapolation to the chiral point.

Our strategy as follows: We determine the estimators R_X on a line of constant physics on all gauge ensembles of tab. 11.1. Afterwards we interpolate the results to the chiral point at each lattice spacing. In the last step, we determine an interpolation formula to describe the results across the covered range of couplings.

11.1.1 Topological charge

As discussed in the motivation for the use of open boundary conditions in section 3.4 and in section 6.4 on quantities from the gradient flow in the Schrödinger functional, we expect a severe topological freezing when going to smaller and smaller lattice spacings. Therefore, we have to monitor the topological charge on all ensembles. In this chapter and in contrast to the work for the matching of QCD and HQET, we use the Wilson flow to smooth the gauge fields for the determination of Q .

We show the Monte Carlo histories of the topological charge for three representative ensembles $A1k_4$, $B1k_4$ and $D1k_4$ in figure 11.1. We can observe a strongly fluctuating charge on $A1k_4$. For $B1k_4$ we see that the charge remains in one sector before it changes again. At the finest lattice spacing considered in this study, the charge is frozen in some sectors for large fractions of the Monte Carlo history.

On the right hand side of figure 11.1, we show histograms with the distribution of Q . The gray bars are bins of width 1 centered around integer values, according to our definition of integer topological charges on the lattice in eq. (6.50). We can see that we have sampled a distribution around $Q = 0$, but the width of the distribution shrinks for decreasing lattice spacings. At the finest lattice spacing we essentially only sample $Q \in \{-1, 0, 1\}$.

We also included a histogram with red bars. Here, the bin width has been chosen to be 0.2 and with this resolution, we are able to investigate the fine structure of the non-integer topological charge. At the coarsest lattice spacing we observe a continuous distribution with peaks which are shifted towards zero, compared to the integer values.

*For better visibility
we show only a
fraction of each run.*

Table 11.1: Overview of the simulation parameters of the $N_f = 3$ ensembles (labeled by ID) that represent our data. Subsequent columns refer to the lattice dimensions L^3T/a^4 , the inverse gauge coupling $\beta = 6/g_0^2$, the light (sea) quark hopping parameter κ_1 , the number of replica N_r , the number of configurations per replicum, both in total (N_{cfg}) and in the subset of configurations with zero topological charge ($N_{\text{cfg}}^{(0)}$), and the corresponding PCAC sea quark masses. Most ensembles have configurations separated by 8 molecular dynamic units (MDU), except for A1k3 and D1k4 that have 4 and 16 MDU, respectively. Compared to the data base of [65, 67], we have generated and used the nearly chiral ensembles A1k3, A1k4, B1k4 and D1k4, and significantly increased statistics for E1k1 and E1k2. Reproduced from [74].

ID	$\frac{L}{a}$	$\frac{T}{a}$	β	κ_1	N_r	N_{cfg}	$N_{\text{cfg}}^{(0)}$	am_{11}	$am_{11}^{(0)}$
A1k1	12	17	3.3	0.13652	20	2560	935	-0.00166(61)	-0.00278(80)
A1k3	12	17	3.3	0.13648	5	1719	614	0.00262(130)	0.00079(118)
A1k4	12	17	3.3	0.13650	20	12080	4424	0.00030(29)	-0.00110(36)
E1k1	14	21	3.414	0.13690	32	4800	1694	0.00308(22)	0.00262(26)
E1k2	14	21	3.414	0.13695	47	7050	2653	0.00034(18)	-0.00022(22)
B1k1	16	23	3.512	0.13700	3	3328	1336	0.00562(14)	0.00549(21)
B1k2	16	23	3.512	0.13703	2	1151	395	0.00481(19)	0.00444(25)
B1k3	16	23	3.512	0.13710	2	2048	938	0.00164(16)	0.00107(20)
B1k4	16	23	3.512	0.13714	1	3482	1401	0.00002(14)	-0.00057(19)
C1k2	20	29	3.676	0.13700	4	1904	857	0.00619(7)	0.00600(11)
C1k3	20	29	3.676	0.13719	4	1934	1249	-0.00086(8)	-0.00109(11)
D1k2	24	35	3.810	0.13701	2	803	357	0.00084(8)	0.00079(10)
D1k4	24	35	3.810	0.137033	8	5313	3469	-0.00002(3)	-0.00007(3)

With the decrease of the lattice spacing, we obtain single distributions which are centered around integer values. As expected, we obtain integer values of the topological charge in the continuum limit.

When we defined the gradient flow coupling in the Schrödinger functional, we introduced the projection on the sector of vanishing topological charge via

We use
 $c = \sqrt{8t}/L = 0.35$
 in this chapter.

$$\langle O \rangle = \frac{\langle O \hat{\delta}(Q) \rangle}{\langle \hat{\delta}(Q) \rangle}. \quad (11.1)$$

We will use this projection for all observables obtained from Ward identities on the ensembles listed in 11.1. We expect that results from the $Q = 0$ sector are the same as in the case of the inclusion of all sectors, if they are sampled properly. For the coarse ensembles we have to neglect a significant part of the statistics when the projection is applied. In exchange we do not have to worry about biased results on the finest ensembles, where the sampling is insufficient.

In this chapter, we will focus on the projected results but also quote the results from full statistics. This will allow us to judge our decision, when we compare the final results.

11.2 RESULTS

In this section we will discuss our analysis and the results at all three stages of the calculation, namely the analysis on single ensembles, the chiral extrapolation at fixed lattice spacing and the interpolation of the results at the chiral point in the range of considered couplings.

11.2.1 Mass parametrizations and determination of the estimators

We start our discussion of the results with the description of the analysis on the base of single ensembles. Since the B1 ensembles are in an intermediate coupling range, where we still see a fluctuation of the topological charge and at the same time expect only moderate cut-off effects, we will use them to illustrate the steps of the analysis.

Since we are interested in the chiral limit of the R_χ , we try to sample a region of $a\Delta$ close to the unitary point. To be able to compare our results from the new strategy with the results from time slice averages, we determine not only m_{12} and m_{22} for each choice of Δ , but also m_{33} . This means that we determine two heavy propagators per choice of Δ . On each ensemble we perform measurements for $O(15)$ choices of Δ which are chosen such that $0 \leq L\Delta_{22} \leq 1$ holds. In addition to these measurements including partially-quenched quarks, we perform a measurement of at the unitary point, i.e., we determine m_{11} at $m_{q,1}$ which is fixed by the simulation.

We perform the
 measurements with
 the *sfcf* program
 [52].

We define the PCAC masses as the time slice averages in the central third of the time extent, i.e., in the region $x_0 \in [L/(2a), L/a]$. The

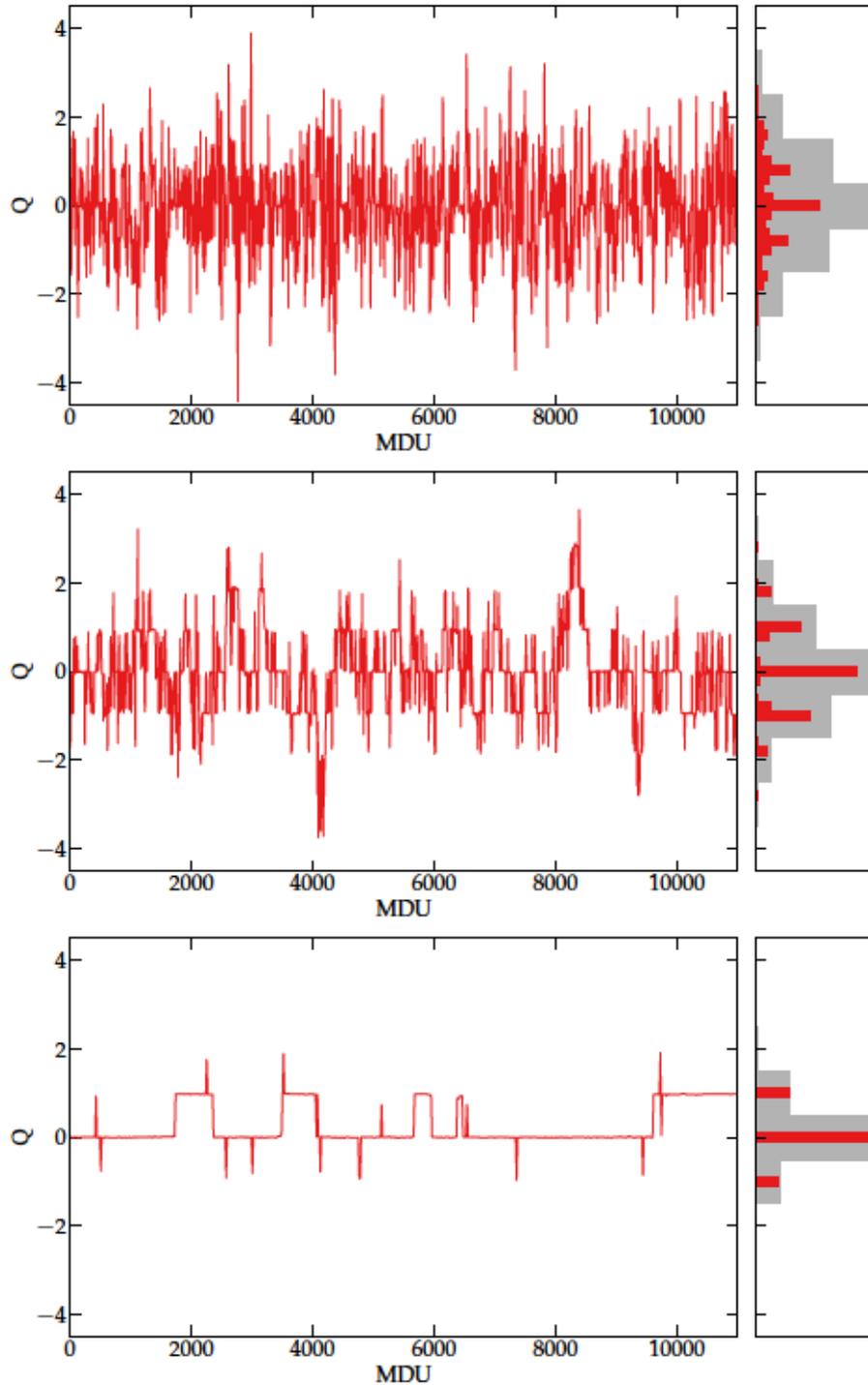


Figure 11.1: Beginning of the Monte Carlo histories and distributions of the topological charge Q for decreasing lattice spacing to monitor the topology freezing (top to bottom: ensembles A1k₄, B1k₄ and D1k₄). The gray histogram counts appearances of Q belonging to different integer sectors $\nu \in \mathbb{Z}$ according to $\delta_{Q-\nu,0}$ of eq. (6.50). The finer-spaced (red) histograms reveal the fine-structure of the (non-integer) topological charge distribution with Wilson fermions and gradient-flow smoothing ratio $c = 0.35$. The histograms have been determined on the full statistics.

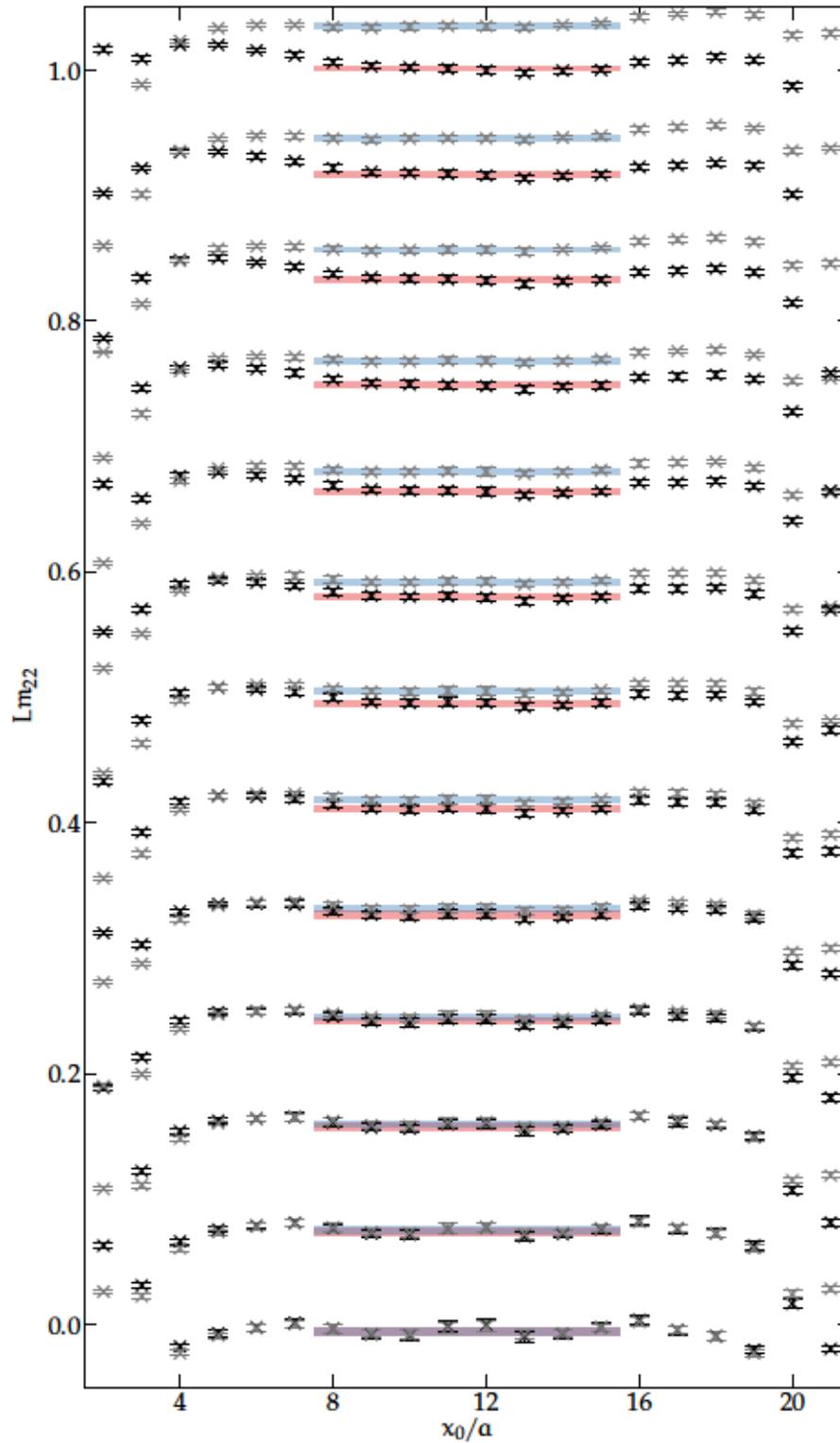


Figure 11.2: Time dependence of exemplary PCAC masses in the considered range on B1k4. Black points and red plateaus show the masses determined with improved derivatives. Gray points and blue plateaus depict masses from standard derivatives. The errors on single data points are barely visible.

values $m_{ij}(x_0)$ are calculated according to eq. (4.26) with the improved discretized derivatives as defined in A.3 and c_A from reference [67].

In figure 11.2, we show the time slice dependent PCAC masses $m_{22}(x_0)$ for different values of κ_2 . The black points show the quark masses on each time slice determined with the improved derivatives. The red regions show the plateaus which are determined from a time slice average. These plateaus enter the analysis. At the boundaries, we can see considerable cut-off effects, leading to a deviation from a straight line. Within the central region where we define our plateau, these boundary effects are not visible anymore. Slight variations within this plateau range are averaged out.

In gray, we show the the PCAC masses determined from the standard discretized derivatives together with the corresponding plateaus in blue. For the lightest quark masses, we can barely see any difference between both methods, except for the cut-off effects at the boundaries. With increasing quark mass, the difference between both plateaus increases. We therefore seem to observe a mass-dependent cut-off effect. In fact, the absolute size of the difference scales with the lattice spacing. We will investigate the influence of these cut-off effects on our final observables in section 11.3.

From our measurements of the correlation functions f_A and f_P , we have $O(15)$ values of non-degenerate quark masses m_{12} and $O(30)$ values of degenerate quark masses which we will name m_{22} from now on, unless we explicitly indicate something different. This choice leads to a higher density of measurements of m_{22} close to the unitary point. Since we cover a range up to $L\Delta_{22} = 1$ we are able to safely determine results for a line of constant physics with heavy valence quarks, in addition to the line of constant physics using massless valence quarks. We define these two lines of constant physics by

$$\text{LCP-0: } L = \text{const}, \quad Lm_{11} = 0, \quad L\Delta_{22} = 0, \quad (11.2)$$

$$\text{LCP-1: } L = \text{const}, \quad Lm_{11} = 0, \quad L\Delta_{22} = 1. \quad (11.3)$$

Based on the measured PCAC masses, we are able to determine the coefficients D_i and N_i of the parametrization formulae (10.13–10.14) from a simultaneous fit to the data. We find that it is advantageous to keep the parameter m_{11} in eqs. (10.13–10.14) as a free parameter compared to constraining it to the mean value of the measured PCAC mass at $m_{q,1}$.

Since the coefficients at the order $(a\Delta)^2$ are necessary for the determination of the fit coefficients at the unitary point and we like to absorb fluctuation in the data in higher order fit coefficients, we choose interpolating fits of degree three.

In figure 11.3 we present as representative example the PCAC masses and the interpolating curves from the fit on ensemble B1k4. In the lower part of the figure we show the data points and the interpolations. On this scale, the error bars of the data points as well as

Remember that
 $m_{22}(\Delta) =$
 $m_{33}(2\Delta)$ *holds.*

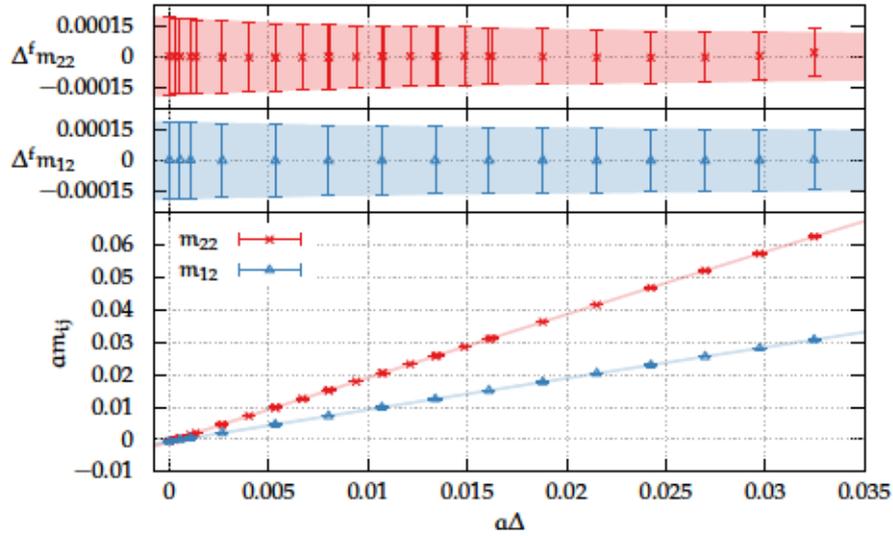


Figure 11.3: *Lower plot:* Example of a combined mass fit (ensemble B1k4) depicting the fitted data points and curves for m_{22} and m_{12} . Error bars and bands are too small compared to the scale of the plot. *Upper plots:* Differences between the measured PCAC masses m_{ij} and the fitted curves, $\Delta^f m_{ij} = a m_{ij}(\kappa_2) - a m_{ij}(\Delta)$ together with the statistical uncertainties of the data points and the uncertainty band of the curves. Reproduced from [74].

the uncertainty bands of the curves are hidden. This underlines the high level of precision that we achieve in the determination of PCAC masses. To be able to judge on the quality of the interpolation formula, we show the difference between data points and interpolation curves

$$\Delta^f m_{ij} \equiv a m_{ij}(\kappa_2) - a m_{ij}(\Delta) \quad (11.4)$$

together with the uncertainty bands of the interpolating formulas centered around zero. It can be seen that the interpolations describe the data very well with the chosen fit degree and we are able to parameterize the current quark masses in the complete range.

Based on the reassuring results, we can use the coefficients in $m_{12}(\Delta)$ and $m_{22}(\Delta)$ to determine the R_X according to eqs. (10.16–10.18) for both lines of constant physics. In tables B.4 – B.7 we list the results for the sea quark masses and the three estimators for all ensembles. Our main results are collected in table B.4, where we gathered the results for LCP-0 and the sector of vanishing topological charge. Table B.6 contains the results for LCP-1 and B.5 and B.7 contain the results based on the analysis on all topological sectors.

11.2.2 Chiral interpolations

We employ these results to perform interpolations in the sea quark mass to the chiral point at fixed lattice spacing. As it is already evident from the numbers, the variations of the estimators depending

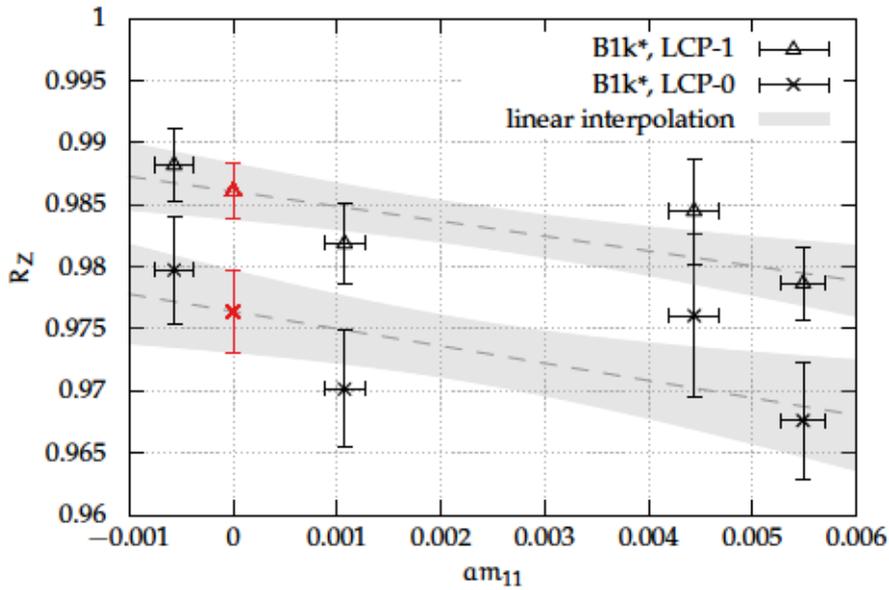


Figure 11.4: Chiral interpolation ($am_{11} \rightarrow 0$) of R_Z for $\beta = 3.512$ for the two lines of constant physics defined in eqs. (11.2) and (11.3) projected onto the sector of vanishing topological charge. The red triangles represent the resulting values in the chiral limit. Reproduced from [74].

on different sea quark masses is rather mild, compared to the size of the uncertainties. In figure 11.4 we exemplarily show the chiral interpolation of R_Z based on the results of the B1 ensembles. From eq. (4.27) and eqs. (10.2 – 10.4) we expect leading effects of $O(a\text{Tr}[M_q])$ when we vary the sea quark mass. Indeed, we observe a linear behavior in the sea quark mass with a small slope for both LCP.

In general, the errors from the heavier line of constant physics are smaller. The differences between both LCP are significant, also for the results at the chiral point. We expect this difference to be an $O(a\Delta)$ effect and we will investigate this in depth in section 11.3.

We collect the results for the R_X at the point of massless sea quarks in tables 11.2 and 11.3. These could now be used for the renormalization and improvement of quark masses at the values of β cited in the tables. Since we intend to use the results on the large-volume CLS ensembles, where slightly different couplings have been used, we proceed to the determination of interpolating curves for the three estimators.

Before we proceed to this stage of the analysis, we want to point out that we could have taken a different path to arrive at the result for the estimators at the point of chiral sea quarks. Since we define the estimators from the fit coefficients m_{11} , D_i and N_i , we can turn our analysis around and perform a chiral interpolation of the single fit coefficients first. In the next step, we can determine the R_X at the chiral point from eqs. (10.16–10.18) and the chirally interpolated fit parameters. In figure 11.5 we show the chiral interpolation of

For LCP-1, the chiral point corresponds to massless sea quarks and heavy valence quarks.

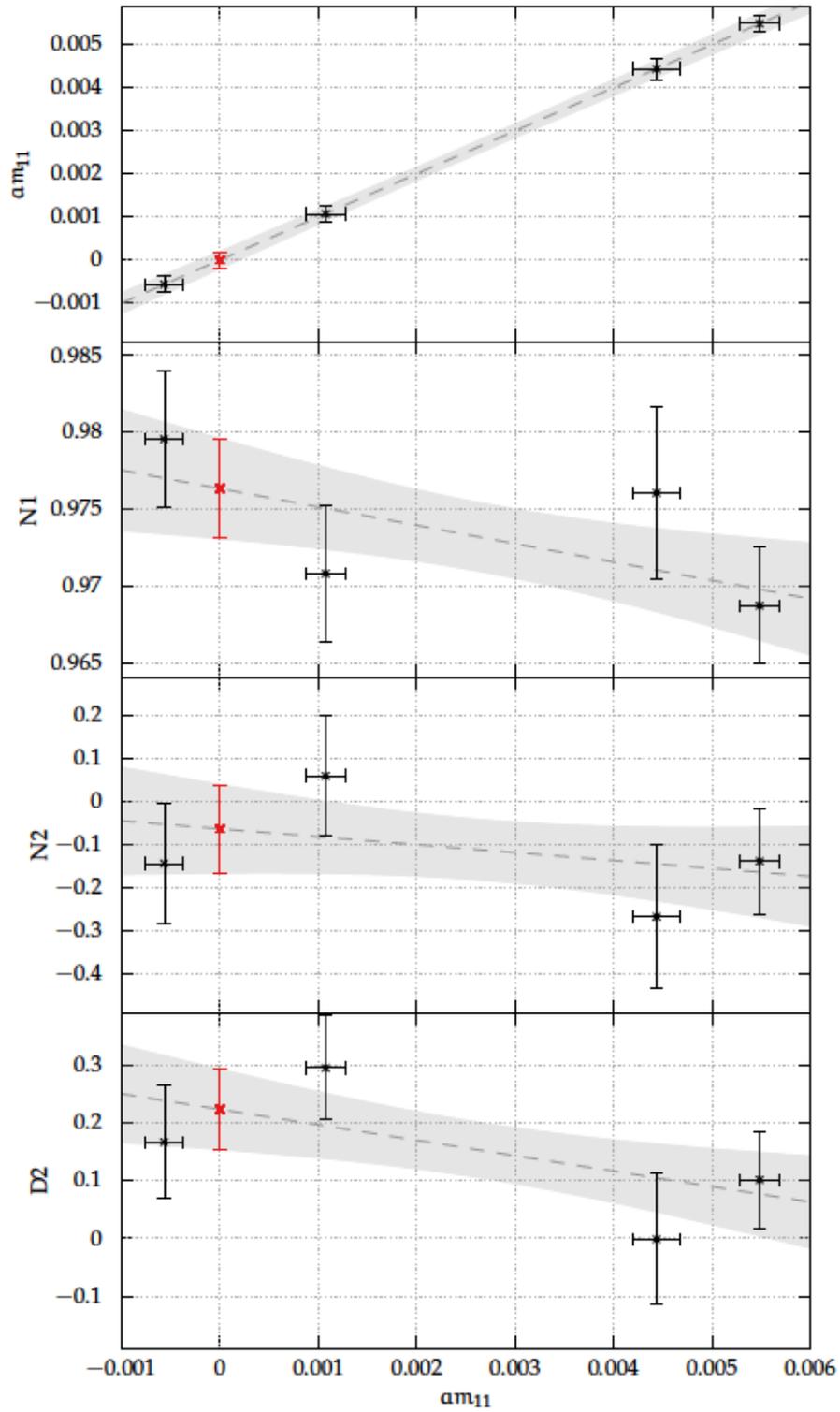


Figure 11.5: Chiral interpolation of the fit parameters am_{11} , N_1 , N_2 and D_2 for the B1 ensemble group. The black crosses show the data from the fits on the single ensembles and the red crosses mark the chiral point.

Table 11.2: Chirally extrapolated LCP-0 results, both for the vanishing topological charge sector, $R_X^{(0)}$, and without zero-charge projection, $R_X^{(\text{all})}$. Only the former are plotted in figure 11.4.

β	$R_{\text{AP}}^{(0)}$	$R_{\text{AP}}^{(\text{all})}$	$R_m^{(0)}$	$R_m^{(\text{all})}$	$R_Z^{(0)}$	$R_Z^{(\text{all})}$
3.300	-0.769(101)	-0.656(55)	1.303(90)	1.244(43)	0.7462(56)	0.7468(28)
3.414	-0.812(53)	-0.770(53)	0.291(53)	0.364(44)	0.8762(40)	0.8719(37)
3.512	-0.515(49)	-0.536(36)	-0.291(39)	-0.177(32)	0.9764(33)	0.9672(26)
3.676	-0.291(46)	-0.279(37)	-0.671(43)	-0.583(35)	1.0588(31)	1.0536(23)
3.810	-0.156(20)	-0.144(17)	-0.738(19)	-0.700(18)	1.0882(11)	1.0866(10)

Table 11.3: Chirally extrapolated LCP-1 results, both for the vanishing topological charge sector, $R_X^{(0)}$, and without zero-charge projection, $R_X^{(\text{all})}$. Only the former are plotted in figure 11.4.

β	$R_{\text{AP}}^{(0)}$	$R_{\text{AP}}^{(\text{all})}$	$R_m^{(0)}$	$R_m^{(\text{all})}$	$R_Z^{(0)}$	$R_Z^{(\text{all})}$
3.300	-0.356(24)	-0.376(10)	-0.025(19)	-0.002(7)	0.7896(36)	0.7846(16)
3.414	-0.362(13)	-0.363(12)	-0.264(12)	-0.237(10)	0.8992(26)	0.8950(24)
3.512	-0.227(12)	-0.244(9)	-0.469(11)	-0.429(9)	0.9861(23)	0.9785(18)
3.676	-0.125(14)	-0.133(12)	-0.643(14)	-0.607(12)	1.0611(23)	1.0564(17)
3.810	-0.070(7)	-0.071(6)	-0.684(7)	-0.669(6)	1.0884(8)	1.0871(8)

the fit parameters entering the estimators R_X for LCP-0. The first interpolation shows the fit parameter $a m_{11}$ versus the measured value for the sea quark masses which we also denote by $a m_{11}$. Since we did not constrain the fit to the measured values, we can check the validity of the fit with the intercept at $a m_{11} = 0$.

From the other three interpolations, we can see that linear interpolations are appropriate in all cases. Since the parameters are based on physical properties of the quark masses, namely the expansion parameters of the renormalized masses in $(a\Delta)$, they do not vary much, when $\text{Tr}[M_q]$ is changed.

Since we decided to use the results of the chirally interpolated estimators instead of the fit parameters to determine the estimators at the chiral point, we do not give all results for the fit parameters in forms of tables. In any case, the results at the chiral point match far beyond the statistical precision. We checked for all five ensemble groups that the behavior of the interpolations are similar to the ones presented in fig. 11.5.

11.2.3 Interpolation in the coupling

The determination of the correct functional form for an interpolation of the results in table 11.2 in the bare coupling g_0^2 is a non-trivial task. For all three estimators, perturbative formulae to one-loop exist [233, 234], but we have no reason to believe that these perturbative predictions hold in the considered coupling range. Therefore, we have to try to find the functional form that best describes the behavior of the R_X . Several fits to different classes of functions are tested to find the optimal parametrizations.

Sometimes, an extensive test of different start parameters is necessary to obtain a decent result.

In general, simple polynomial fits provide a good result and a correct error estimation for the fit parameters, since they are linear in their coefficients. Unfortunately, we have to deal with another complication. As stated above, the finest CLS ensembles have been generated at an inverse coupling of $\beta = 3.85$ which is slightly outside of the considered range. We therefore try to find a functional form that is trustworthy beyond the fit range. Due to the behavior of the data, polynomial interpolations tend to bend strongly for $\beta > 3.810$.

We decide to constrain the interpolating functions to the perturbative predictions towards the limit $g_0^2 \rightarrow 0$. With this choice and after an extensive search for the best ansatz, we describe the functional behavior by the following formulas:

$$R_{AP}(g_0^2) = -0.0010666 g_0^2 \times \{1 + \exp(p_0 + p_1/g_0^2)\}, \quad (11.5a)$$

$$R_m(g_0^2) = -0.5 - 0.0762933 g_0^2 \times \frac{1 + q_0 g_0^2 + q_1 g_0^4}{1 + q_2 g_0^2}, \quad (11.5b)$$

$$R_Z(g_0^2) = 1.0 + 0.0703169 g_0^2 \times \frac{1 + z_0 g_0^2 + z_1 g_0^4}{1 + z_2 g_0^2}. \quad (11.5c)$$

As indicated above, the leading coefficients are set to the values predicted by perturbation theory [233, 234]. We choose the same fit forms for both lines of constant physics since the behavior in the coupling is not fundamentally different between both. For LCP-0, we obtain the fit parameters

$$(p_j) = (16.7457, -19.0475), \quad (11.6a)$$

$$(q_j) = (3.53337, -2.48944, -0.516695), \quad (11.6b)$$

$$(z_j) = (0.703413, -0.769835, -0.478372), \quad (11.6c)$$

with covariance matrices

$$\text{cov}(p_i, p_j) = \begin{pmatrix} 3.49591 & -6.07560 \\ -6.07560 & 10.5834 \end{pmatrix}, \quad (11.7a)$$

$$\text{cov}(q_i, q_j) = \begin{pmatrix} 94.5681 & -57.5056 & 0.859064 \\ -57.5056 & 34.9883 & -0.525367 \\ 0.859064 & -0.525367 & 0.009086 \end{pmatrix} \times 10^{-2}, \quad (11.7b)$$

$$\text{cov}(z_i, z_j) = \begin{pmatrix} 4.22703 & -2.54941 & 0.231607 \\ -2.54941 & 1.537772 & -0.139695 \\ 0.231607 & -0.139695 & 0.013179 \end{pmatrix} \times 10^{-2}. \quad (11.7c)$$

For LCP-1, i.e., the line of constant physics including the partially-quenched heavy valence quarks, we find:

$$(p_j) = (15.6049, -18.4592), \quad (11.8a)$$

$$(q_j) = (2.66968, -1.93055, -0.468542), \quad (11.8b)$$

$$(z_j) = (0.729908, -0.780933, -0.467403), \quad (11.8c)$$

with covariance matrices

$$\text{cov}(p_i, p_j) = \begin{pmatrix} 1.50497 & -2.63930 \\ -2.63930 & 4.63683 \end{pmatrix}, \quad (11.9a)$$

$$\text{cov}(q_i, q_j) = \begin{pmatrix} 74.2042 & -44.4131 & 2.10959 \\ -44.4131 & 26.5860 & -1.26398 \\ 2.10959 & -1.26398 & 0.062898 \end{pmatrix} \times 10^{-2}, \quad (11.9b)$$

$$\text{cov}(z_i, z_j) = \begin{pmatrix} 2.94708 & -1.76762 & 0.182059 \\ -1.76762 & 1.06029 & -0.109193 \\ 0.182059 & -0.109193 & 0.011597 \end{pmatrix} \times 10^{-2}. \quad (11.9c)$$

The resulting curves together with the data points are presented in figure 11.6. The numbers in the covariance matrices are inflated by a factor of two compared to the outcome of the fits. This is done for two reasons. First, by inflating the errors of the interpolating curves, these errors are similar to the errors on the single data points. Since we do not have any insight on the true functional form of the curves, we only want to describe the data in the best possible way.

Another motivation for the inflation of the error is the prediction of the values at $\beta = 3.85$. Since this point is outside the fit range, the effect of choosing different ansätze for the interpolations is enhanced, compared to the other couplings. By inflating the error, we are able to

Table 11.4: Interpolated values of our estimators for couplings employed in CLS simulations along the two renormalised trajectories LCP-0 and LCP-1 considered in this work. Statistical uncertainties are as described in the text and match the confidence band in figure 11.6.

β	LCP-0			LCP-1		
	R_{AP}	R_m	R_Z	R_{AP}	R_m	R_Z
3.85	-0.155(36)	-0.781(38)	1.0975(25)	-0.073(12)	-0.708(15)	1.0971(18)
3.70	-0.258(42)	-0.640(31)	1.0591(23)	-0.119(14)	-0.630(11)	1.0612(17)
3.55	-0.432(46)	-0.358(47)	0.9937(42)	-0.196(14)	-0.498(15)	1.0015(30)
3.46	-0.590(53)	-0.044(65)	0.9320(50)	-0.265(14)	-0.376(17)	0.9468(35)
3.40	-0.726(67)	+0.290(76)	0.8758(52)	-0.324(17)	-0.266(17)	0.8981(35)
3.34	-0.893(95)	+0.810(98)	0.8019(68)	-0.398(23)	-0.124(22)	0.8355(44)

cover the spread of different curves at $\beta = 3.85$ and therefore provide reliable numerical values within errors. We give an overview of the values at the couplings of the CLS simulations [60, 142, 214, 235, 236] obtained from the fits in table 11.4.

As it can be seen from figure 11.6, we achieve a good description of the data. As cross-check, we like to compare our results with existing determinations from the literature. For $b_A - b_P$, values based on the coordinate space method extracted directly on the CLS ensembles are available from [71]. We show our results together with their data points and their interpolation in fig. (11.7). It can be seen that both determinations feature the same qualitative behavior. In [235], two values for Z determined on the CLS ensembles can be found. As it can be seen in the lower part of figure 11.7, these values nicely align with the curve, determined from our data.

Based on our determination of Z , we are able to determine the ratio Z_S/Z_P which is used in quark mass renormalization. Together with the interpolation formula for Z_A [65, 66], the scale independent ratio can be determined via

$$\frac{Z_S}{Z_P} = \frac{1}{Z_S Z_P} = \frac{1}{Z_A Z}. \quad (11.10)$$

A comparison of the values based on R_Z with the results from a dedicated determination of Z_S/Z_P from Ward identities can be found in figure 7 of [237]. There, it can be seen that the differences between both methods scale to zero linear in a^3 towards the continuum limit. Since both determinations have been performed on the same set of ensembles, they are correlated. These correlations have not been taken into account.

In [71], b_A and b_P have been determined separately which is of importance for the improvement of matrix elements.

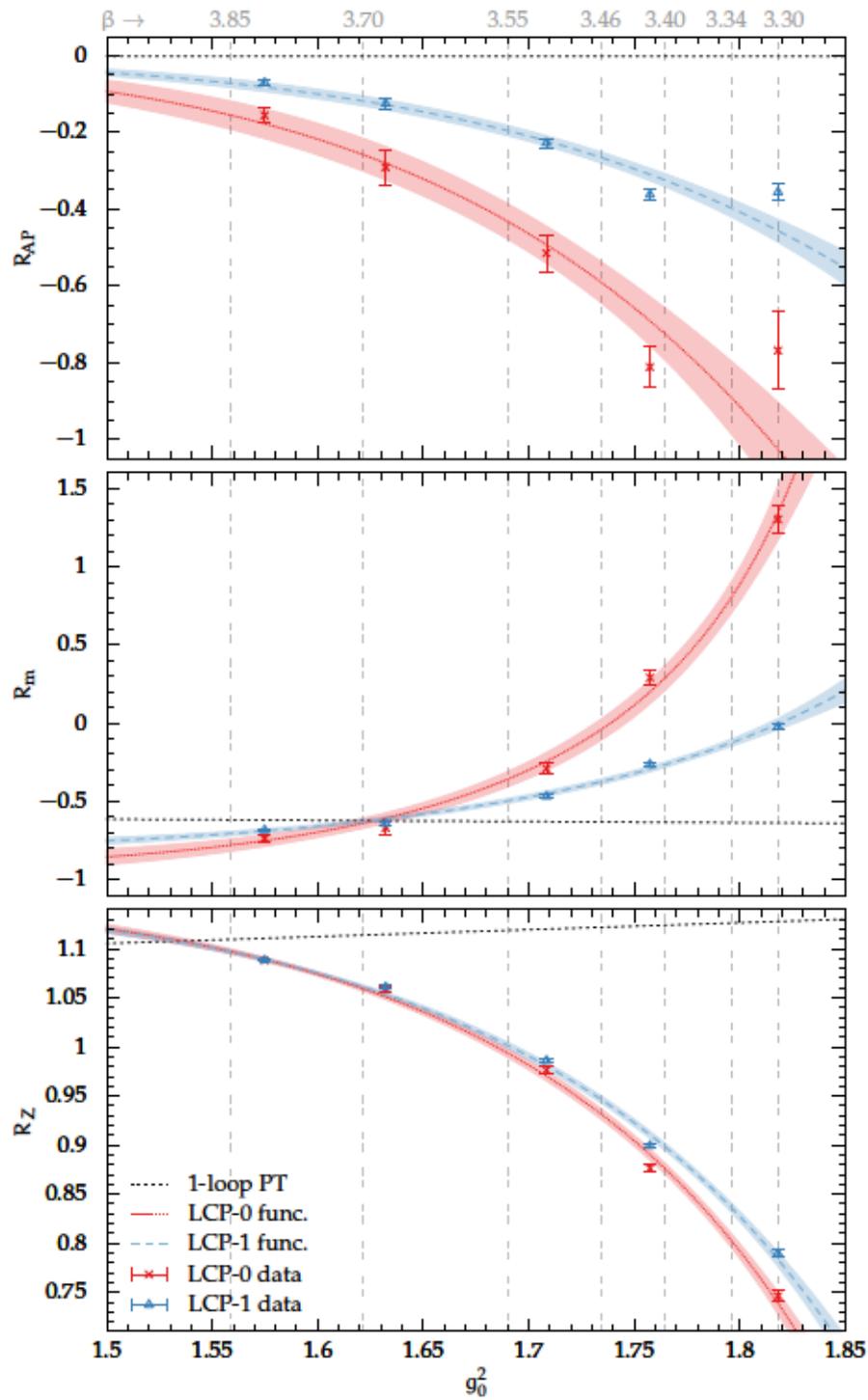


Figure 11.6: Estimators $R_X^{(0)}$ for LCP-0 and LCP-1 and their interpolating functions according to eqs. (11.5). Vertical lines indicate CLS g_0^2 -values, while straight dashed lines represent the one-loop perturbative estimates. Reproduced from [74].

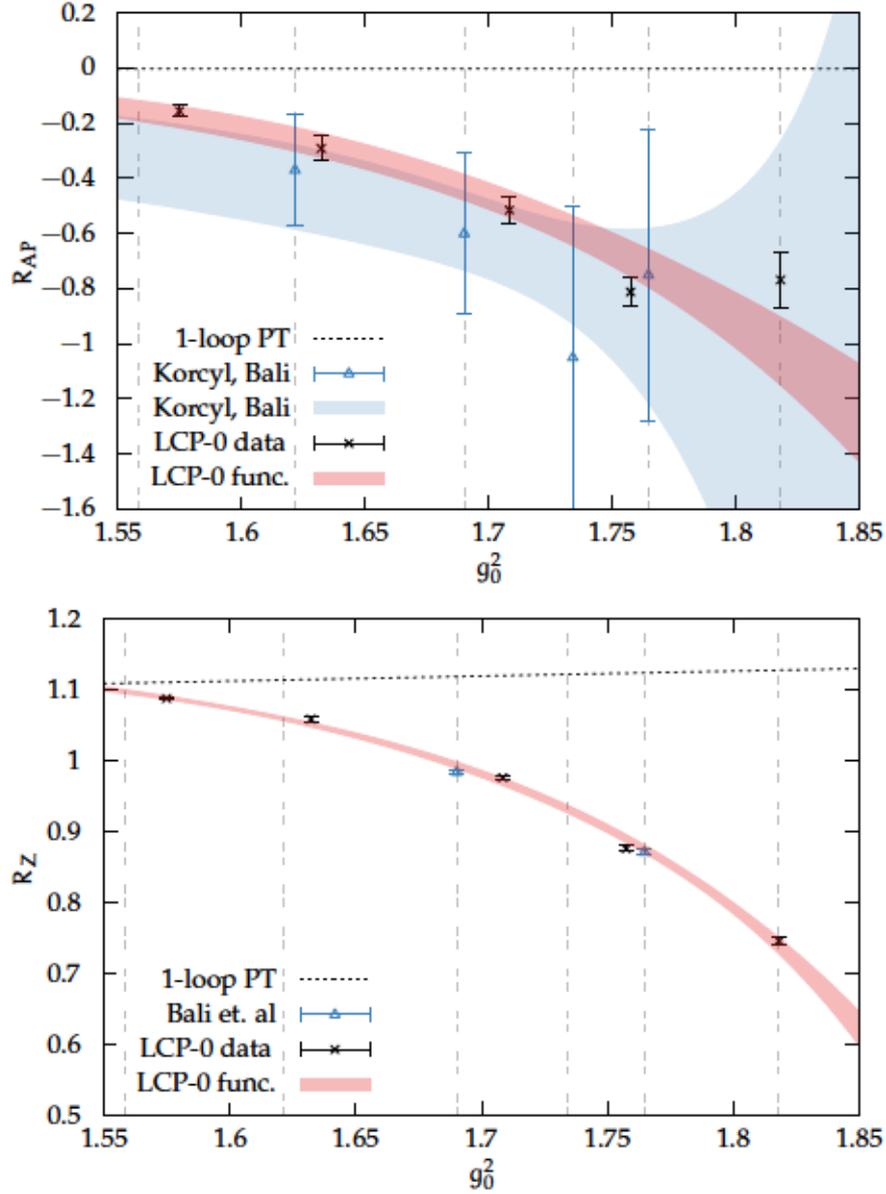


Figure 11.7: *Top:* The chirally extrapolated estimator $R_{AP}^{(0)}$ (in the sector of trivial topology) as a function of the squared coupling, g_0^2 , together with the values and the curve determined by Korcyl and Bali in [71]. Reproduced from [74]. *Bottom:* The estimator R_Z together with the values determined by Bali et. al in [235]. The dashed line shows the one-loop perturbative prediction. The vertical lines indicate the couplings used in the CLS large volume simulations [60, 142, 235].

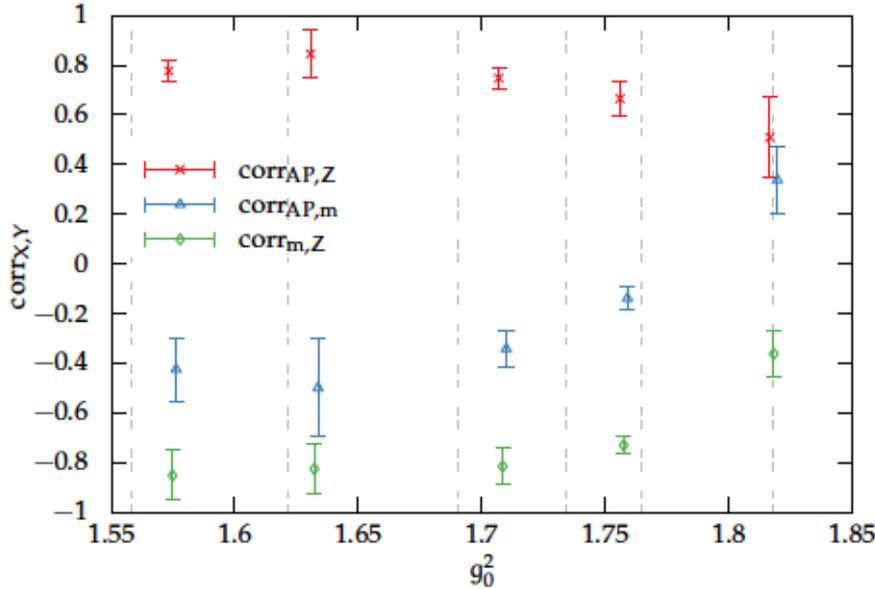


Figure 11.8: Correlations between the estimators R_{AP} , R_m and R_Z for the inverse couplings used in our simulations. The points are shifted horizontally to improve visibility. Vertical dashed lines indicate typical g_0^2 -values of CLS large-scale simulations. Reproduced from [74].

11.2.4 Errors and correlations

Let us further comment on the use of the errors of the different estimators. Since Z renormalizes the quantity of interest, taking into account its error, e.g., in the computation of quark masses, is a necessity. The situation might be different for b_m and $b_A - b_P$. As far as we can judge from the literature, the error on b_X and c_X quantities is not taken into account. For c_{SW} , c_t and \tilde{c}_t which enter as parameters into simulations, the inclusion of an error would mean the generation of ensembles with a slightly varied input parameter. For the b_X parameters which multiply in general $O(a)$ suppressed values, we assume the impact of statistical errors to be fairly small.

More important for the decision to neglect the error may be our LCP setup which we explained above. Many ambiguities which vanish in the continuum limit, are larger than the statistical errors. Choosing a different LCP would result in a different approach to the continuum limit which is equally valid. With our quoted errors, we try to cover all statistical and residual systematic uncertainties within our LCP.

As soon as one chooses to include the errors on all quantities entering a renormalized quark mass, one might be concerned about the correlation of the observables R_Z , R_{AP} and R_m . Since the estimators are based on the same parameters N_i and D_i these correlations are non-negligible. In order to further quantify this, we have investigated the size of these correlations on our ensembles.

Table 11.5: Correlations between the estimators R_{AP} , R_m and R_Z for the β -values used in our simulations after an interpolation to the chiral point. Reproduced from [74].

β	$\text{corr}_{AP,m}$	$\text{corr}_{m,Z}$	$\text{corr}_{AP,Z}$
3.300	0.34(14)	-0.36(10)	0.51(17)
3.414	-0.14(5)	-0.73(4)	0.67(8)
3.512	-0.34(8)	-0.82(8)	0.75(5)
3.676	-0.50(20)	-0.83(11)	0.85(10)
3.810	-0.43(13)	-0.85(11)	0.78(5)

In table 11.5 and figure 11.8 we give estimators for these correlations, i.e.,

$$\text{corr}_{X,Y} = \frac{\text{cov}_{X,Y}}{\sqrt{\text{cov}_{X,X}\text{cov}_{Y,Y}}}, \quad X \neq Y \in \{AP, m, Z\}. \quad (11.11)$$

We evaluated the correlations on the single ensembles and did interpolations to the chiral point for each lattice spacing. Since the correlations only marginally change for different LCP, we choose to quote only the values for LCP-0 in the sector $Q = 0$.

From figure 11.8, we can see that the correlations are significant and relatively large. Nevertheless, even if the values in tab. 11.5 are used to determine the joint error of quantities like $(b_A - b_P)/Z$ or $b_m - (b_A - b_P)$ which are common in applications, we see that the effect of the correlations is way below the effect of the inflation of the statistical errors by a factor of two in the interpolation formula.

Based on these observations, we argue that if one likes to use the errors on the R_X , as they are given by the covariance matrices in eqs. (11.7, 11.9), it is safe to neglect correlations. Since other parameters as c_A and Z_A have been determined on the same ensembles and the correlation with these observables is unknown, this might be the most consistent way.

11.3 AMBIGUITY CHECKS

As discussed above, the comparison of our results with values determined in other studies is reassuring. Nevertheless, we like to investigate the ambiguities that manifest themselves in our approach if the line of constant is chosen differently. Based on eqs. (10.2–10.4) we expect these ambiguities to scale to zero linear in a for R_{AP} and R_m and linear in a^2 for R_Z .

As explained before, we consider all choices that we have made to fix the line of constant physics as part of the definition of the R_X on the lattice. We will inspect the impact of three such choices in this

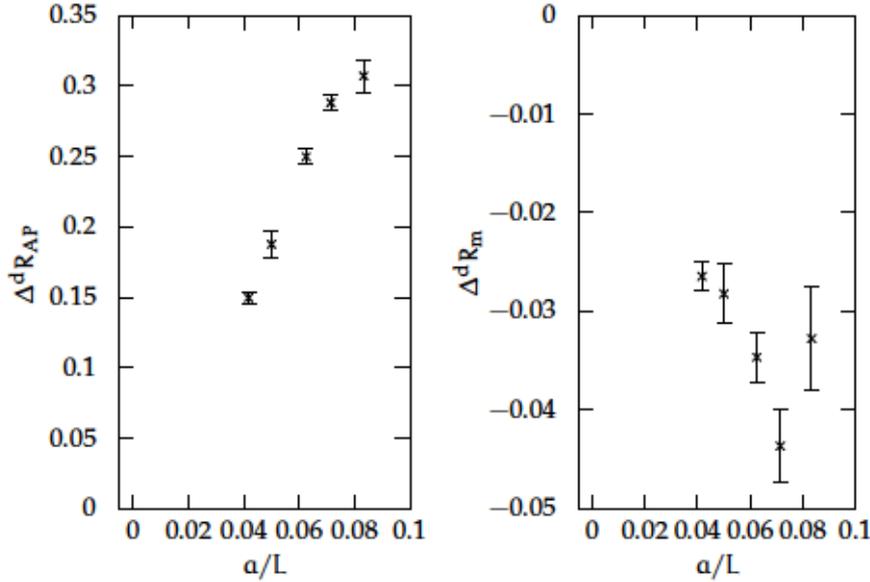


Figure 11.9: $O(a)$ ambiguities of R_{AP} and R_m due to different definitions of the lattice derivative (improved vs. standard). Reproduced from [74].

section. To take the correlations between the determination based on two different LCP into account properly, the differences have been determined within the Γ -method.

We start with the definition of the discretized derivatives that enter the PCAC masses. Since the determination of current quark masses is the core of our analysis, it is important to check the influence of the rather uncommon choice to use the improved derivatives. We define the difference between the estimators that have been determined with improved derivatives (as they enter our final results) and the ones determined with the standard choice by

$$\Delta^d R_X \equiv R_X|_{\text{imp}} - R_X|_{\text{std}} . \quad (11.12)$$

We show the scaling of this difference towards zero for R_{AP} and R_m in figure 11.9. For R_m , the ambiguities are rather small. The value at the coarsest lattice spacing might be influenced by a statistical fluctuation. Since the lattice spacing is rather large for this data point, we also might see higher orders in the cut-off effects.

For R_{AP} , the differences nicely scale, but their absolute values are large, compared to the size of the estimator. Effects of similar size have first been observed in the quenched studies [223, 227] and in the two-flavor study [229]. The estimator R_{AP} seems to be especially sensitive to the choice of the derivative. Since all differences scale to zero and our results match those of [71], we do not see a reason to be concerned.

For the next check we turn to the choice of the degree in the polynomial fits. We decided to use a fit of degree three to achieve a good

The values from standard derivatives are also consistent with the curve from [71].

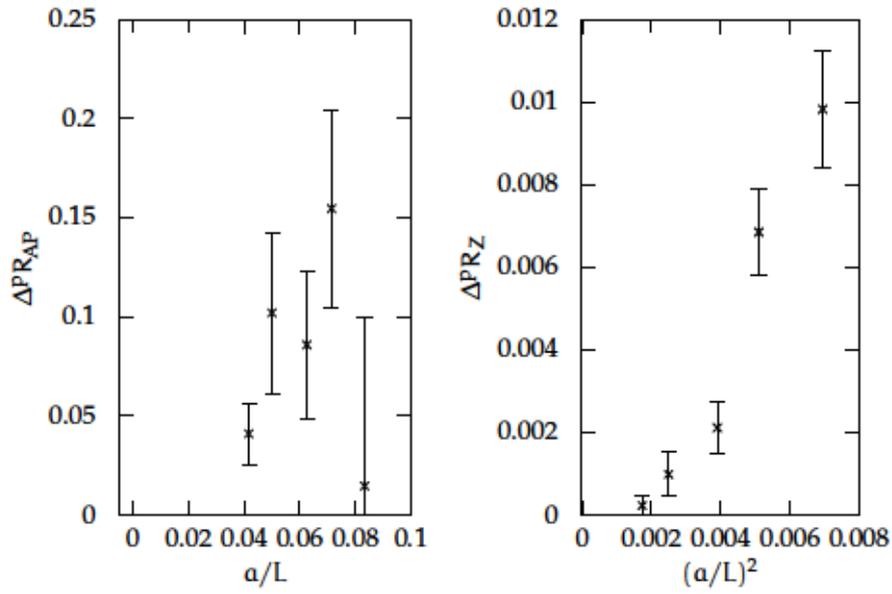


Figure 11.10: $O(a)$ ambiguity for R_{AP} and R_Z , originating from third- vs. fourth-degree polynomial fits. Reproduced from [74].

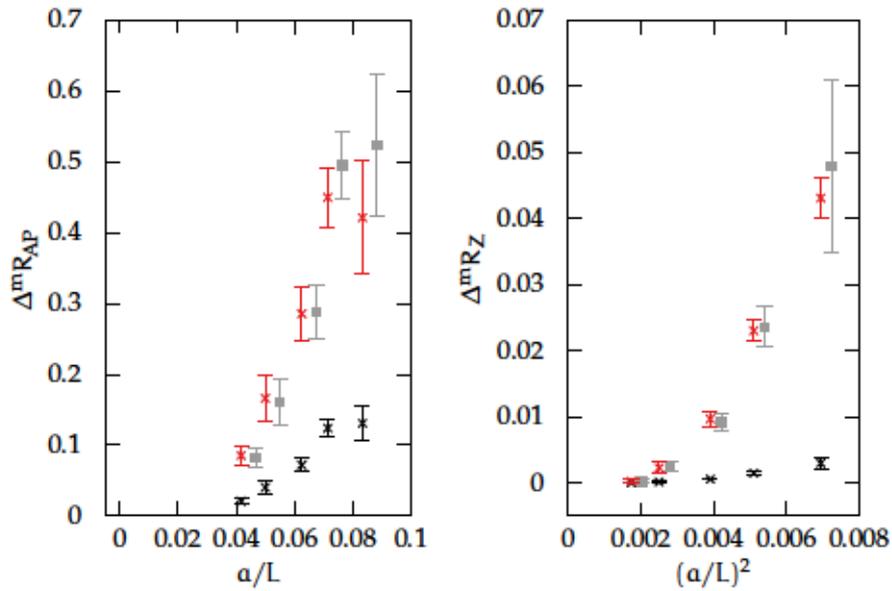


Figure 11.11: $O(a)$ ambiguities for R_{AP} and R_Z between different valence lines of constant physics. Black crosses show the differences between $L\Delta_{22} = 0.25$ and $L\Delta_{22} = 0$, while red crosses do so between $L\Delta_{22} = 1$ and $L\Delta_{22} = 0$. The gray boxes show the differences between $L\Delta_{22} = 0.25$ and $L\Delta_{22} = 0$ scaled by 4 and 4^2 respectively and are slightly displaced horizontally to enhance visibility. Reproduced from [74].

description of the data, as shown in figure 11.3. The inclusion of terms of order $(a\Delta)^4$ should lead to similar results, provided that the physics at low orders is still extracted correctly. In figure 11.10 we show the difference

$$\Delta^{\text{PR}}R_X \equiv R_X^{(\text{deg}=3)} - R_X^{(\text{deg}=4)}, \quad (11.13)$$

for R_{AP} and R_Z . For the latter, the differences seem to scale to zero even faster than the expected rate of $(a/L)^2$. For R_m , the effects vanish linearly. For the coarsest lattice spacing, these effects are zero within errors which might be due to a fluctuation or cut-off effects of higher orders.

As last check, we investigate the ambiguities between LCP with different valence quark masses. Since our polynomial fits cover the range $0 \leq L\Delta_{22} \leq 1$ we could have chosen to fix the quark mass to any value within this range. Determinations of the R_X based on different heavy quark masses are expected to differ by quark mass dependent cut-off effects. To investigate these effects, we look at the difference

$$\Delta^{\text{m}}R_X \equiv R_X|_{L\Delta_{22}>0} - R_X|_{L\Delta_{22}=0} \quad (11.14)$$

between LCP-1 and LCP-0 in figure 11.11. We also add the comparison with a line of constant physics, where we keep $L\Delta_{22} = 0.25$ to visualize the quark mass dependence of the cut-off effects.

As it can be seen for the exemplary chosen estimators R_{AP} and R_Z , the ambiguities seem to vanish faster than the expected rates. As expected, the cut-off effects are enhanced for larger quark mass differences. To allow for a quantitative comparison, we add a third set of points. The grey boxes show the differences between $L\Delta_{22} = 0.25$ and LCP-0 multiplied by 4 in the case of R_{AP} and 4^2 for R_Z . We see that these are nearly equal to the differences between LCP-1 and LCP-0. We therefore have numerical evidence that the cut-off effects scale as expected from eqs. (10.2–10.4).

All three LCP chosen in this comparison are equally valid. In the case of LCP-0, quark mass dependent cut-off effects are absent due to the definition at the unitary, chiral point. The inclusion of mass dependent effects in the estimators as done in LCP-1 might help to cancel higher order effects in the application. This has to be checked numerically in separate computations.

11.4 ESTIMATORS FROM TIME SLICE AVERAGES

As extension to this chapter, we like to take a brief look at the determination of the estimators R_{AP} , R_m and R_Z from time slice averages, as it was done in [227–229]. In figure 11.12 we show the estimators determined from time dependent PCAC masses $m_{ij}(x_0)$, as defined by eqs. (10.2–10.4) for $L\Delta_{22} \approx 0.5$. We show the estimators based on the improved derivatives in black and based on standard derivatives

in gray. The corresponding PCAC mass m_{22} is shown in figure 11.2. Since the Ward identities are valid on every single time slice, we expect plateaus for the estimators, as soon as we are far away from the boundaries. This is approximately the case for all three estimators. For the two time slices in the middle of the plateau, we see a strong fluctuation and for R_m , we see an upward trend. These fluctuations and deviations can be smoothed out by an average in the central region.

As we have already seen it in figure 11.9, R_{AP} is affected rather strong by the choice of the derivative. This is also the case for the determination based on the time slice averages.

The major improvement of our new strategy, when it is compared to the estimators from time slice averages, becomes apparent when we look at the estimators towards the unitary limit. This is done in figure 11.13, where we show the estimators from time slice averages, based on several choices of $L\Delta$. Going to smaller and smaller values of $L\Delta$, the errors of the data points grow more and more until a reliable determination of the estimators becomes impossible.

The reason for this behavior can be found in the definition of the estimators from ratios of differences of quark masses. When these differences become small, statistical fluctuations lead to numerical instabilities. Therefore, the unitary limit cannot be reached with this method.

This is not the case, when the estimators are based on the continuous parametrizations. The red curves in fig. 11.13 show the continuous description of the estimators based on eqs. (10.16–10.18). We can see that the values from both methods nicely coincide at the upper end of the considered range. Towards the unitary point, the estimators from time slice averages begin to deviate from the continuous curves. This happens, before the errors start to grow significantly faster than for the continuous curve. Therefore, numerical inaccuracies might influence the estimators already at non-vanishing quark masses.

11.5 CONCLUSIONS

We have determined the improvement coefficients $b_A - b_P$ and b_m and the renormalization constant $Z = Z_m Z_P / Z_A$ in finite volume simulations in the Schrödinger functional in the strongly coupled region of three-flavor QCD. Together with the results for the improvement and renormalization of the axial current [65–67] and for the renormalization factor of the pseudoscalar density [64] these results allow for the determination of renormalized and $O(a)$ improved quark masses from the large-volume CLS simulations with $2 + 1$ flavors.

Based on a similar work in the theory with two quark flavors [229], we have refined the strategy of the calculation such that results could be obtained in a fully massless setup. We complement those with

results at finite quark mass which include higher order cut-off effects and may lead to a better scaling of heavy quark observables.

Our results have been obtained on a line of constant physics such that a smooth dependence of the bare coupling is provided. We determined interpolation formulas with carefully estimated systematic uncertainties to allow for the use of our results within a range of bare couplings g_0^2 . Our results are in line with the findings based on a coordinate-space renormalization scheme [71, 72] but significantly more precise.

Our results from LCP-0 have been employed in the determination of the light and strange quark masses in $2 + 1$ flavor lattice QCD [77] where mass-dependent $O(a)$ cut-off effects have been found to be small. We will apply the findings of this chapter in the determination of the charm quark mass in the $2 + 1$ flavor QCD in chapter 14 where we face significant discretization effects.

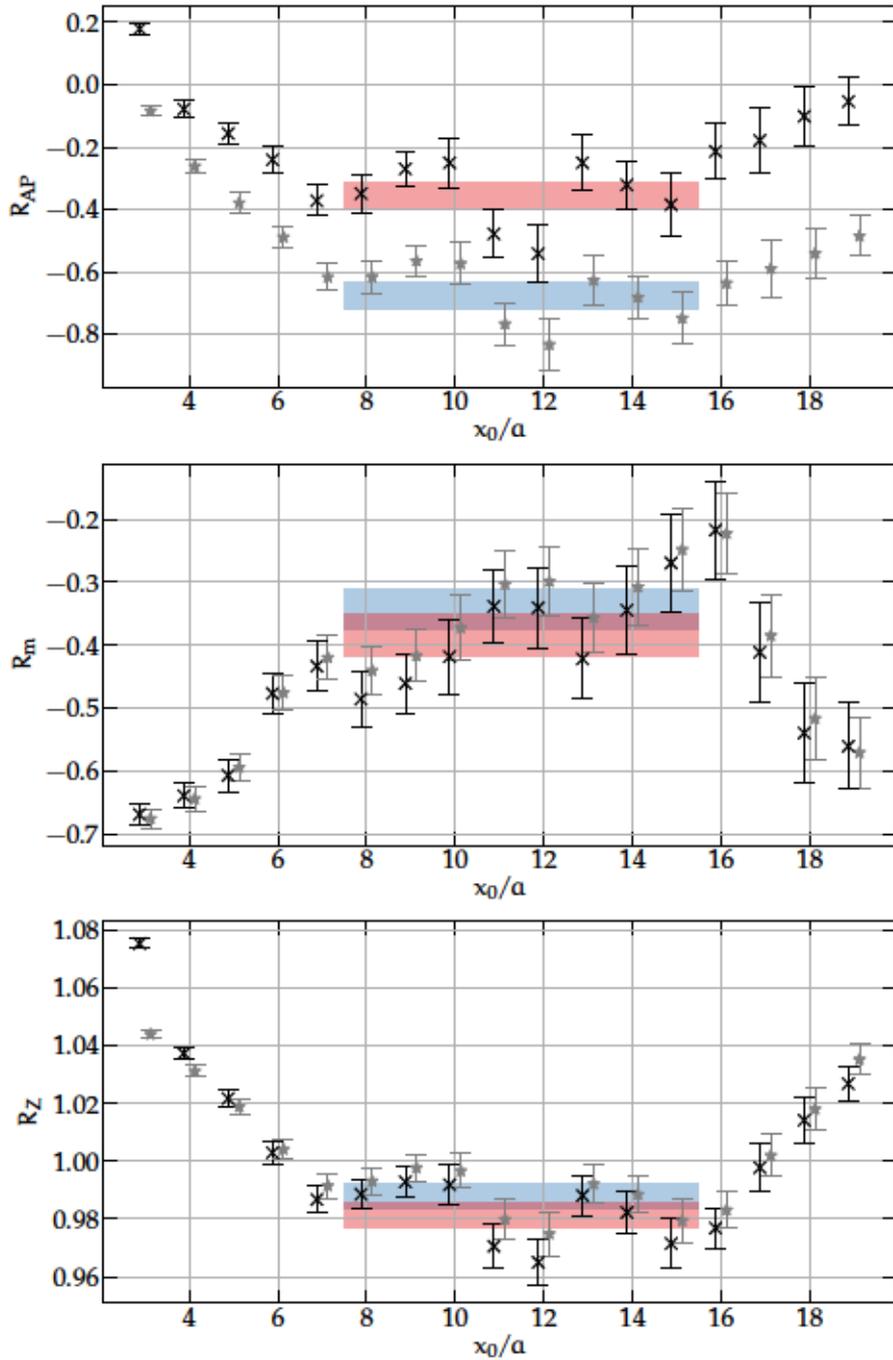


Figure 11.12: The x_0 -dependence of the estimators R_{AP} , R_m and R_Z together with the associated plateau regions for the ensemble B1k4 and valence quarks with hopping parameter $\kappa_2 = 0.13594$ corresponding to $L\Delta_{22} \approx 0.5$. Black crosses and red plateaus show the determination with improved derivatives, gray stars and blue plateaus the results from standard derivatives. Data points are slightly displaced horizontally, to allow for better visibility.

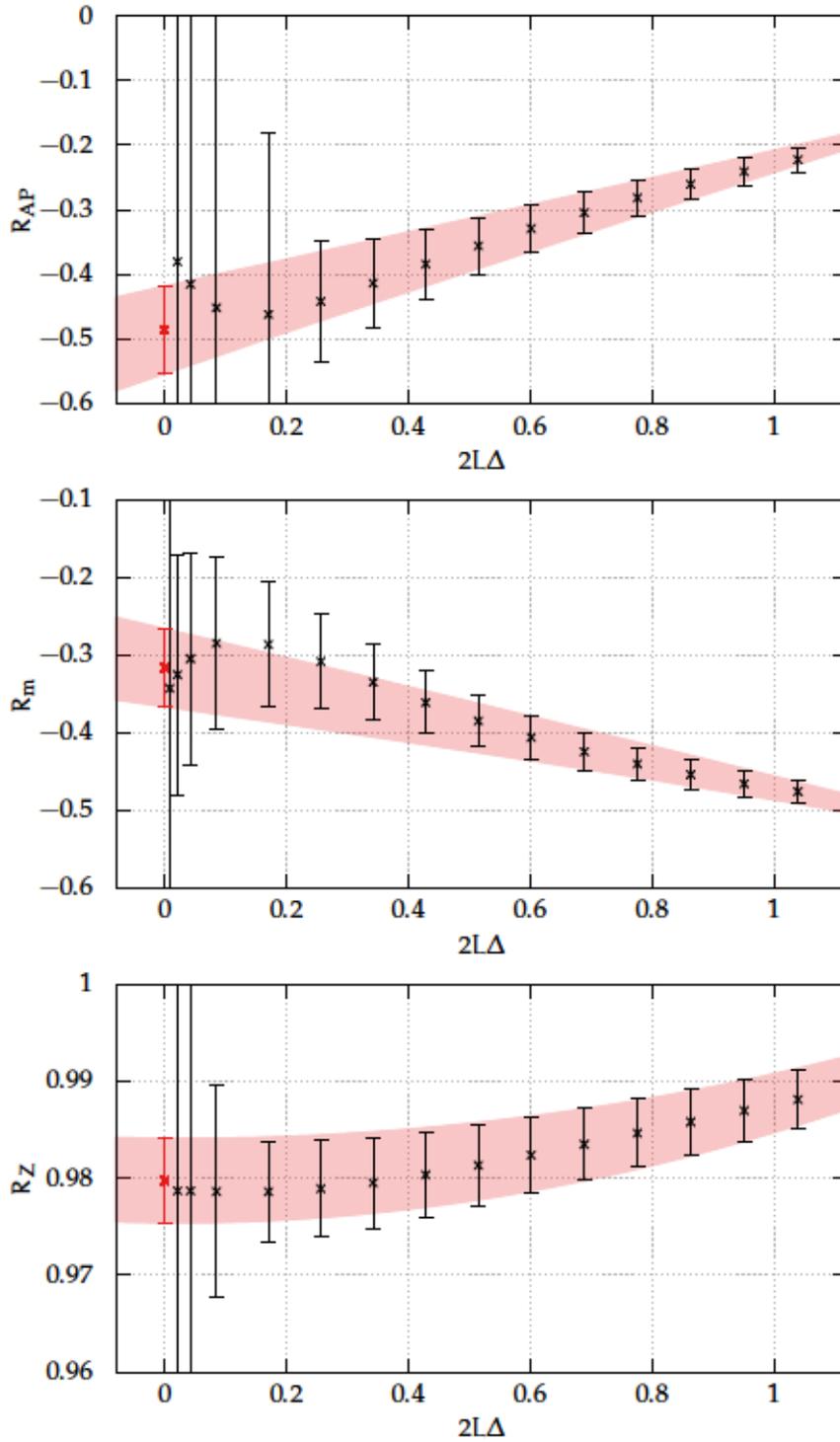


Figure 11.13: Estimators R_{AP} , R_m and R_Z plotted against $2L\Delta$ for B1k4. The black points are obtained from time slice averaged, as described in this section. The red curves show the continuous parametrization according to eqs. (10.16–10.18). The value at the unitary point obtained from the continuous curve is marked by the red cross.

QUARK MASS RENORMALIZATION FOR THE MATCHING OF QCD AND HQET

We now turn to the determination of the improvement coefficients $b_A - b_P$ and b_m and the renormalization constant Z in a more weakly coupled regime. For our matching of HQET to QCD, we need to determine renormalized quark masses in a region around the bottom quark mass.

The line of constant physics for the matching of QCD and HQET consists of a light part which is fixed at the time of the ensemble generation and a heavy part. Since all heavy quarks are treated as partially quenched, we tune to the heavy line of constant physics after the generation of gauge ensembles has been finished. For the determination of the matching parameters we want to fix the renormalization group invariant heavy quark mass to the bottom quark mass,

$$LM_h = LM_b. \quad (12.1)$$

At the same time, we like to determine the exact value of the bottom quark mass from measurements in the effective theory on the large-volume simulations. Therefore, we do not know the exact value of LM_b when we determine the matching parameters.

To allow us to extract the matching parameters satisfying the LCP defined by eq. (12.1) a posteriori, we determine them for several choices of renormalized quark masses which are chosen such that they enclose the physical b quark mass. The matching parameters at the physical b quark mass can then be extracted via an interpolation in $z \equiv LM_h$.

We define the renormalized quark masses which we require to remain constant along the LCP, via the subtracted quark mass, i.e., via eq. (4.17). Therefore, we need to determine all renormalization factors and improvement coefficients that enter the computation of renormalized quark masses from the bare subtracted quark mass. From eq. (4.17) we can then extract the heavy quark hopping parameters κ_h which we use to determine the matching parameters in the full theory.

Since we are well outside the range of couplings considered in chapter 11, we do not believe the interpolation formulas of this section to hold here. Also, since we are in a weakly coupled region, we can expect far more precise results from a dedicated determination in the coupling region of the matching. The same holds for the renormalization constant Z_A , where interpolation formulas for the CLS coupling region are available from [65, 66]. Therefore, we will perform a determination of Z_A along the lines of [65]. The renormalization

constant of the pseudo scalar density Z_P is scale dependent. Hence, we can not use tabulated results and we determine Z_P on our line of constant physics along the lines of [64, 218].

Since our efforts to generate the matching ensembles on a line of constant physics are described in full detail in chapter 8, we just give a short overview of the ensembles that have been employed for our calculations. We perform the measurements on the ensembles generated to match the LCP condition in eq. (8.1) in the volume $L_0 = 0.25$ fm. In this way, we keep the cost moderate and at the same time avoid correlations with the matching parameters which will be determined in $2L_0$.

In the two-flavor project [229], dedicated ensembles with $T/L = 3/2$ have been generated to determine the improvement coefficients. We decided to use the existing ensembles with $T = L$ which have been generated for the tuning towards the line of constant physics. For the smallest lattices, we therefore have to check whether we are able to sufficiently avoid boundary effects which will be present on a significant part of the bulk time slices.

Table 12.1: Ensembles used in this chapter. All ensembles feature $T = L$. N_r is the number of replica runs, while N_{cfg} gives the total number of configurations, separated by τ_{ms} MD units, for which the coupling and Schrödinger functional correlation functions are evaluated to compute the sea PCAC quark mass am_1 . The trajectory length is $\tau = 2$ MDU. All configurations are in the sector of vanishing topological charge.

L_0/a	β	κ	N_r	$\frac{\tau_{\text{ms}}}{\text{MD}}$	N_{cfg}	\bar{g}_{GF}^2	$L_0 m_1$
12	4.3030	0.1359947	7	8	9669	3.9461(41)	-0.00032(36)
16	4.4662	0.1355985	7	10	5887	3.9475(61)	+0.00043(34)
20	4.6017	0.1352848	18	10	8478	3.9493(63)	+0.00100(21)
24	4.7165	0.1350181	3	16	7303	3.9492(64)	+0.00012(17)
32	4.9000	0.1345991	3	20	5014	3.949(11)	+0.00543(34)

We list the ensembles that are used for this study in table 12.1. The five ensembles are simulated on a line of constant physics that is defined by

$$\bar{g}_{\text{GF}}^2 = 3.949, \quad m_1 = 0, \quad (12.2)$$

which translates into a physical extent of $L_0 \approx 0.25$ fm. Since we tuned the sea quark masses to values $L_0 m_1 \lesssim 0.005$, we do not perform a chiral extrapolation but assume to be at the chiral point.

Since the lattice spacing is in a range $0.0078 \text{ fm} < a < 0.021 \text{ fm}$, and we are in finite volume with a very small physical lattice extent, the topological charge is frozen and all configurations are in the sector of

vanishing topological charge. Therefore, we are able to use the full statistics with the implicit projection on the sector of $Q = 0$.

12.1 PCAC MASSES

Since we are interested in the renormalization of heavy quarks, we do not focus on results at the unitary point but cover a range of heavy valence quark masses up to the bottom quark mass. Thus, we perform measurements in the range $0 \lesssim Lm_{q,2} \lesssim 5$. Since we will use the parametrization of quark masses, eqs. (10.13–10.14), for the determination of $b_A - b_P$, b_m and Z we do not have to determine explicit values m_{33} . Instead, we perform measurements to determine m_{12} and m_{22} for eleven different hopping parameters κ_2 such that

$$L(m_{q,2} - m_{q,1}) \in \{0, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0\}, \quad (12.3)$$

where $m_{q,1}$ is the sea quark mass. We obtain ten values of m_{22} and m_{12} together with the PCAC mass at the unitary point αm_{11} .

Since we changed the geometry, compared to the determination in [74], we have to decide for a plateau range which is used to average the current quark masses. After an extensive study of the time dependence of the quark masses for various hopping parameters, we decide to define the plateau region by the central quarter of the lattice, i.e., $x_0 \in [\frac{3}{8}L, \frac{5}{8}L]$. It can be seen for the two exemplary choices in figure 12.1 that dominant boundary effects only form outside this plateau region. As in the former study, we choose improved discretized derivatives for the determination of the PCAC masses.

These masses are used to determine the parameters of the expansions in eqs. (10.13–10.14) via a simultaneous fit. Since we do not tightly constrain the region close to the unitary point by the measured partially-quenched quark masses, we decide to fix the leading term of the parametrizations to the measured value of αm_{11} .

The small statistical errors on the quark masses together with the large fit range leads to an insufficient quality of the parametrization, when the series is truncated after order $(a\Delta)^3$. The inclusion of terms of order $(a\Delta)^4$ leads to a significant improvement with small deviations between data points and interpolating curve. We decide to include terms of $O((a\Delta)^5)$ and obtain an excellent description of the data. When we include terms multiplying $(a\Delta)^6$, the parameter N_6 and D_6 cannot be determined safely anymore, but the quality of the description of the data remains the same.

We show the differences $\Delta^f m_{ij}$ between the measured PCAC masses m_{22} and the interpolating curves based on different degrees of the polynomials on the ensemble with $L_0/a = 12$, where the cut-off effects are most prominent, in fig. 12.2. By construction, the curves coincide with the data point at vanishing quark mass.

We explain shortly, how to determine αm_b on this LCP.

To a very good approximation $\kappa_1 \approx \kappa_{cr}$.

If more data points at small quark masses are included, the quality of the parametrization at large quark masses deteriorates.

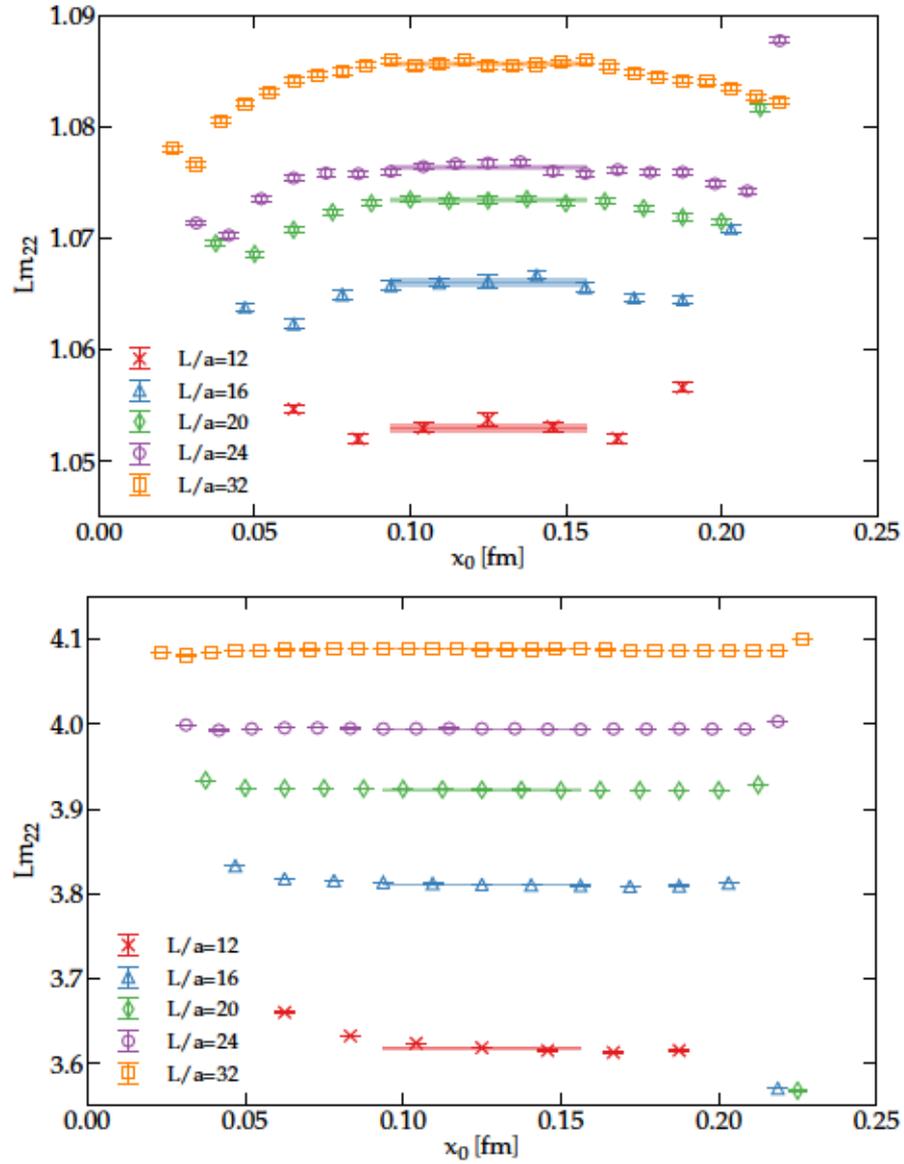


Figure 12.1: Time dependence and plateaus of exemplary PCAC masses for two choices of $a\Delta_q$ for the five considered ensembles.

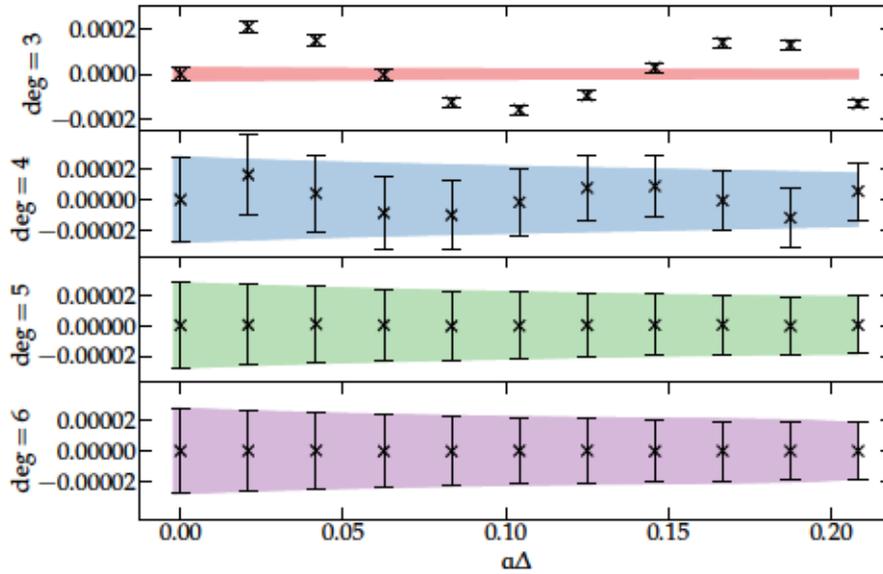


Figure 12.2: Differences $\Delta^f m_{ij} = a m_{ij}(\kappa_2) - a m_{ij}(\Delta)$ between the measured PCAC masses m_{22} and the fitted curves together with the statistical uncertainties of the data points and the uncertainty band of the curves for polynomial interpolations of degrees three to six on the ensemble with $L_0/a = 12$.

The correctness of the fits can be checked by a comparison of the estimators based on the fit coefficients N_i and D_i with the estimators obtained by time slice averages as done in section 11.4. Based on our choice of heavy quark hopping parameters, we are able to determine the estimators from time slice averages for five different heavy quark masses with $Lm_{33} \in \{0.5, 1.0, 1.5, 2.0, 2.5\}$. The determinations of the estimators based on both methods coincide for the chosen degree of the fit.

12.2 ESTIMATORS

Having fixed all parameters of the mass parametrizations, we are able to determine the estimators R_X at any value in the considered range of PCAC masses. Since we did not cover the region close to the unitary point densely and we perform this analysis to renormalize heavy quark masses, we do not determine the R_X at the unitary point.

Instead, we choose a different path and evaluate the estimators at heavy quark masses. This is motivated by a study in [229], where it was found that estimators defined at larger quark masses led to an improved scaling behavior of effective energies defined at heavy quark masses. The inclusion of mass-dependent cut-off effects into the definition of the renormalization procedure therefore improved the precision of the continuum limit. Since we want to optimize the scaling of the matching parameters around the bottom quark mass,

we consider the estimators at heavy quark masses. In the following we will consider different choices of the line of constant physics for the valence quarks.

Since we believe to be able to achieve a significantly better precision than in the more strongly coupled region and we are able to determine the renormalization constants Z_A and Z_P in the same analysis as the estimators R_X , we decide to fix the renormalized quark mass of the valence quark rather than the bare one in the definition of the lines of constant physics for the extraction of the improvement coefficients. Our LCP condition is

$$\text{LCP: } L = \text{const}, \quad Lm_{11} = 0, \quad \frac{Z_A}{Z_P}L\Delta_{22} = \text{const}, \quad (12.4)$$

and we are able to take all correlations in the combination $\frac{Z_A}{Z_P}L\Delta_{22}$ into account. The final value of $\frac{Z_A}{Z_P}L\Delta_{22}$ is not defined a priori and in the following we will carefully investigate the mass dependence of the estimators. We postpone the description of the determination of Z_A and Z_P to section 12.4. An overview of the results is given in table 12.2.

In figure 12.3 we show the continuous description of the three estimators in the considered range of heavy quark masses. Since we chose the same bare subtracted quark masses for all five ensembles, the range of renormalized quark masses differs between the curves. It can be seen that the deviations from the leading linear respectively quadratic behavior of the curves grow with the lattice spacing. This is based on mass-dependent cut-off effects of higher order.

From the smooth descriptions of the estimators R_X , we are able to fix the heavy quark mass in the LCP to any value in the given range and a priori, we do not know which value suppresses best the cut-off effects in the matching observables. An evaluation close to the physical bottom quark mass might be most beneficial. In section 7.3.3, we determined $z_b = L_1 M_b \approx 17.4$ and we can convert this to

$$\frac{Z_A}{Z_P}L_0\Delta_{22} = \frac{L_0}{L_1} \frac{z_b}{h(L_0)} + O(\alpha) \approx 5.9 \quad (12.5)$$

in the volume L_0 , using the running factor $h(L_0) = 1.4744(87)$ determined in eq. (8.44).

The best choice for the LCP, in the sense that most mass-dependent cut-off effects of the matching parameters are subtracted, can not be determined without performing an actual test. In the two-flavor project [229], a scaling test using effective heavy-light energies has been done. Before we decide for the best candidates, we investigate the coupling dependence of the estimators in a range of heavy quark masses.

In figure 12.4 we show the coupling dependence of the estimators R_{AP} , R_m and R_Z determined for different choices of the renormalized heavy quark mass. We decide to cover a range from slightly above the

For the charm quark mass, we determine $\frac{Z_A}{Z_P}L_0\Delta_{22} \approx 1.3$.

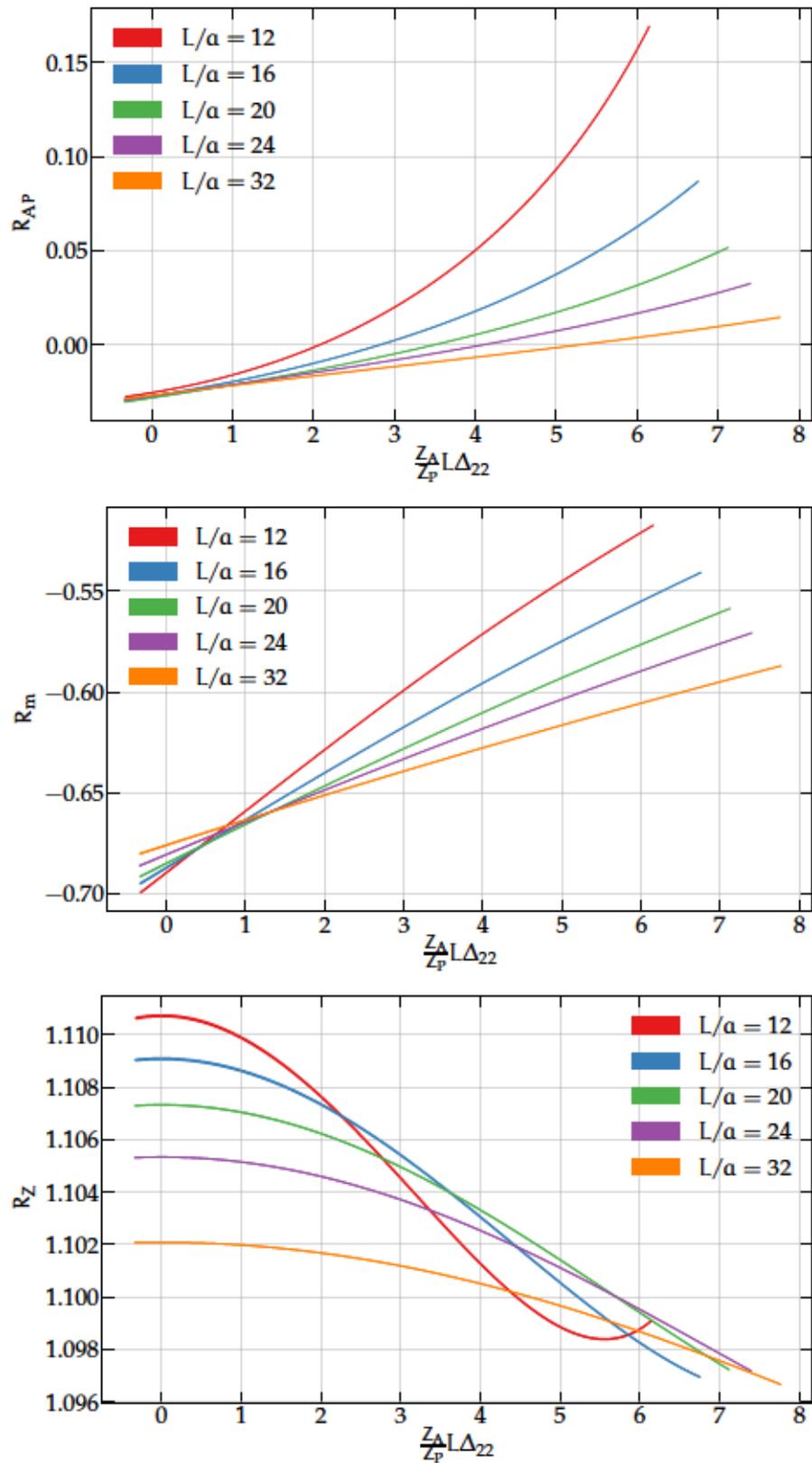


Figure 12.3: Mass dependence of the estimators R_{AP} , R_m and R_Z for all five ensembles along the LCP.

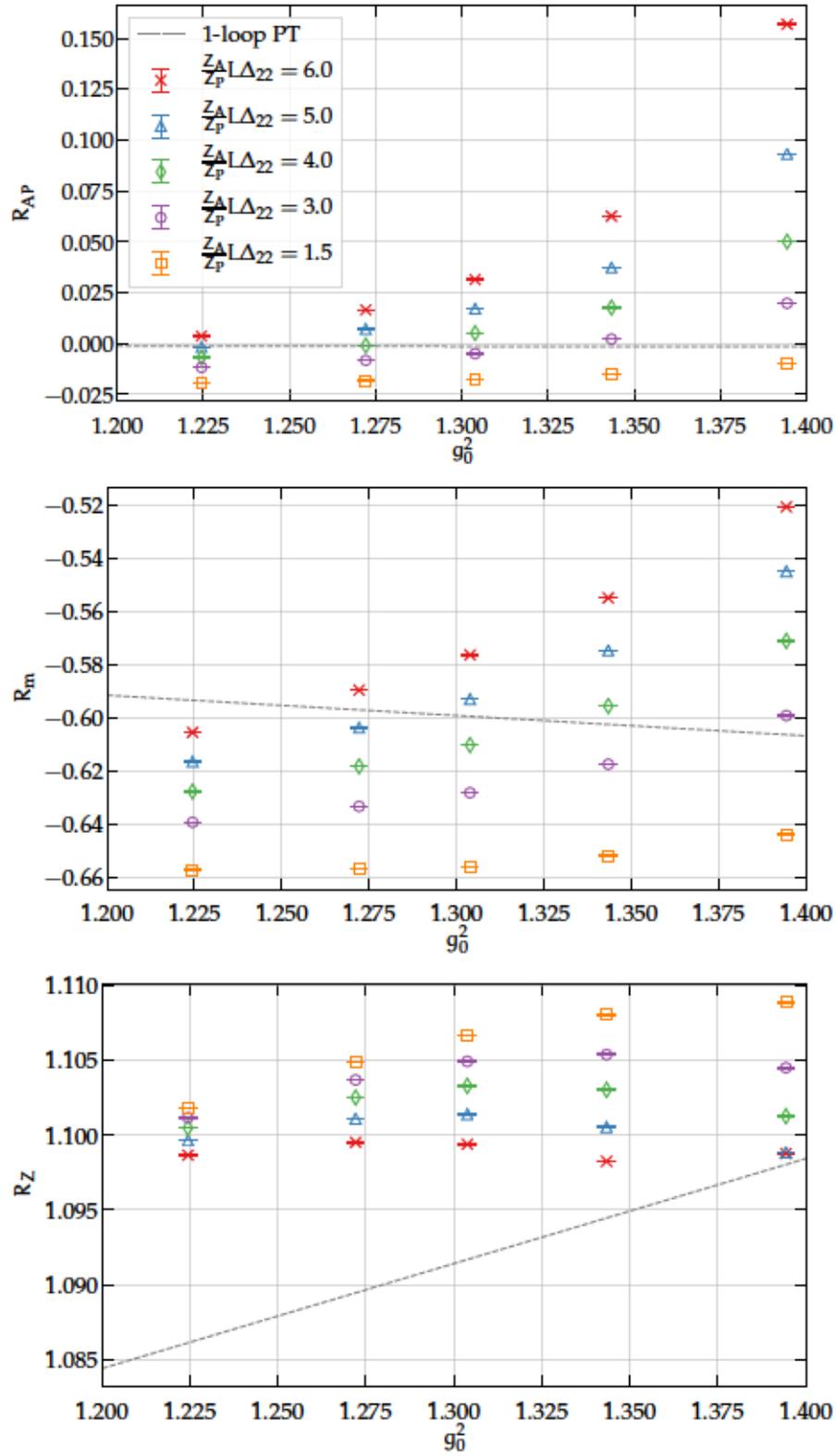


Figure 12.4: Coupling dependence of the estimators R_{AP} , R_m and R_Z for different heavy quark LCP together with the one-loop result from perturbation theory [233, 234].

charm quark mass up to the bottom quark mass. For comparison, we also show the one-loop result from perturbation theory [233, 234].

Compared to the results in the coupling region of the CLS simulations, fig. 11.6, we are able to achieve a remarkable statistical precision. Although we have no reason to believe perturbation theory to hold in this regime, our results fall in the same region as the curves determined from perturbation theory. We are able to achieve a smooth behavior of the estimators towards smaller couplings for most of the different sets of estimators.

However, the determination of R_Z at coarse lattice spacings and large quark masses raises our attention. In the lower plot of fig. 12.4, we can see that the dependence of the estimators evaluated at $\frac{Z_p^\Delta}{Z_p} L_0 \Delta_{22} = 6$ is not as smooth as the others. This is connected to the functional form of R_Z in dependence of the valence quark mass for the coarsest lattice spacing. In fig. 12.3 we see a bending of the curve for $L/a = 12$ at large quark masses. We suspect that higher order cut-off effects become visible for $am_{q,2} > 0.3$ which corresponds to $\frac{Z_p^\Delta}{Z_p} L_0 \Delta_{22} \approx 5$ for $L_0/a = 12$.

Since this is the region where we like to determine the matching parameters, it is possible that the inclusion of higher order effects into the improvement coefficients is beneficial. At the same time we can expect that the mass-dependent cut-off effects of the matching parameters become sizable at this coarse lattice spacing and we might have to exclude it from the continuum limit. From this perspective, it is reassuring that these higher order effects do not manifest themselves for $L_0/a = 16$, where the upper end of the interpolation range is at $am_{q,2} \approx 0.3$.

12.3 AMBIGUITIES

Any redefinition of the line of constant physics is expected to result in ambiguities of order $O(a^n)$ which vanish in the continuum limit. We investigate the size of these ambiguities and their behavior towards the continuum limit to check the choices that we have made in the definition of the LCP. Higher order cut-off effects, as they appear in R_Z at large quark masses, manifest themselves by a deviation from the leading order behavior towards larger values of a .

We will start by the investigation of ambiguities based on choices that have been made at the stage of the analysis, as we have done it in chapter 11. We begin with the investigation of the impact of the choice of derivatives on the final results. For R_{AP} significant ambiguities have been observed when switching from improved to standard derivatives in the quenched and two-flavor studies [223, 227, 229] and also in the coupling region of the CLS ensembles [74]. The same observation can

One set of parameters in fig. 12.4 corresponds to all points on a vertical line in fig. 12.3.

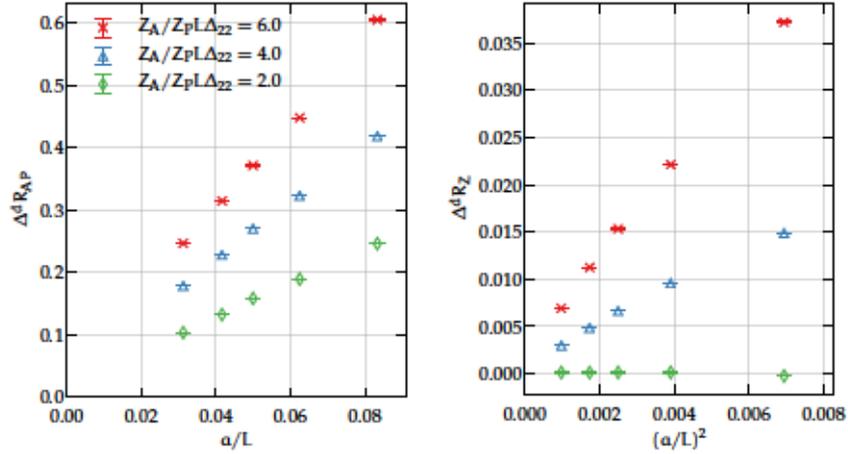


Figure 12.5: $O(a)$ ambiguities of R_{AP} and R_Z due to different definitions of the lattice derivative (improved vs. standard).

be made in the coupling region under investigation. We show the differences

$$\Delta^d R_X \equiv R_X|_{\text{imp}} - R_X|_{\text{std}} . \quad (12.6)$$

in figure 12.5. The differences for R_{AP} are especially large considering that the absolute values of the coefficient are almost vanishing. To get a more complete picture, we show the dependence of R_{AP} determined with standard derivatives in figure 12.6. It can be seen that the functional form differs significantly from the curves in fig. 12.3.

Despite these differences of the functional forms, we observe a linear scaling to zero in fig. 12.5, as we would expect it from ambiguities. Since the determination of R_{AP} is only a byproduct of our project, we now turn to R_Z and R_m . Both are less affected by the choice of the derivatives. We show the ambiguities in the determination of R_Z for three choices of fixed renormalized quark masses in the right panel of fig. 12.5. It can be seen that the size of the cut-off effect increases with the heavy quark mass. All effects scale to zero linear in $(a/L)^2$, as we expect it for this observable from eq. (10.4).

The investigation of the choice of the degree of the polynomial fits in the quark mass parametrization is done going from degree five, the first one describing the data to the required precision, to degree six. The correlated difference between the determinations of the estimators R_X vanishes towards the continuum limit. Since the absolute size of the deviation is significantly smaller than the size of the statistical 1σ errors of the estimators, we do not investigate this ambiguity further.

We have already gained an insight into the mass dependence of the estimators from fig. 12.3 and it has become apparent that the slope of the the functional form decreases when going to finer lattice spacings.

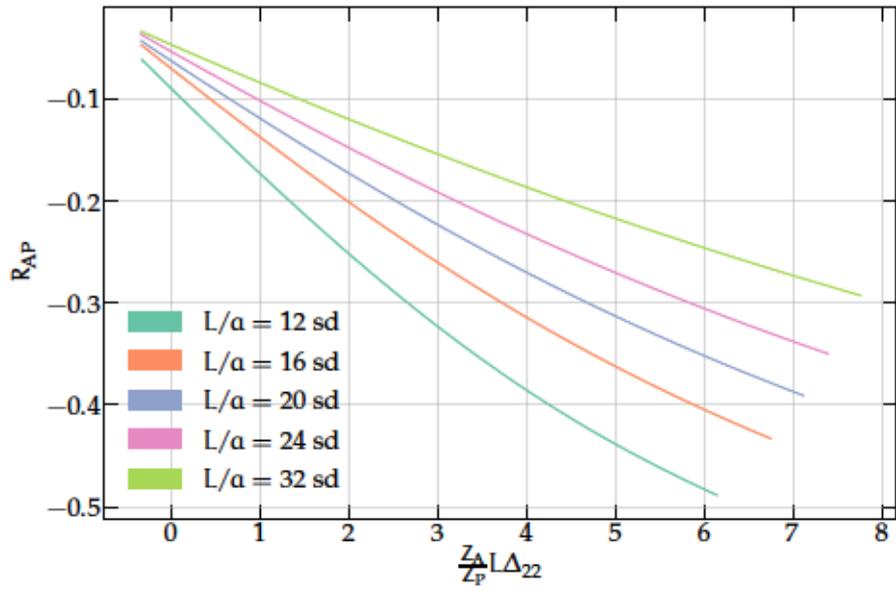


Figure 12.6: Mass dependence of the estimator R_{AP} determined from standard derivatives for all five ensembles.

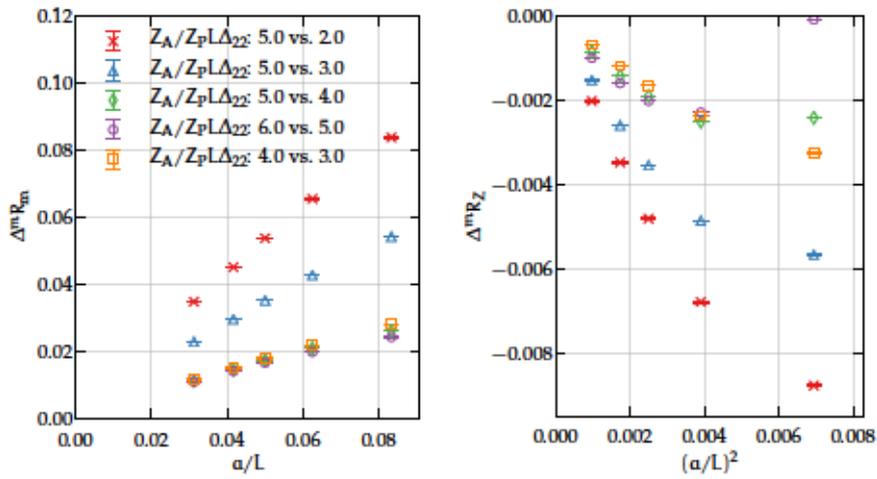


Figure 12.7: $O(a)$ ambiguities for R_{AP} and R_Z between different valence lines of constant physics.

We can investigate this behavior more systematically by looking at the difference

$$\Delta^m R_X \equiv R_X|_{L\Delta_{22}=m_{h_1}} - R_X|_{L\Delta_{22}=m_{h_2}} \quad (12.7)$$

between the determinations of the estimators R_X on two different heavy lines of constant physics.

We show $\Delta^m R_m$ and $\Delta^m R_Z$ for a number of different combinations in fig. 12.7. For $\Delta^m R_m$ we can see that the ambiguities vanish linearly for all LCP combinations. The absolute size of the cut-off effects is dominated by the difference of the two heavy quark masses and not by the absolute size of one of the quark masses. Of course, this only reflects the quasi linear dependence of R_m on the heavy quark mass which is visible from fig. 12.3.

For $\Delta^m R_Z$ we see significant deviations from the leading $O(a^2)$ behavior of the cut-off effects at coarse lattice spacings. These deviations are enhanced if the absolute size of the heavy quark mass is large. Nevertheless, significant deviations appear only at the coarsest lattice spacing and a good scaling is obtained for the other four data points.

Having looked at ambiguities that are based on the definition of the LCP at the stage of the measurements and the analysis, we now turn to an investigation of effects introduced on the stage of the ensemble generation. We can compare the results obtained from the ensembles in table 12.1 with the results determined with the exact same procedure but on a different set of ensembles.

For this task, we have generated three ensembles with $L_0/a \in \{12, 16, 20\}$ and $T = 2L_0$. The gradient flow couplings and sea quark masses determined on these ensembles are listed in tab. 12.4. The difference in the geometry results in a different setup of the line of constant physics and we can investigate the resulting cut-off effects.

We study the difference

$$\Delta R_X \equiv R_X|_{T=L_0} - R_X|_{T=2L_0} \quad (12.8)$$

for all three estimators and different choices for the renormalized heavy quark mass difference $Z_A/Z_P L\Delta_{22}$ in fig. 12.8. Except for the geometry of the ensemble and the resulting changes in the coupling and the sea quark mass, the setup is the same for the determination of the R_X on both ensembles. Due to the increased computational cost, we are limited to the three coarsest resolutions.

We observe the scaling of the ambiguities towards zero for all three estimators. Higher order cut-off effects become visible for larger mass differences in R_m and R_Z . As far as we can judge from three data points, these effects appear to be present only for the coarsest lattice spacing.

We add another check by performing the measurements to determine the estimators R_X on the ensembles of table 8.1, i.e., the ensembles that will be used for the matching with HQET. Due to the

The shift in the gradient flow couplings when going from $T = L$ to $T = 2L$ matches the parametrization given in [219].

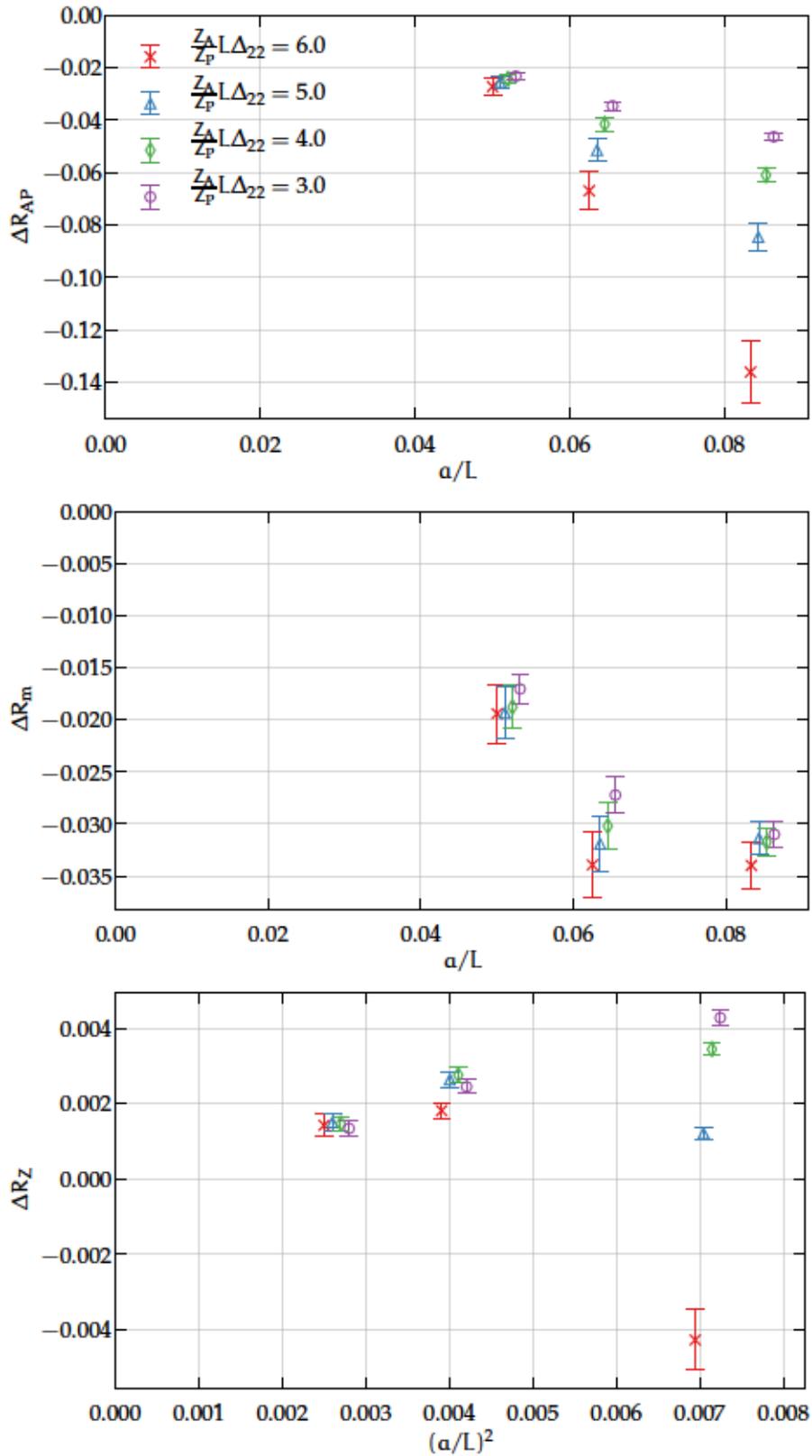


Figure 12.8: Differences between the determination on the ensembles of tab. 12.1 with $T = L_0$ and the ensembles of tab. 12.4 with $T = 2L_0$ for the estimators R_{AP} , R_m and R_Z . The data points are displaced horizontally for better visibility.

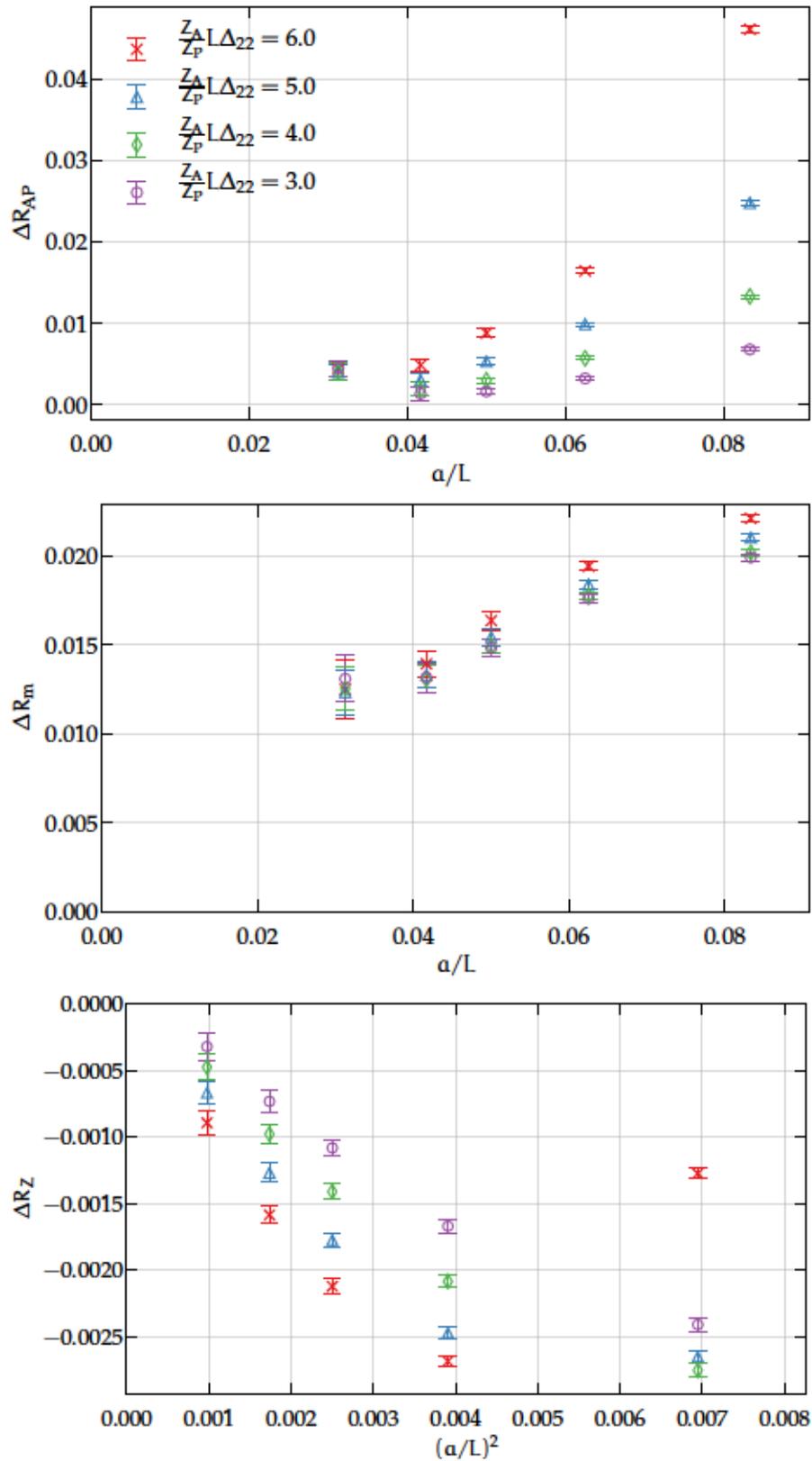


Figure 12.9: Differences between the determination on the ensembles of tab. 12.1 with $L = L_0$ and the ensembles of tab. 8.1 with $L = 2L_0$ for the estimators R_{AP} , R_m and R_Z .

high cost of the inversion on these lattices and since we are only interested in a check of our determination in the volume L_0 , we have not performed the measurements on the full statistics. For the ensembles with $L_1/a \in \{24, 32\}$, we have performed the measurements on all configurations. On the ensemble with $L_1/a = 40$ we have performed measurements on 2500 configurations and on the ensembles with $L_1/a = 48, 64$ we have used 600 configurations.

We show the differences

$$\Delta R_X \equiv R_X|_{L=L_0} - R_X|_{L=2L_0} \quad (12.9)$$

between the determinations on both sets of ensembles in figure 12.9. The addition of two more data points at finer resolution does not change the conclusions drawn in the last paragraph and we see the expected scaling towards zero. For ΔR_{AP} and ΔR_m at the finest lattice spacing, we see a small shift that is independent of the valence quark mass. We suspect that this is an $O(a\text{Tr}[M_q])$ effect. Since $L_1 m_{11}(L_1/a = 64) \approx 0.15$ is significantly larger than the sea quark masses on the other ensembles, the influence of the non-vanishing sea quark mass is amplified for this data point. However, the absolute size of the differences is rather small and we are confident that the estimators determined in the volume L_0 will be well applicable for fixing the heavy quark mass on the matching ensembles.

We will have to expect higher order cut-off effects at the coarsest lattice spacing. In our determination of the improvement coefficients $b_A - b_P$ and b_m and the renormalization constant Z , these are most prominent for R_Z at the coarsest lattice spacing. The qualitative behavior of the mass dependence of R_Z , i.e., the bending of the curve for large quark masses and coarse lattice spacings as depicted in fig. 12.3, is the same for all three sets of ensembles. Therefore, we can conclude that this effect is not induced by the finite extent of the lattice but rather by large values of $a\Delta_{22}$.

The distance between two configurations is increased for $L_1/a = 48, 64$ compared to the other three ensembles.

12.4 RENORMALIZATION OF THE PSEUDOSCALAR DENSITY AND THE AXIAL CURRENT

To perform the renormalization of bare subtracted quark masses, we further need to determine the renormalization constants Z_P and Z_A along the line of constant physics of the matching procedure.

12.4.1 Renormalization of the pseudoscalar density

We determine the scale dependent renormalization constant of the pseudoscalar density $Z_P(g_0^2, \mu = 1/L)$ along the lines of [64, 218]. In the Schrödinger functional, we define Z_P via

$$Z_P(g_0^2, L/a) = c \frac{\sqrt{3f_1}}{f_P(L/2)} \quad (12.10)$$

Table 12.2: Normalization constants Z_A and Z_P determined for the matching of HQET and QCD.

β	L_0/a	$Z_A(g_0^2)$	$Z_P(g_0^2, 1/L)$
4.3030	12	0.831798(50)	0.57835(32)
4.4662	16	0.838819(53)	0.56972(45)
4.6017	20	0.844883(25)	0.56502(53)
4.7165	24	0.849695(26)	0.56002(48)
4.9000	32	0.857020(22)	0.55390(70)

where c is chosen such that $Z_P = 1$ at tree-level. The Schrödinger functional correlation functions f_1 and f_P are defined in section 3.3.3 and we will use time reversal symmetry to include data from g_P . We define the renormalization constant in a massless scheme, i.e., we use the sea quark hopping parameter for the quark propagators.

Numerical values for the constant c in our setup have first been given in [218] and c differs from unity by effects of order $O(a/L)^2$. To be able to determine Z_P on all ensembles used in our work, we have calculated c for a range of values L/a . This determination has been performed using Schrödinger functional gauge fields with gauge links set to unity. On these non-interacting gauge fields, we have used a code based on openQCD to calculate f_P and f_1 and to subsequently determine c via

$$c = \frac{f_P(L/2)}{\sqrt{3}f_1}. \quad (12.11)$$

We list these values of c in table 12.3. No further analysis is required for the determination of Z_P . We show the results of our calculation in table 12.2. We achieve a precision at sub-per mil level which is about the same precision as for bare PCAC masses.

12.4.2 Renormalization of the axial current

We also need to determine the renormalization constant of the non-singlet axial current. This has been done in [65] in the Schrödinger functional with $N_f = 3$ flavors on the ensembles of chapter 11. Since the determination of Z_A has only been a small part of our work, we refer to [51, 65, 238] for all derivations concerning the definition of Z_A which are based on Ward identities.

Table 12.3: Normalization constant c defined by $Z_P = 1$ at tree-level determined from unit gauge fields in the Schrödinger functional with $T = L$, vanishing boundary fields $C = C' = 0$ and $\theta_k = \theta = 0.5$.

L/a	c	L/a	c	L/a	c
6	0.98969662	28	0.99952193	50	0.99985002
8	0.99417664	30	0.99958352	52	0.99986134
10	0.99626479	32	0.99963393	54	0.99987142
12	0.99740297	34	0.99967572	56	0.99988044
14	0.99809059	36	0.99971074	58	0.99988854
16	0.99853742	38	0.99974038	60	0.99989584
18	0.99884400	40	0.99976568	62	0.99990246
20	0.99906343	42	0.99978746	64	0.99990846
22	0.99922584	44	0.99980634	68	0.99991891
24	0.99934941	46	0.99982281	72	0.99992767
26	0.99944559	48	0.99983727	96	0.99995931

The normalization constant Z_A in a massless Schrödinger functional renormalization scheme is defined by

$$Z_A = \lim_{m \rightarrow 0} \left[\frac{f_1}{F_{AA}^I(x_0, y_0) - 2m \cdot \tilde{F}_{PA}^I(x_0, y_0)} \right]^{\frac{1}{2}} \quad (12.12)$$

with the improved correlation functions

$$F_{AA}^I = F_{AA}(x_0, y_0) \quad (12.13)$$

$$+ \alpha c_A [\tilde{\delta}_{x_0} F_{PA}(x_0, y_0) + \tilde{\delta}_{y_0} F_{AP}(x_0, y_0)] \quad (12.14)$$

$$+ \alpha^2 c_A^2 \tilde{\delta}_{x_0} \tilde{\delta}_{y_0} F_{PP}(x_0, y_0), \quad (12.15)$$

$$\tilde{F}_{PA}^I = \alpha \sum_{x'_0=y_0}^{x_0} w(x'_0) [F_{PA}(x_0, y_0) + \alpha c_A \partial_{y_0} F_{PP}(x_0, y_0)] \quad (12.16)$$

with the central difference operator $\tilde{\delta}$, F_{XY} as defined in eq. (3.41) and the implementation of the trapezoidal rule by

$$w(x'_0) = \begin{cases} 1/2 & \text{if } x'_0 = y_0 \text{ or } x'_0 = x_0 \\ 1 & \text{if } y_0 < x'_0 < x_0 \end{cases}. \quad (12.17)$$

The mass m is given by the PCAC mass of the sea quarks. When the Wick contractions are performed it turns out that six diagrams contribute to this definition of Z_A . Among those, there are two disconnected diagrams and it has been shown in appendix A of [238]

As in the case of Z , ambiguities for Z_A are of $O(\alpha^2)$.

that these two diagrams contribute to $O(\alpha^2)$ and vanish in the limit of massless quarks. The inclusion of these diagrams therefore only changes cut-off effects of $O(\alpha^2)$ which is beyond the order we are considering. We do not take these two diagrams into account. This definition of the renormalization constant is labeled Z_A^{con} in [65].

In [65] it has been found that the dependence of Z_A on the quark mass is very mild when the Ward identity eq. (12.12) is used. To maximize the distance between the insertion points and keep their physical distance fixed, we choose $x_0 = \frac{2}{3}T$ and $y_0 = \frac{1}{3}T$. We will check explicitly that we do not face any problems due to small distances of the insertion points on the smallest lattices. In [65] wave functions have been used in the definition of the operators, eq (3.29), to project onto the ground state. We do not use this procedure in our case.

We summarize the results for our determination of Z_A in table 12.2. We see that the statistical errors are of the same order as for Z . In figure 12.10 we show the comparison of our work with the results from [65, 66]. We see that their data have been obtained in a different region of the coupling. The form of the interpolation curves in the regime of our data heavily depends on the chosen fit ansatz. In red we show the curve as it has been provided in appendix C.2 of [66]. If one parameter less was chosen for this polynomial fit which is constrained to perturbation theory, it would overlap with the blue curve from [65].

From this observation we draw the conclusion that both interpolation curves are not valid for our coupling range, unless a systematic error is added. The spread of all possible curves obtained from the data in the CLS region would give rise to a significant systematic error of Z_A in our coupling region. In any case, the errors on the interpolating curves are already rather large.

In contrast, the statistical error on our data points is too small to see it on the chosen scale of the figure. The dedicated determination of Z_A therefore leads to a drastic reduction of the error. We do not have to determine an interpolation formula, since our data points have been determined at the values of the couplings which are used in the matching procedure. It remains to check that no systematic error has been introduced by the small distance (in units of a) between the insertion points and the boundaries. For this task we have generated three ensembles of sizes $L_0 \in \{12, 16, 20\}$ with the geometry $T = 2L_0$. We give an overview of these ensembles together with the measured sea quark masses and the values for the renormalization constant Z_A in table 12.4.

In figure 12.11 we show the difference

$$\Delta Z_A = Z_A(T = L) - Z_A(T = 2L) \quad (12.18)$$

together with a linear fit to the functional form $\Delta Z_A = l_0 + l_1(a/L)^2$. It can be seen that the already small differences vanish with the expected

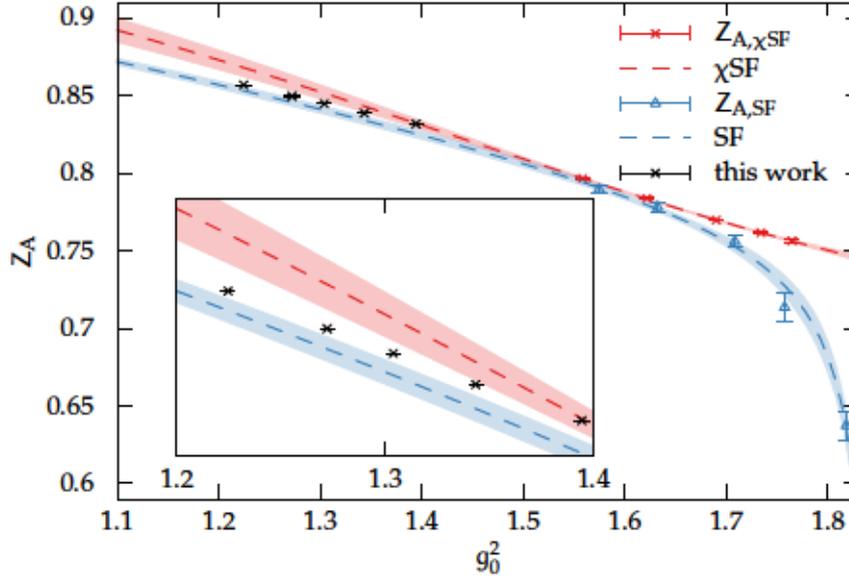


Figure 12.10: Results from our work together with the data and interpolation formulae from [65, 66]. In red, we show $Z_{A,\text{sub}}^l$ from the L_2 -LCP of [66] (χSF) together with the fit from appendix C.2 of the same reference. In blue we show the data points from [65] (SF) together with the corresponding interpolation curve given in eq. (4.1). The region around our data points is magnified in the inset.

rate towards the continuum limit. We conclude that the small time extent of the lattices in tab. 12.1 does not pose a problem and we do not have to generate rather expensive ensembles with $L = 24, 32$ and an enlarged time extent.

12.5 HOPPING PARAMETERS FOR CONSTANT RENORMALIZED QUARK MASSES

The goal of our computations in this chapter is the determination of different sets of hopping parameters such that the renormalization group invariant heavy quark mass remains constant towards the continuum limit. From the definition of the renormalization group invariant quark mass from bare subtracted quark masses based on eqs. (4.17, 8.40),

$$z = L_1 M_h = L_1 h(L_0) \frac{Z Z_A}{Z_P(L_0)} (1 + b_m a m_{q,h}) a m_{q,h}, \quad (12.19)$$

we can determine the heavy quark hopping parameter κ_h via

$$\kappa_h(z, L_1) = \left[\frac{1}{\kappa_{\text{cr}}} - \frac{1}{b_m} \left(1 - \sqrt{1 + z \frac{4b_m Z_P(L_0) a}{L_1 h(L_0) Z Z_A}} \right) \right]^{-1}. \quad (12.20)$$

Table 12.4: Ensembles created for an explicit check of the ambiguities in the determinations of Z_A , Z , b_m and $b_A - b_P$ and their sea quark mass and gradient flow coupling together with the results for the renormalization constant Z_A .

L_0/a	T/a	β	κ	N_r	$\frac{\tau_{\text{ms}}}{\text{MD}}$	N_{cfg}	\bar{g}_{GF}^2	$L_0 m_1$	Z_A
12	24	4.3030	0.1359947	5	8	5321	4.2024(76)	+0.00676(46)	0.83393(30)
16	32	4.4662	0.1355985	2	10	3461	4.2251(99)	+0.00473(42)	0.83926(46)
20	40	4.6017	0.1352848	10	10	2000	4.214(15)	+0.00499(43)	0.84571(33)

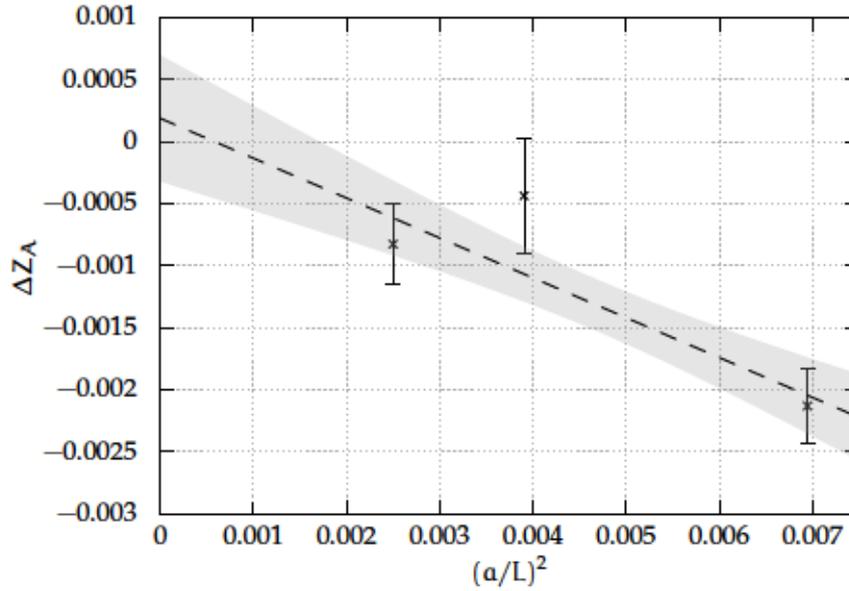


Figure 12.11: Difference ΔZ_A as defined in eq. (12.18). The uncertainty band of the two-parameter fit to the data is depicted by the gray region.

Assuming $\kappa_1 = \kappa_{\text{cr}}$ for our ensembles which are close to chiral point, we have all ingredients to determine κ_h for the five ensembles used in our matching procedure.

The final results for the heavy hopping parameters depend on our exact definition of the line of constant physics and the choice of the valence quark mass in the determination of b_m and Z has a significant impact. In figure 12.12 we show the functional form of $\kappa_h(z, L_1/a = 24)$ for different sets of parameters.

For small quark masses there is barely any difference between the curves obtained from different sets of parameters. Going to the region of the b quark mass and beyond, we can see significant differences. The curves stop at the point where the square root in eq. (12.20) becomes imaginary. We can see that this point is shifted towards larger RGI

We expect $z_b \approx 17.4$.

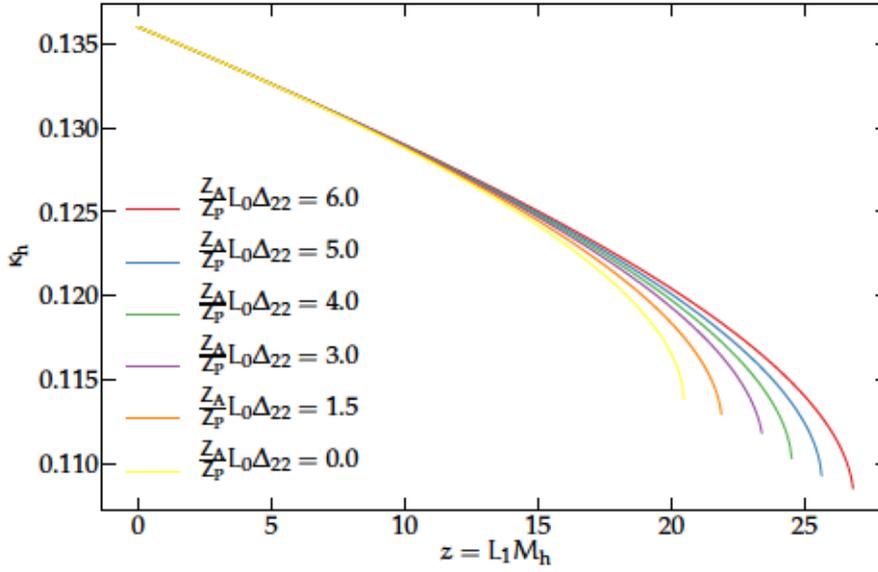


Figure 12.12: Hopping parameters $\kappa_h(z, L_1/a)$ for fixed renormalized quark masses $z = L_1 M_h$ determined from eq. (12.20) with Z_A and Z_P from tab. 12.2 and Z and b_m from different heavy LCP on the coarsest ensemble with $L_1/a = 24$.

masses if larger quark masses have been used in the determination of b_m and Z .

In the study of the two-flavor theory [215] a range of RGI masses up to $z = 21$ has been used. The RGI value of the bottom quark mass was $z = 13.25(22)(13)$ [239] which is smaller than in our study due to a smaller matching volume. In [215] four lattices spacings have been used for the continuum extrapolation of the matching parameters and the determination of κ_h has not been possible for $z > 18$ for the coarsest lattice spacing.

To cover the same range of RGI quark masses as in [215], we have to determine κ_h in a range up to $z \approx 27$. From figure 12.12 it is visible that a determination of κ_h at this value of z is not possible at the coarsest lattice spacing regardless of the set of parameters b_m, Z . Nevertheless, it seems appropriate to choose a relatively large value of $\frac{Z_A}{Z_P} L \Delta_{22}$ to maximize the range of RGI quark masses where the determination of z is possible.

For all other resolutions, the determination of κ_h is possible in the full range of RGI quark masses and the difference between hopping parameters from different sets of b_m, Z vanishes with the lattice spacing. In comparison to the study in [215], where four resolutions have been used, we have set up the matching ensembles with five different resolutions. Therefore, the exclusion of the coarsest ensemble does not affect the validity of the continuum extrapolation.

In table 12.5 we list the values of $\kappa_h(z, L/a)$ determined for a range of RGI quark masses and all five resolutions using Z and b_m evaluated at

Table 12.5: Hopping parameters for fixed renormalized quark masses $z = L_1 M_h$ determined from eq. (12.20) with Z_A and Z_P from tab. 12.2 and Z and b_m evaluated at $\frac{Z_A}{Z_P} L \Delta_{22} = 5$.

z	$\kappa_h(z, 24)$	$\kappa_h(z, 32)$	$\kappa_h(z, 40)$	$\kappa_h(z, 48)$	$\kappa_h(z, 64)$
0	0.1359947	0.1356071	0.1352848	0.1350181	0.1345991
6	0.1318899(24)	0.1326433(24)	0.1329742(23)	0.1331341(17)	0.1332264(18)
9	0.1297062(37)	0.1310919(37)	0.1317762(35)	0.1321636(26)	0.1325248(27)
11	0.1281830(47)	0.1300251(46)	0.1309586(43)	0.1315043(32)	0.1320509(33)
13	0.1265911(58)	0.1289277(56)	0.1301238(52)	0.1308343(38)	0.1315716(40)
15	0.1249114(70)	0.1277943(67)	0.1292698(61)	0.1301525(45)	0.1310866(46)
17	0.1231159(85)	0.1266185(78)	0.1283939(71)	0.1294577(51)	0.1305954(53)
19	0.1211585(103)	0.1253913(91)	0.1274932(81)	0.1287485(58)	0.1300977(60)
21	0.1189541(129)	0.1241007(105)	0.1265638(92)	0.1280233(66)	0.1295930(67)
23	0.1163070(175)	0.1227294(122)	0.1256011(105)	0.1272803(74)	0.1290808(74)
25	0.1124223(342)	0.1212513(143)	0.1245989(118)	0.1265171(82)	0.1285604(82)
27		0.1196226(171)	0.1235490(134)	0.1257311(91)	0.1280313(90)

$\frac{Z_A}{Z_P} L \Delta_{22} = 5$. Since we have determined Z_A , Z_P , Z and b_m on the same set of ensembles, we are able to perform a fully correlated analysis without having to rely on tabulated values. The only external input is the running factor $h(L_0)$ and its error is only taken into account after the continuum extrapolation has been performed. Therefore, it has not been propagated into the uncertainties in tab. 12.5.

12.6 VERIFICATION OF IMPROVEMENT

To verify the cancellation of $O(a)$ effects by the results of section 12.2 and to judge on the reduction of mass-dependent cut-off effects by the determination of b_m and Z at heavy valence quark masses, we have to perform an explicit scaling test. In [228, 229] this has been done using effective heavy-light meson energies. There, the hopping parameter of the heavy quark has been fixed using two different sets of parameters b_m , Z . In [229] the investigation of renormalized current quark masses has been added as test for the universality of the continuum limit in the $O(a)$ improved theory.

We perform both of these tests with data that has been determined on the ensembles of tab. 8.1, i.e., the matching ensembles on the QCD side. We start with an investigation of the scaling of renormalized quark masses and use the quark mass parametrizations to determine

PCAC masses at given hopping parameters. As explained in section 4.3.4, we expect renormalized quark masses determined from bare subtracted quark masses to differ from renormalized current quark masses by $O(a^2)$ effects if both are properly renormalized and improved.

We proceed as follows. We determine the hopping parameters $\kappa_h(z, L/a)$ for fixed values of z and different sets b_m, Z . From the parametrization of the current quark masses, eqs. (10.13–10.14), we are able to determine the quark masses am_{12} and am_{22} for the different values of κ_h without performing dedicated measurements. From am_{12} and am_{22} we determine renormalized quark masses via

$$z_{hh} = h(L_0)Z_m [1 + (b_A - b_P)am_{q,h}] L_1 m_{22}, \quad (12.21)$$

$$z_{hl} = 2h(L_0)Z_m \left[1 + (b_A - b_P) \frac{1}{2}(am_{q,h} + am_{q,l}) \right] L_1 m_{12} - z_{ll}, \quad (12.22)$$

for the ensembles with $L_1/a \in \{24, 32, 40, 48, 64\}$.

Since we are free to choose the value of z , we have a wide variety of tests at our disposal. Further investigations can be performed by adding another set of renormalized quark masses where the dependence on the critical hopping parameter has been removed (4.31),

$$\tilde{z}_{hh} = h(L_0)Z_m [1 + (\tilde{b}_A - \tilde{b}_P)am_{22}] L_1 m_{22}, \quad (12.23)$$

$$\tilde{z}_{hl} = 2h(L_0)Z_m [1 + (\tilde{b}_A - \tilde{b}_P)am_{12}] L_1 m_{12} - \tilde{z}_{ll}. \quad (12.24)$$

In the determination of the PCAC masses, we are free to choose improved or standard derivatives. Where we define the RGI quark masses z_{hi} from standard derivatives, we also use these derivatives in the determination of b_m, Z .

Based on the concept of $O(a)$ improvement, we expect the renormalized current quark masses to differ from the target value z by cut-off effects of $O(a^2)$. The extrapolation of the renormalized current quark masses in a^2 therefore has to match z . When performing this test, we experience a systematic effect due to the non-vanishing sea quark masses on the $2L_0$ ensembles. Since the light quark hopping parameter is not exactly the critical one, we effectively determine the hopping parameters κ_h for the difference of the subtracted quark masses $m_{q,h} - m_{q,l}$. We adjust the renormalized current quark masses accordingly via

$$z_{ij} \rightarrow z_{ij} - z_{ll} \quad (12.25)$$

in the remainder of the section. This shift is of $O(0.03)$ in the renormalized quark masses. Therefore it is not significant at the scale of the bottom quark mass. Nevertheless, we will use its size to estimate systematic effects in the matching observables due to the non-vanishing sea quark mass.

From all tests we hope to get information on the reduction of mass-dependent cut-off effects based on different sets of improvement coefficients at different RGI quark masses z . In figure 12.13 we show an exemplary investigation of the scaling of renormalized quark masses at the scale of the bottom quark mass, i.e., $z = 17.4$. We show the results for the renormalized quark masses together with an unconstrained fit to the data points at the three finest lattice spacings.

We choose three different sets b_m, Z to investigate the influence of light and heavy quark masses in the setup of the LCP and to see whether the cut-off effects observed for R_Z at large quark masses have an influence on the scaling. In figure 12.13 we study three different definitions of renormalized quark masses from current quark masses.

In the upper part of the plot, we show the renormalized quark mass determined from a heavy-heavy correlation function with improved derivatives. The determination from standard derivatives is shown in the lower part. In the middle we show the determination from heavy-light correlation functions. It can be seen that, as one could expect it, the determination using improved derivatives generally leads to smaller cut-off effects. Nevertheless, in all three cases we see a good scaling to the continuum limit.

To judge on higher order cut-off effects, we show the result of two-parameter fits linear in $(a/L)^2$ to the three points at $L_1/a = 40, 48, 64$. In all cases under investigation, we see a deviation from this leading behavior for the data point at the coarsest lattice spacing. The data point at $L_1/a = 32$ seems to show enhanced deviations from the leading order behavior for the set of improvement coefficients determined at $\frac{Z_A}{Z_P} L\Delta_{22} = 1.5$. In the case of the renormalized quark mass from heavy-heavy correlation functions and improved derivatives, the sign of the cut-off effects of higher order changes when increasing $\frac{Z_A}{Z_P} L\Delta_{22}$ in the determination of the improvement coefficients. To allow for a better comparison of different renormalization procedures, we show three variants of PCAC masses determined from improved derivatives in figure 12.14.

Since the scale of the figures at the bottom quark mass hides deviations from the leading behavior, we also show the renormalized current quark mass at the scale of the charm quark in figure 12.15. Also in this case we see that the deviation from the leading order behavior sets in at $L_1/a = 32$ but using the three points close to the continuum, a clean extrapolation is obtained.

Despite deviations from a linear behavior in (a/L) for coarser lattice spacings, all sets of improvement coefficients seem to effectively cancel effects of $O(a)$. Therefore, we have successfully checked that $O(a)$ improvement works as expected using our results. The reduction of mass-dependent higher order cut-off effects differs among different sets of parameters but also among different sets of observables. Therefore it is hard to judge on the quality of different sets.

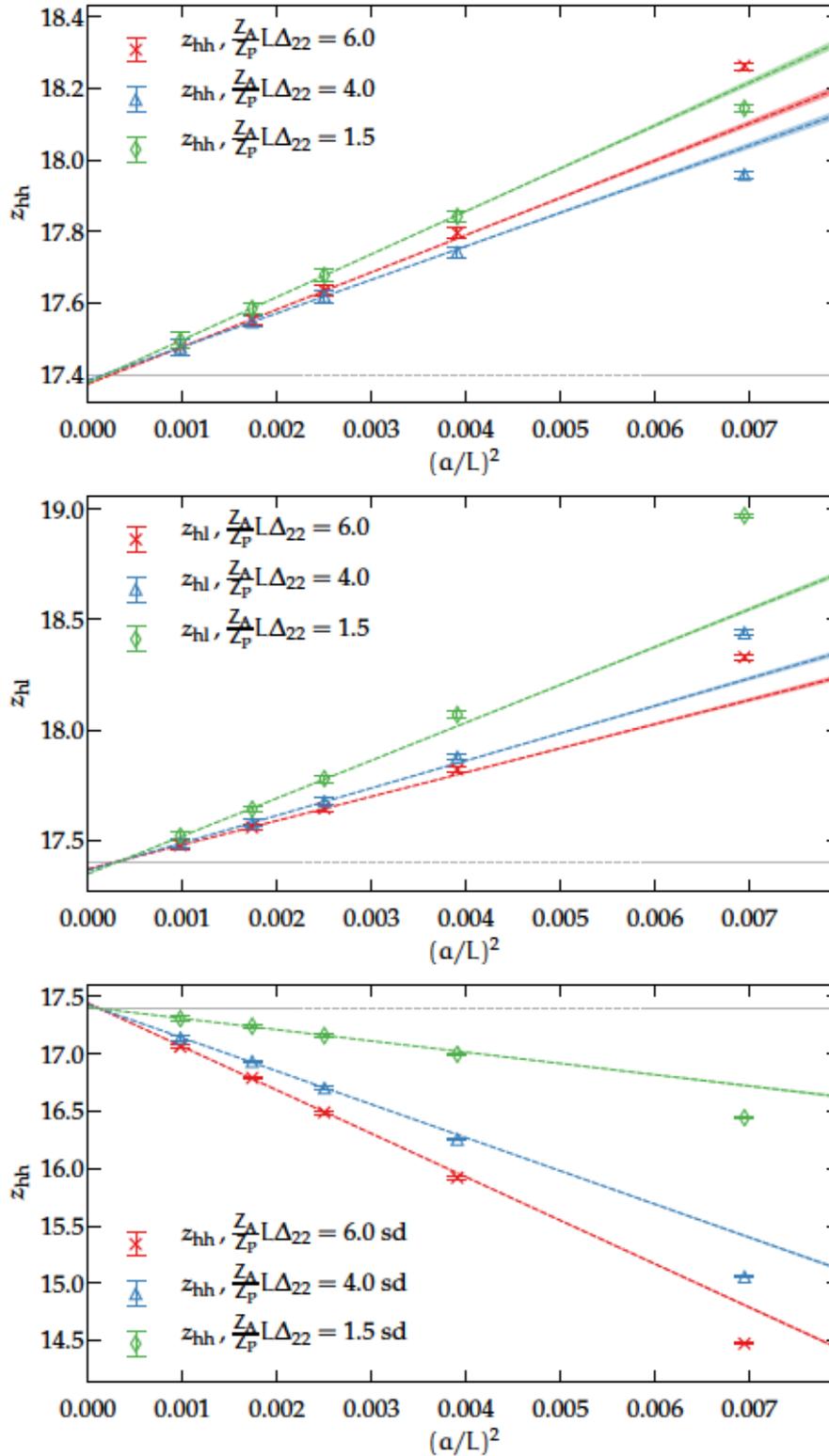


Figure 12.13: Scaling of PCAC masses evaluated at $\kappa_h(17.4, L/a)$ for different LCP. Dashed colored lines represent the unconstrained fit and the dashed grey line denotes the target z , close to the bottom quark mass. *Top:* Heavy-heavy PCAC mass. *Middle:* Heavy-light PCAC mass. *Bottom:* Heavy-heavy PCAC mass with standard derivatives.

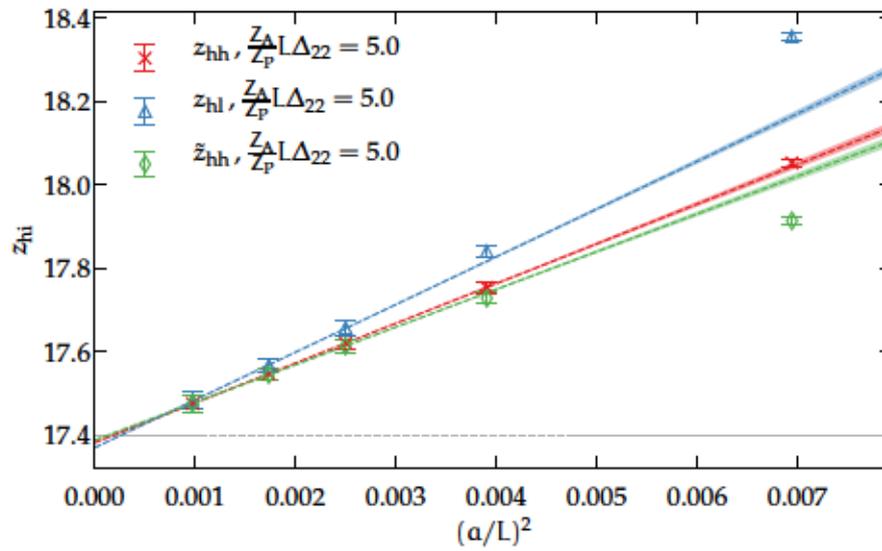


Figure 12.14: Scaling of PCAC masses on ensembles of tab. 8.1 evaluated at $\kappa_h(17.4, L/a)$ for different definitions of the renormalized PCAC mass. Dashed colored lines represent the unconstrained fit and the dashed grey line denotes the target z , close to the bottom quark mass.

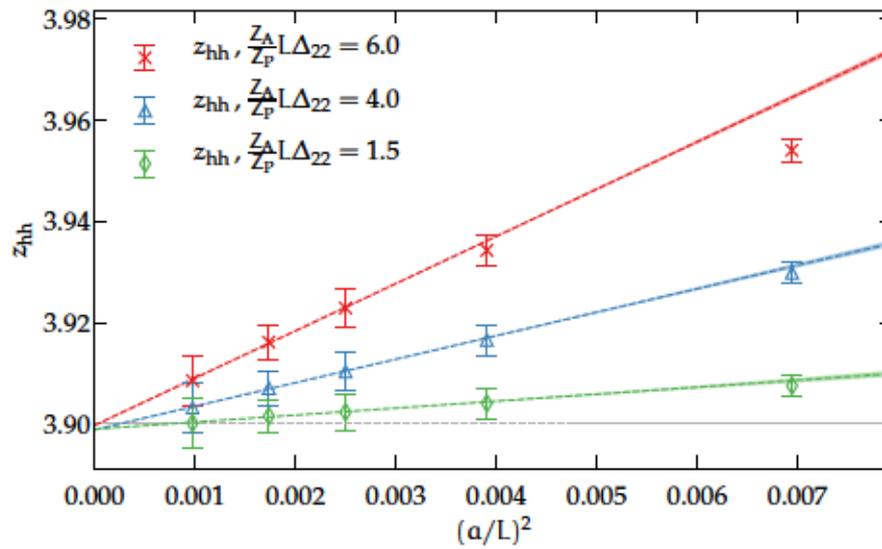


Figure 12.15: Scaling of PCAC masses on ensembles of tab. 8.1 evaluated at $\kappa_h(3.9, L/a)$ for different LCP. Dashed colored lines represent the unconstrained fit and the dashed grey line denotes the target z , close to the charm quark mass.

If we compare the upper plot of fig. 12.13, where the scale of the bottom quark is shown, with fig. 12.15, where the charm quark mass is shown, we see that the slope of the continuum extrapolations is smaller when the set of improvement coefficients has been evaluated at the same scale. For the the scale of the bottom quark mass, this corresponds to $Z_A/Z_P \approx 5.7$ and for the charm quark to $Z_A/Z_P \approx 1.3$. This observation supports the argument to use improvement coefficients evaluated at the scale of the bottom quark to optimize the scaling of the matching parameters at the same scale.

We perform another scaling test using the effective energy

$$L_1 \Gamma_{\text{PS}}^{\text{rs}}(z, g_0) = -L_1 \delta_0 \ln \left[f_A^{\text{rs}, I}(x_0, z) \right] \Big|_{x_0=L/2}, \quad (12.26)$$

from the heavy-heavy and the heavy-light current. All multiplicative renormalization factors cancel in this quantity. The same set of measurements on the $2L_0$ ensembles as above is used. To allow for the determination of Γ_{PS} at given values of z , we interpolate the energy in the quark mass difference $a\Delta$. We expect the square of the energy to scale with the quark mass and obtain an excellent parameterization of the data using the fit ansatz

$$(L_1 \Gamma_{\text{PS}})^2(a\Delta) = c_0 + c_1 a\Delta + c_2 (a\Delta)^2 + c_3 (a\Delta)^3. \quad (12.27)$$

Based on the parameterization of the effective energy, we are able to evaluate $L_1 \Gamma_{\text{PS}}(a\Delta)$ at given values of κ_h determined from eq. (12.20) just as we have done it for the PCAC masses. Afterwards, we can investigate the scaling of the effective energy at various mass scales.

In fig. 12.16 we show the scaling of the effective energy from the heavy-heavy and the heavy-light current at the scale of the b quark for different sets of improvement coefficients. In all cases we observe a scaling in $(a/L)^2$ and can conclude that the leading order cut-off effects are canceled. The size of higher order cut-off effects depends on the chosen set of improvement coefficients.

At the scale of the bottom quark, where we will determine the matching parameters, the cut-off effects are best suppressed using the set of improvement coefficients determined at the same scale. Together with the data points we show continuum extrapolations of the energies. The number of point that can be included in a fit in a^2 depends on the chosen set of improvement coefficients. The largest fit ranges and therefore the most precise results of the continuum limit are achieved using the set of improvement coefficients evaluated at $Z_A/Z_P L_0 \Delta_{22} = 6$. In general, all observations made for the renormalized current quark masses concerning the cancellation of higher order cut-off effects hold for the effective pseudoscalar energies.

The symmetric derivative is defined in app. A.

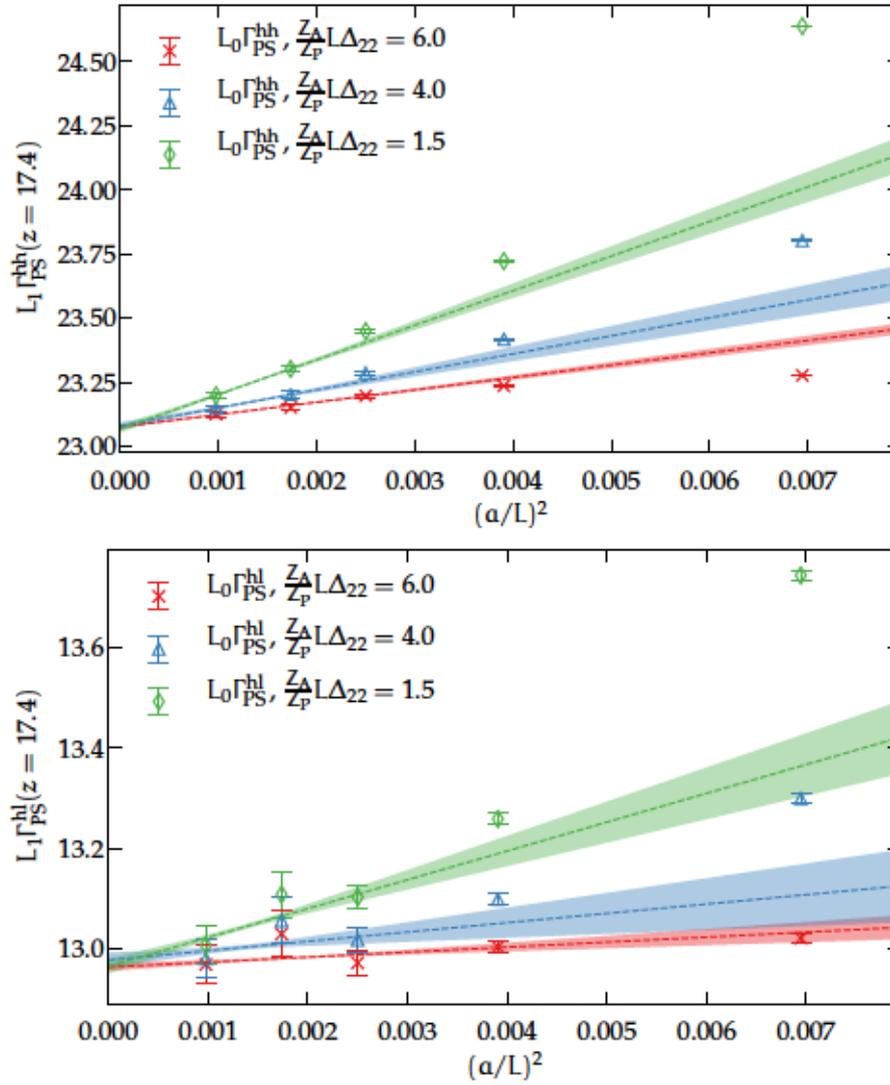


Figure 12.16: Scaling of $L_1 \Gamma_{PS}$ evaluated at $\kappa_h(17.4, L/a)$ for different LCP together with the continuum extrapolation. *Top*: Heavy-heavy current. *Bottom*: Heavy-light current.

12.7 CONCLUSIONS

We have presented the determination of all improvement coefficients and renormalization constants that are needed to fix the renormalization group invariant heavy quark mass in the matching of QCD and HQET. This calculation is an important step towards the matching in the three-flavor theory.

We have determined the coefficients b_m and Z defined in a large range of mass scales and tested the cancellation of $O(a)$ cut-off effects as well as the suppression of mass-dependent higher order cut-off effects for two observables defined in the matching volume. No remaining $O(a)$ effects can be observed in the test observables. Both tests suggest that the summation of mass-dependent cut-off effects to all orders in the definition of b_m and Z from the setup with a heavy valence quark leads to the reduction of higher order cut-off effects and therefore allows us to include results at relatively coarse lattice spacings in the continuum extrapolation of the matching parameters.

The careful analysis of all ambiguities in our data makes us confident that no hidden systematic effects are present. Therefore, the results of this chapter are ready to be used to determine the heavy quark hopping parameters κ_h for the determination of the matching coefficients for selected values of the heavy quark mass on the ensembles described in chapter 8.

Part IV

THE MASS OF THE CHARM QUARK

CLS ENSEMBLES

When physical quantities (as opposed to renormalization constants) are extracted from lattice simulations, the spatial extent of the lattices has to be large enough to accommodate all relevant degrees of freedom. In the case of mesonic observables, the lattice has to be large enough to allow mesons to propagate freely without feeling the boundaries [240–244]. Interactions around the spatial torus of the lattice lead to $O(\exp(-m_\pi L))$ effects and the squeezing of the wave function leads to effects of $O(1/L^n)$ with $n = 2 - 3$ [14]. For QCD one can deduce that $m_\pi L \geq 4$ and $L > 3$ fm are appropriate parameters to sufficiently suppress finite-volume effects for hadron masses.

The generation of ensembles with large physical volumes and small quark masses is computationally demanding and is mostly performed by larger collaborations. In our work, we are able to use the ensembles which are generated by members of the Coordinated Lattice Simulations (CLS) effort [60, 235, 236]. We work on the ensembles with $N_f = 2 + 1$ flavors of dynamical fermions. The effect of a dynamical charm quark in low-energy observables is expected to be suppressed below the available accuracy [245]. This is different for charmed observables [100, 246] and therefore a systematic uncertainty is present in physical observables including a charm quark from $2 + 1$ flavor lattice QCD.

Since we perform our work on renormalization and improvement with respect to this set of large volume simulations, the gauge and fermion actions of the CLS ensembles coincide with those of our Schrödinger functional ensembles. In contrast to our small-volume simulations, open boundary conditions in the time direction are used. These allow for a smooth variation of the topological charge. The generated configurations are of size $N_t \times N_s^3$ with open boundaries at 0 and $N_t - 1$, resulting in a physical time extent of $T = (N_t - 1)a$.

All of the algorithmic techniques described in section 5.1 have been used in the generation of the ensembles. The use of the rational approximation for the strange quark and twisted mass reweighting of the light quarks results in the necessity of two reweighting factors. Five lattice spacings in a range from 0.087 fm down to 0.04 fm allow for a save continuum extrapolation. These small lattice spacings in combination with $O(a)$ improvement are needed to perform the continuum extrapolations using Wilson quarks.

The pion masses on the CLS ensembles are mostly above the mass of the physical pion. Physical quark masses have been simulated on two ensembles at coarse and intermediate lattice spacing. Since simulations

In QCD, the pion is the lightest mesonic degree of freedom. In other theories, a different lightest particle mass has to be considered.

There exists a set of ensembles with $N_f = 2$.

The inclusion of the sign of the rational determinant leads to a third reweighting factor, which has not yet been included in our results [102].

and measurements at small quark masses are quite expensive, the chiral extrapolation is constrained by other ensembles spanning a range of pion masses $\in [180 \text{ MeV}, 420 \text{ MeV}]$.

Depending on the objective, different subsets of ensembles can be chosen. Charm observables suffer from large cut-off effects and profit from small lattice spacings. Their dependence on the light degrees of freedom is small. Therefore the ensembles with pion masses below 200 MeV can be left out of the computation. For static-light quantities, the situation is complementary. Here, it is beneficial to include small quark masses and to leave out the finest lattice spacing. An overview of all ensembles used in our work is given in figure 13.1.

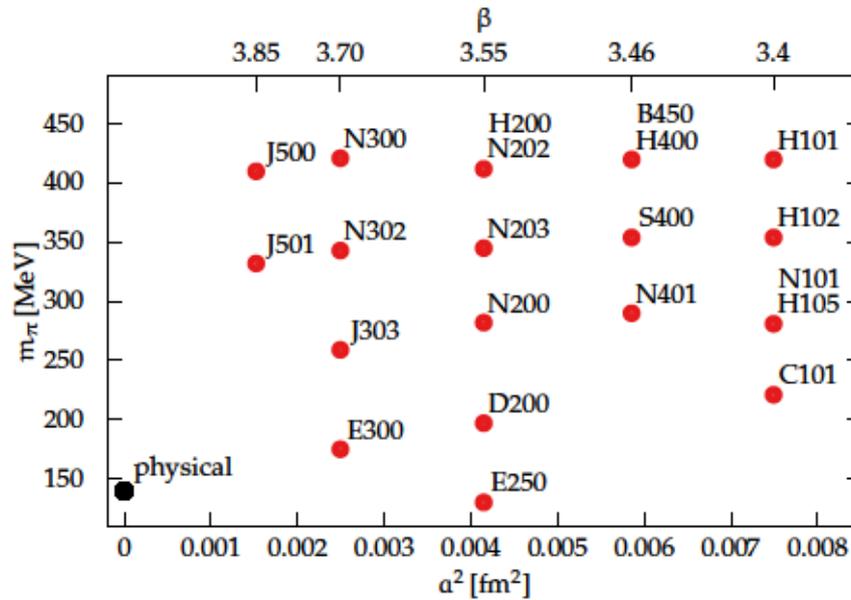


Figure 13.1: CLS ensembles on the $\text{Tr}[M_q] = \text{const.}$ trajectory used in this work in the landscape of lattice spacing and pion mass. The corresponding bare inverse coupling β is given on the top and the kaon mass is implicitly defined by the trajectory. For some points in the parameter space, two ensembles with the same physical parameters and different geometries have been generated.

The Monte Carlo trajectory of each ensemble has to be long enough to allow for the correct determination of the statistical errors. In [60] the lower bound of $O(50) \times \tau_{\text{exp}}$ is given. The exponential autocorrelation time on the CLS ensembles, which we use to attach an exponential tail according to eq. (5.84), is estimated using the formula

$$\tau_{\text{exp}} \approx 14(3) \frac{t_0}{a^2} \quad (13.1)$$

which has been given in [60] based on an analysis of the square of the topological charge and the Yang-Mills action density from the Wilson flow.

13.1 THE CHIRAL TRAJECTORY

After the extraction of physical observables on a set of ensembles we have to extrapolate the results to the physical point which is defined by physical quark masses and vanishing lattice spacing. How this point is approached via a combined chiral-continuum extrapolation depends on the setup that has been chosen for the generation of the ensembles.

When simulations with $N_f = 2 + 1$ or $N_f = 2 + 1 + 1$ flavors are performed, most lattice collaborations use a setup where the renormalized strange quark mass is kept at its physical value while the light quark masses span a range above the physical ones. Keeping the strange quark mass constant and directly at the physical value is quite difficult and deviations of the strange quark mass from its physical value have to be expected for the majority of the ensembles. This can be compensated for by a quark mass reweighting [109, 247] or by the simulation of new ensembles with adjusted parameters.

Such setup with a constant renormalized strange quark mass exists within the set of CLS ensembles [235] but the renormalization and improvement pattern within this setup has further complications compared to our discussion in chapter 4. As an example we can consider the improved coupling \tilde{g}_0^2 introduced in eq. (4.4). When the strange quark mass is kept constant while the light quark masses are decreased at fixed bare coupling g_0^2 , the renormalized coupling and the lattice spacing change at $O(a)$.

It has been realized by the QCDSF collaboration [248] that it can be beneficial to keep the sum of bare quark masses $\text{Tr}[M_q] = m_u + m_d + m_s = 2m_l + m_s$ constant when the light quark masses are decreased towards their physical values. Since gluonic observables and the central value of hadron multiplets depend strongly on $\text{Tr}[M_q]$ and only subdominantly on $m_l - m_s$, a smooth extrapolation without large changes in the scale can be expected and less corrections have to be applied. As further advantage, mass dependent cut-off effects of order $O(a\text{Tr}[M_q])$ do not change for a fixed value of the bare coupling.

Within the set of ensembles with a constant sum of quark masses, which we will refer to as $\text{Tr}[M_q] = \text{const}$, ensembles at different lattice spacings are matched via the dimensionless quantities

$$\phi_2 = 8t_0 m_\pi^2 \quad \text{and} \quad \phi_4 = 8t_0 \left(m_K^2 + \frac{1}{2} m_\pi^2 \right) \quad (13.2)$$

with t_0 defined from the Wilson flow as defined in eq. (6.40). To leading order in chiral perturbation theory, these quantities are proportional to the sum of the corresponding quark masses [249, 250]

$$\phi_2 \propto m_u + m_d, \quad \phi_4 \propto m_u + m_d + m_s. \quad (13.3)$$

The renormalized quark mass matrix has been introduced in eq. (4.16) and if the bare sum of quark masses is held constant as definition of

the chiral trajectory, this does not imply a constant physics condition. Since the deviations of $\text{Tr} [M_q]_R$ from a constant value at constant $\text{Tr} [M_q]$ have turned out larger than one would expect it from $O(a)$ effects [77, 142], the chiral trajectory of the CLS ensembles has been redefined to

$$\phi_4 = \text{const.} = \phi_4^{\text{phys}}. \quad (13.4)$$

The physical values of ϕ_4 and ϕ_2 determined from the physical value of t_0 and isospin corrected pion and kaon masses without QED effects have been determined in [142],

$$\phi_4^{\text{phys}} = 1.119(21) \quad \phi_2^{\text{phys}} = 0.0804(8). \quad (13.5)$$

In the our computations, the goal is to evaluate all observables at $\phi_4 = \phi_4^{\text{phys}}$ such that the dependence on the quark masses can be parametrized by ϕ_2 and physical quark masses are reached when $\phi_2 = \phi_2^{\text{phys}}$.

13.2 QUARK MASS SHIFTS

Due to the setup of the ensembles via $\text{Tr} [M_q]$, deviations of ϕ_4 from its physical value are present and have to be corrected accordingly. We perform this correction via a shift of the sea quark masses. The form of this shift is given by the first term of the Taylor expansion of the observable with respect to the quark mass m_f [142, 220],

$$\langle \mathcal{O} \rangle|_{m_f} = \langle \mathcal{O} \rangle|_{m_f} + \Delta m \frac{d\langle \mathcal{O} \rangle}{dm_f}. \quad (13.6)$$

The total derivative of an observable \mathcal{O} with respect to a change in the quark mass m_f is given by

$$\frac{d\langle \mathcal{O} \rangle}{dm_f} = \left\langle \frac{\partial \mathcal{O}}{\partial m_f} \right\rangle - \left\langle \mathcal{O} \frac{\partial S}{\partial m_f} \right\rangle - \mathcal{O} \left\langle \frac{\partial S}{\partial m_f} \right\rangle \quad (13.7)$$

where the partial derivative of the observable with respect to the quark mass as well as the derivative of the action with respect to the quark mass have to be taken into account. Reweighting can be straightforwardly applied in the definition of the expectation values as given by eq. (5.24).

If the expectation value is defined in terms of the fermion and gauge fields and the observable is defined from fermion fields, it may be necessary to consider additional Wick contractions in the second term of eq. (13.7). However, if we consider the expectation value after the fermions have been integrated out, observables based on correlation functions are defined from propagators and no additional Wick contractions have to be performed.

The derivative of the action with respect to the quark masses can be determined from the trace of the quark propagator [251],

$$\left\langle \frac{\partial S}{\partial m_f} \right\rangle = \sum_x \langle \bar{\Psi}_f(x) \psi_f(x) \rangle = - \sum_x \langle \text{Tr} [\mathcal{S}_f(x, x)] \rangle^{\text{gauge}}, \quad (13.8)$$

where the trace acts on spin and color indices and ψ_f and \mathcal{S}_f denote the quark fields and the propagator of quark flavor f .

Whereas the derivative of the action with respect to the quark masses has to be evaluated once per configuration, we have to explicitly determine the derivatives of the propagators used in our observables with respect to a change of the sea quark masses present in the observables. In our project on the determination of the charm quark mass, we perform this calculation using the `mesons` code [252].

Having the partial derivatives at hand, we have to determine the size of the shift $\Delta m \equiv \Delta m_s \equiv \Delta m_l$ that is necessary to obtain $\phi_4 = \phi_4^{\text{phys}}$. Afterwards we shift all observables involved in our computations accordingly. This shift is performed on the level of correlation functions and propagates into all derived observables.

DETERMINATION OF THE CHARM QUARK MASS

Quark masses are fundamental parameters of the Standard Model. As such, their knowledge is of general interest. Furthermore, the masses of charm and bottom quarks are used as inputs for tests of the Standard Model via weak decays and CKM matrix elements [253] and therefore their precise knowledge is especially important.

However, the experimental determination of quark masses is difficult since quarks are confined in bound states and cannot be isolated. The mass of these bound states is not just a simple sum of its parts, but depends on an interplay of non-perturbative dynamics. QCD provides the theoretical background to compute Standard Model parameters but the solution of the non-linear equations is highly non-trivial. We use lattice QCD to solve the underlying equations in order to extract Standard Model parameters as quark masses.

As explained before, we have to live with some limitations regarding our simulations. In our $2 + 1$ flavor simulations, we do not take into account the effect of charm quark loops and for observables at the charm scale, this is the source of a significant systematic uncertainty. In [100, 246], an upper bound of 5% for the size of this uncertainty has been given. However, the comparison of charm quark determinations based on $2 + 1$ flavors [210, 254, 255] and $2 + 1 + 1$ flavors [256–259] given in [209] shows an excellent agreement.

We will determine the charm quark mass from bare current quark masses and perform a non-perturbative renormalization and improvement [31, 48, 218, 260]. The same strategy has been applied in works of the ALPHA collaboration in the quenched and two-flavor theory [261–263]. A preliminary account of our work with $2 + 1$ flavors has been presented in [264].

In this work, we will concentrate on the extraction of physical observables on the CLS ensembles and discuss our strategies to circumvent numerical problems. The results obtained in this work will be the basis of a fully correlated combined chiral-continuum extrapolation. We will show such extrapolation and provide a preliminary result for the charm quark mass at the physical point. A detailed investigation of the systematic effects of this extrapolation will be presented in [265] and in our forthcoming publication [266].

14.1 CALIBRATION OF THE MEASUREMENTS

We like to determine the mass of the charm quark on ensembles with $N_f = 2 + 1$ flavors of dynamical fermions. Therefore, we have to

The deviations within the results from $2 + 1 + 1$ flavors are significantly larger than the difference between the averaged values from three and four flavor theory.

introduce charm observables in a partially quenched setup where no charm quark loops are present in the path integral.

The addition of the charm quark as physical parameter to the simulations requires to fix the hopping parameter of the heavy valence quarks such that the renormalized valence quark masses on each ensemble match the physical charm quark mass. As in the case of the light quarks, we like to employ meson masses for this calibration, since they are relatively easy to compute and provide a good signal.

Different paths to fix the physical charm quark can be taken. For a comparison of different approaches see [267]. There, the masses of the D , the D_s and the η_c meson have been investigated to fix κ_c and all approaches have shown different advantages and disadvantages. The D meson is relatively noisy and has a strong dependence on the light quark masses which change considerably towards the chiral limit. The signal of the D_s meson is much cleaner but in the case of [267], where the strange quark is fixed to values close to the physical one, a correction for a slight mistuning of the strange quark mass had to be applied. The η_c defined from the connected Wick contraction has a very clean signal but it differs from the physical meson by disconnected contributions. These contributions are assumed to be smaller than the systematic corrections due to neglecting QED and isospin effects.

Our setup is different compared to [267] since we keep the sum of light and strange quark masses constant on the chiral trajectory. Therefore, the mass of the D_s meson changes towards the chiral limit due to the increase in the strange quark mass. To arrive at a clean observable with a small dependence on the sea quark masses, we will consider the flavor averaged meson mass

$$m_{\overline{D}} \equiv \frac{2}{3}m_D + \frac{1}{3}m_{D_s}, \quad (14.1)$$

to fix the physical charm quark mass. The physical value of this flavor-averaged mass without QED and isospin effect is given by [2, 209, 268]

$$m_{\overline{D}}^{\text{phys}} = 1899.4(2) \text{ MeV} \quad (14.2)$$

In eqs. (7.28–7.29) we have introduced the fine structure of mesons with one heavy quark in the framework of HQET. If we assume the charm quark to be heavy enough for HQET to be applicable, we can use our insights to eliminate the spin contribution to the first order corrections via the spin average

$$m_{\overline{H}} = \frac{1}{4}(3m_V + m_{PS}) = m_h + \bar{\Lambda} - \frac{\lambda_1}{2m_h}, \quad (14.3)$$

where m_V and m_{PS} are the masses of vector and pseudoscalar mesons and m_h is the mass of the heavy quark. Since the fine structure

vanishes in the static limit and is sensitive to short distance effects, we assume that we can reduce cut-off effects when the next-to-leading order contribution is reduced.

We will combine the spin average with the flavor average and define the spin-flavor averaged mass to be

$$m_{\overline{D}} = \frac{1}{12}(6m_{D^*} + 2m_D + 3m_{D_s^*} + m_{D_s}). \quad (14.4)$$

With this definition, we have to take into account the masses of the vector mesons m_{D^*} and $m_{D_s^*}$ which are comparably noisy. Therefore it is not clear if the assumed reduction of cut-off effects compensates the increase in the statistical error. In our analysis, we will use the flavor averaged and the spin-flavor averaged meson mass to fix the physical charm quark mass and we will compare both approaches in the final chiral-continuum extrapolation.

So far we have not explained how we fix the heavy quark mass from the meson masses. We perform measurements for two different heavy quark masses in the region around the physical charm quark mass. For both heavy valence quarks, we determine the flavor and spin-flavor averaged heavy meson masses and use those to determine the hopping parameter of the charm quark κ_c . Using initial values provided by [269], we interpolate linearly to the physical quark mass. The linear dependence of m_{D_s} on $1/\kappa_h$ is visible in the results of [270], where measurements at four different heavy quark masses have been performed. If we are able to interpolate rather than extrapolate from our two choices of heavy quark hopping parameters κ_{h_1} and κ_{h_2} will become clear only after the measurements have been performed. The hopping parameters used in our measurements are listed in table 14.1.

We do not perform a third set of measurements at the extracted value of κ_c but rather interpolate all physical observables based on the two heavy quark masses to the physical point defined by κ_c . This interpolation is combined with the chiral-continuum fit.

Table 14.1: Hopping parameters of the partially quenched heavy valence quarks for the determination of κ_c provided by [269].

β	κ_{h_1}	κ_{h_2}
3.40	0.123147	0.124056
3.46	0.125563292	0.126983423
3.55	0.1274374	0.128651119
3.70	0.13018588	0.13062697
3.85	0.13206693	0.13242984

14.2 THE RENORMALIZED CHARM QUARK MASS FROM CURRENT QUARK MASSES

We use the different types of renormalized and improved quark masses introduced in section 4.3 to determine the physical value of the charm quark mass. The calculation of renormalized quark masses from subtracted quark masses according to eq. (4.17) can be performed using the value of κ_c which is determined from the interpolation in the (spin-)flavor averaged meson masses. However, this definition has sizable systematic uncertainties, since the factor $(r_m - 1)$, which cancels the leading order dependence on the sea quark masses, is unknown so far.

There are efforts to determine this coefficient.

The critical hopping parameters that enter the subtracted quark mass have been determined in [68]. Since the quark masses on the $\text{Tr}[M_q] = \text{const.}$ trajectory are far from being chiral, an additional systematic uncertainty is added to the bare subtracted quark mass via the uncertainty in κ_{cr} . Therefore, we do not consider renormalized subtracted quark masses in this project.

Instead, we determine the renormalized quark mass from current quark masses, according to eq. (4.31). The renormalization constants are known from [66] for Z_A , from [64] for Z_P and from our work in chapter 11 for Z . The valence mass-dependent cut-off effect of $O(a)$ can be canceled with $b_A - b_P$ and Z of chapter 11.

We will test, whether LCP-1 removes the cut-off effects more efficiently than LCP-0.

Remnant cut-off effects of $O(a \text{Tr}[M_q])$ remain in our data, but these are suppressed by an order of magnitude compared to the effects of the valence quarks. Nevertheless, we have to check for the existence of these cut-off effects in the final extrapolation to the continuum limit.

We have different options to construct a renormalized charm quark mass from valence and sea quark propagators. The obvious choice is the construction from the current quark mass $m_{\text{hh}'}$ based on two mass-degenerate valence quark propagators. After the interpolation of κ_h to κ_c , the quark mass $m_{\text{cc}'}$ is obtained. We define

$$m_{\text{R,c}}^{(\text{c})} = m_{\text{R,cc}'} \quad (14.5)$$

Since current quark masses from heavy-heavy correlation functions have a clean signal and show long plateaus, provided that distance preconditioning is used in the solver, this mass can be determined to high precision. The valence quark part of the mass-dependent cut-off effects scales with $m_{\text{cc}'}$ and we expect rather large cut-off effects for this definition of the charm quark mass, even though the leading order effects are subtracted.

We hope to reduce the size of the cut-off effects by the use of heavy-light and heavy-strange correlation functions. When we determine light, strange and valence quark propagators, we can define the charm quark mass via

$$m_{\text{R,c}}^{(\text{l})} = 2m_{\text{R,lc}} - m_{\text{R,ll}'}, \quad m_{\text{R,c}}^{(\text{s})} = 2m_{\text{R,sc}} - m_{\text{R,ss}'} \quad (14.6)$$

where we again use eq. (4.31) to renormalize current quark masses. From this definition, the mass-dependent cut-off effects of $O(am_{ij}^2)$ are reduced, compared to the definition from mass-degenerate heavy quarks. In return, we expect a stronger dependence on the light quark masses.

We try to reduce this dependence by the definition of the flavor-averaged charm quark mass

$$\bar{m}_{R,c} = \frac{2}{3}m_{R,c}^{(l)} + \frac{1}{3}m_{R,c}^{(s)}. \quad (14.7)$$

Since the sum of light and strange quark masses remains constant on our set of ensembles, we believe that this definition reduces the dependence of the charm quark mass on the position on the chiral trajectory.

We can construct yet another definition of the renormalized charm quark mass from the ratio difference method, eq. (4.38), where the first flavor is chosen to be the heavy valence quark and the second one is one of the sea quarks. Based on the same argument as above, we define

$$m_{R,c}^{(rd)} = Z_m \left(\frac{2}{3} \frac{r_{l,cl} d_{l,cl}}{r_{l,cl} - 1} + \frac{1}{3} \frac{r_{l,cs} d_{l,cs}}{r_{l,cs} - 1} \right), \quad (14.8)$$

where improved ratio and difference have been defined in eqs. (4.36 – 4.37). The systematic uncertainties which are present in the subtracted quark masses are canceled in the quark mass from the ratio-difference method.

These three definitions of the renormalized charm quark mass are the basis for our chiral-continuum extrapolation. In principle, all of them are valid and we expect them to lead to the same result in the continuum, albeit with different statistical and systematic uncertainties. It remains to be checked whether a combined extrapolation stabilizes the chiral-continuum fit or if we use the different definitions to estimate systematic uncertainties of our final result.

In principle even more variation is possible within this set of quark masses. For example, we can use standard or improved derivatives for the determination of the PCAC masses, we can choose between different sets of improvement coefficients and renormalization constants and we can investigate quark masses from heavy-light correlation functions without a flavor average. All of these choices can be made at the stage of the analysis and will be studied in detail in [265, 266].

To quote renormalization group invariant values of the charm quark mass, the running factor

$$\frac{M}{m_{R,\text{had}}} = 0.9148(88) \quad (14.9)$$

determined in [64] is used for the scale evolution from the CLS ensembles. Its relative error of about 1% will most likely dominate our

A similar result has been obtained in quenched QCD [261].

We have introduced the running of quark masses in section 8.3.

result at the physical point. The error is only added to the total error after the chiral-continuum extrapolation.

The increase in the uncertainty due to the errors on the renormalization factors and the running factor can be avoided when ratios of quark masses are determined. A standard choice is the ratio m_c/m_s . This might be related to the fact that the strange quark mass is held constant in most computations and therefore the chiral extrapolation of this ratio is very mild. In our setup of $\text{Tr}[M_q] = \text{const}$, where the strange quark mass changes significantly towards the chiral limit, such chiral-continuum extrapolation is possible, but we might face large uncertainties due to the slope in the chiral fit. We will also test the determination of $m_c/(2m_l + m_s)$ where we expect the dependence on the light quark masses to be small.

We can employ quark mass ratios to compare our result with others while neglecting the effect of different renormalization procedures. The final value can be translated to physical charm quark masses by using the prior knowledge on physical light quark masses, e.g., from [77].

In the last chapter we have outlined our strategy for the determination of the charm quark mass and described which quantities have to be computed in order to calculate renormalized quark masses and to calibrate the bare parameters of the theory. In this chapter we will describe the calculation of correlation functions and the extraction of the physical quantities on the large-volume CLS ensembles. We will discuss systematic effects and how we try to estimate or circumvent them and we will quote results for quark and meson masses on all ensembles used in our study. An overview of the ensembles that have been considered in this study is given in table B.8.

Together with the data for the quark mass dependence of the action (13.8) and the set of improvement coefficients and renormalization constants determined by the ALPHA collaboration, we have gathered all ingredients for the extrapolations to the physical point.

15.1 CORRELATION FUNCTIONS

In order to calculate current quark masses from the PCAC relation according to eq. (4.25), we need to determine the time component of the axial current A_0 and the pseudoscalar density P on single time slices x_0 . For this task, we use zero-momentum two-point correlation functions, defined by

$$f_{O_1 O_2}^{rs}(x_0, y_0) = -\frac{a^6}{L^3} \sum_{\mathbf{x}, \mathbf{y}} \langle O_1^{rs}(x_0, \mathbf{x}) O_2^{\dagger, rs}(y_0, \mathbf{y}) \rangle, \quad (15.1)$$

where r and s are the flavor indices. We only consider the case of two distinct flavors $r \neq s$. y_0 is the time-coordinate of the source, i.e., the time slice where the source has a non-vanishing norm, and x_0 is the time coordinate of the sink. The summation over the spatial indices leads to the projection onto zero-momentum [14]. The operators O^{rs} are defined by

$$O^{rs}(x) = \bar{\psi}^r(x) \Gamma \psi^s(x) \quad (15.2)$$

where Γ is a product of gamma matrices, representing the correct quantum numbers for the problem at hand. In this section, we will use the operators defined in table 15.1.

Using time reflection symmetry, we can average over correlation functions at the source positions y_0 and $T - y_0$. In this way, we can increase the statistics without introducing complications due to

We work with the flavor non-singlet PCAC relation.

Table 15.1: Operators used in the determination of the charm quark mass, based on eq. (15.2). We restrict ourselves to the time components of axial and tensor current.

State	Operator	Γ
Pseudoscalar	P^{rs}	γ_5
Vector	V_i^{rs}	γ_i
Axial vector	A_0^{rs}	$\gamma_0\gamma_5$
Tensor	T_{0i}^{rs}	$\gamma_0\gamma_i$

different absolute source positions with respect to the open boundaries. We consider the functions

$$f_P^{rs}(x_0, y_0) = \frac{1}{2} [f_{PP}^{rs}(x_0, y_0) + f_{PP}^{rs}(T - x_0, T - y_0)] \quad (15.3)$$

$$f_{A_i}^{rs}(x_0, y_0) = \frac{1}{2} [f_{PA_i}^{rs}(x_0, y_0) - f_{PA_i}^{rs}(T - x_0, T - y_0)] \quad (15.4)$$

$$f_V^{rs}(x_0, y_0) = \frac{1}{6} \sum_{k=1}^3 [f_{V_k V_k}^{rs}(x_0, y_0) + f_{V_k V_k}^{rs}(T - x_0, T - y_0)] \quad (15.5)$$

$$f_{VT}^{rs}(x_0, y_0) = \frac{1}{6} \sum_{k=1}^3 [f_{V_k T_{0k}}^{rs}(x_0, y_0) + f_{V_k T_{0k}}^{rs}(T - x_0, T - y_0)]. \quad (15.6)$$

To obtain $O(a)$ improved observables, we employ the the improved correlation functions,

$$f_{A_i}^{rs,I}(x_0, y_0) = f_{A_i}^{rs}(x_0, y_0) + ac_A \partial_i f_P^{rs}(x_0, y_0) \quad (15.7)$$

$$f_V^{rs,I}(x_0, y_0) = f_V^{rs}(x_0, y_0) + ac_V \partial_0 f_{VT}^{rs}(x_0, y_0), \quad (15.8)$$

which we use to determine the $O(a)$ improved PCAC mass (4.26)

$$am_{rs}(x_0, y_0) = \frac{\tilde{\partial}_0 f_{A_0}^{rs}(x_0, y_0) + ac_A \partial_0^* \partial_0 f_P^{rs}(x_0, y_0)}{2f_P^{rs}(x_0, y_0)}. \quad (15.9)$$

The correlation functions which are defined in eqs. (15.3-15.5) are also used to compute the effective masses of mesons with the quark content rs and the quantum numbers of the respective operators, according to eq. (6.5). For the pseudoscalar mesons, we use the correlation function f_P . A simultaneous fit to f_P and f_{A_0} does not improve the statistical precision since both are highly correlated. For the vector mesons, we average over the three components of f_V .

As explained in section 6.1.2, the source position has to be chosen with some care, in the case of open boundary conditions. We place the source directly at the boundary, i.e., at $y_0 = a$ and consider the boundary states as excited state contributions that vanish for large

The discretized derivatives are defined in app. A.3.

source-sink separations. This has the advantage that possible boundary cut-off effects are expected to manifest themselves in a region which is anyway discarded for the extraction of physical quantities. In this way the plateaus of the pseudo-goldstone bosons, which do not suffer from an exponentially decreasing signal-to-noise ratio, are maximized when $y_0 = a$ is chosen.

We choose stochastic sources with $U(1)$ noise and 16 noise sources per time-slice (a and $T - a$). It has been found in [119] that the correlation functions for light quarks computed from displaced sources in the bulk of the lattice are fully correlated. Therefore, further source positions are not expected to decrease the statistical error of the quantities under investigation.

The same number of sources has been used in [77, 142] for the determination of quantities based on light quarks. A priori we do not know whether the same number of noise sources is appropriate for observables with heavy quarks which are not present in the sea. We will investigate the influence of the number of noise sources on the statistical error of quark and meson masses in section 15.5.

In our computations of heavy quark propagators we have used distance preconditioning to obtain the required level of precision at large source-sink separations. We describe the tuning of the parameter α in section 15.4 where we also investigate the reduction of systematic uncertainties.

15.2 MESON MASSES

To determine the position of an ensemble on the chiral trajectory and the charm quark hopping parameter κ_c , we have to determine the masses of pion and kaon as well as the masses of the heavy mesons involved in the spin-flavor average in eq. (14.4).

As explained in section 6.1, the extraction of the ground state matrix elements and energies of correlation functions is straight forward at large time distances, since we face an exponentially suppressed excited state contamination. If the signal-to-noise-ratio of the correlation function under investigation vanishes exponentially, we have to determine the region, where a safe extraction of the ground state energy is possible.

The extraction of the mass of the ground state mesons can be performed using the correlation function or the effective mass built from it and we choose the latter. We expect our fits to be more stable when the effective mass is used, since one degree of freedom is eliminated compared to the fit to the correlation function.

We can take different paths to determine the ground state energies in the presence of excited state contributions. In the first approach we perform a fit including the excited state contributions in order to determine the source-sink-separation t where the contribution

The variance of hadron propagators and the properties of stochastic sources are explained in [87].

This was, e.g., done in [142]

of all excited states is negligible compared to the statistical errors. Afterwards, we perform a fit to the ground state and can safely neglect all excited state contributions. In the case of the effective mass, this fit to the ground state corresponds to a plateau average.

In the second approach, we also start with a fit including the ground state and the first excited state. In this case, we obtain the energy of the ground state directly from this two-state fit. The fit range has to be chosen properly to include the region where higher order states have decayed and the first excited state can be properly captured. It remains to be tested numerically if one of the two approaches leads to smaller statistical or systematic uncertainties.

We use weighted fits to remove a potential bias from highly fluctuating data points at large source-sink separations.

The two-state fit to the effective mass has three free parameters and is non-linear in one of the parameters such that one could expect larger uncertainties. This can be counteracted by the fact that we can include more data points compared to the plateau fit and that the relative error on effective masses is smaller for shorter source-sink separations (in the case of heavy pseudoscalar and vector mesons). In principle, special care has to be taken in the determination of the errors on the fit parameters in the case of non-linear fits [140]. Since we determine these errors via resampling techniques, this complication is avoided.

Before we explain how we determine the effective masses from the data, we like to look at the more general picture, namely the bare data for the quantities which we use to calibrate our measurements. These are the pion mass m_π , the kaon mass m_K for the chiral trajectory and the flavor averaged mass $m_{\overline{D}}$ or the spin-flavor averaged mass $m_{\overline{D}}$ respectively to fix κ_c .

We show representative effective masses of the pseudo-goldstone bosons in figure 15.1 and of the heavy mesons which contribute to the averages in figure 15.2. For the light mesons, we obtain plateaus over a large fraction of the lattice, since the signal-to-noise ratio is constant for these bound states. At large source-sink separations, the effect of physical states propagating from the boundaries into the bulk becomes visible.

We will shortly define this point.

For the heavy mesons, we show the effective masses up to the point where the statistical errors become too large. We can see that the mesons including the strange quark show a better signal than the mesons including a light quark. The vector mesons suffer from large statistical errors. From this observation, we expect the spin-flavor averaged mass to have significantly larger errors than the flavor averaged mass.

15.2.1 Fit ranges

Since we know the functional form of the exponential contribution of excited states to the correlation functions used to determine the ground state meson masses, we can model these corrections in our fits.

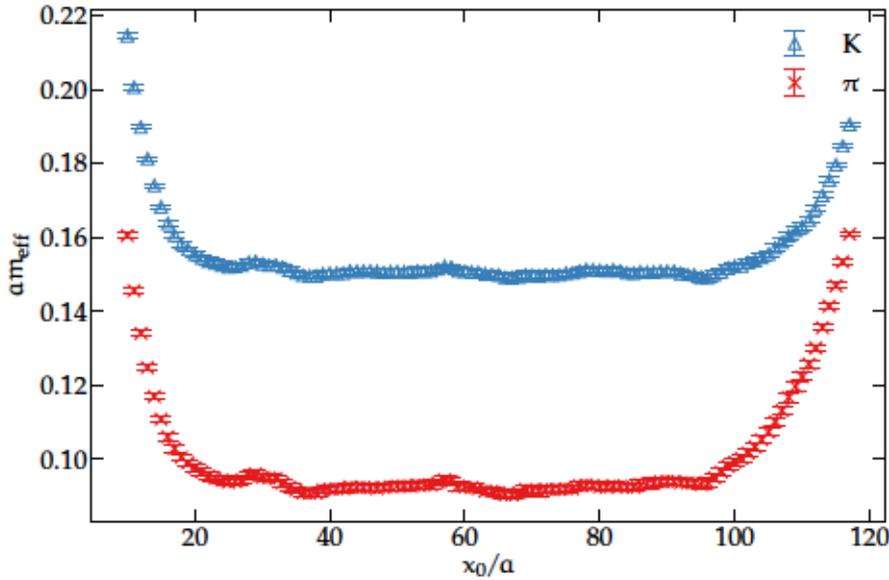


Figure 15.1: Effective masses of the pseudo-goldstone bosons π and K on N_{200}

The discrimination of several terms in a sum of exponential functions is difficult without further knowledge on the size of the exponents. Therefore we try to perform a two-state fit in a region, where all but the lowest excited state have decayed.

The functional form of a two-state fit to effective masses is given by eq. (6.8). Since we do not know where all higher exponential corrections have vanished, we have to test the fit quality for a number of different fit ranges $[t_{\min}, t_{\max}]$.

For this test, we first fix the upper end of the fit interval t_{\max} . In the case of heavy-light mesons, this is done by setting an upper limit on the relative size of the statistical error of the correlation functions. By setting this limit to 2% for pseudoscalar mesons and to 7% for vector mesons, we can exclude regions with excessively fluctuating data points from the fit regions. Since we perform weighted fits and data points with large statistical errors contribute less to the overall fit result, the fit quality is stable under a variation of t_{\max} .

The signal-to-noise-ratio of light pseudoscalar mesons is constant across the temporal extent. Therefore, we can not set t_{\max} by a noise criterion. Instead, we have to exclude the region where boundary states lead to exponential corrections of known functional form. Since this procedure is similar to the determination of the lower bound of the fit region, we will postpone the discussion of the boundary effects and assume that a proper value for t_{\max} has been found such that a plateau of the effective mass is visible up to t_{\max} .

To determine t_{\min} , we perform a number of fits. We start with a fit in the range $[5a, t_{\max}]$ and continue by increasing the lower end of the interval in steps of a . During this process, we monitor the fit quality

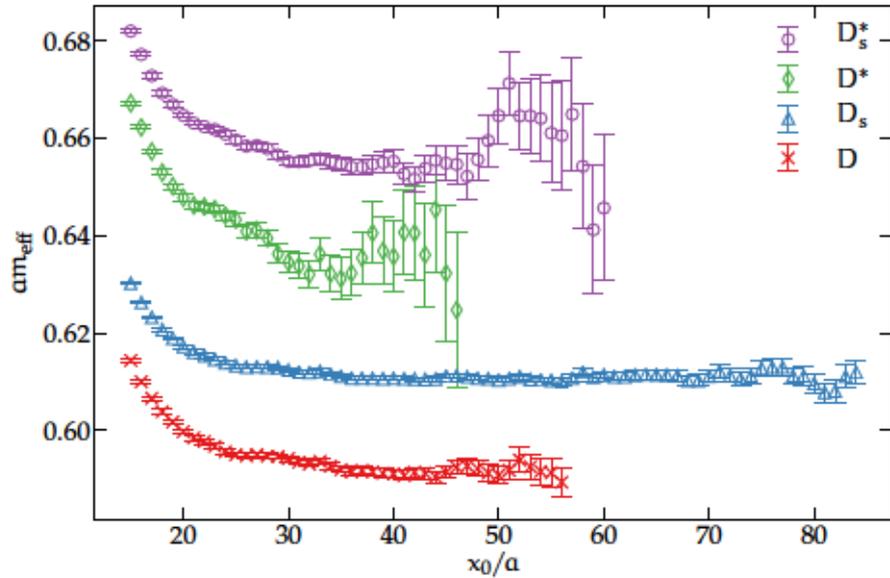


Figure 15.2: Effective masses of the heavy mesons corresponding to D , D_s , D^* and D_s^* including the heavy quark corresponding to κ_{h2} on N_{200} . We exclude data points with large statistical errors.

via the value of $\frac{\chi_{\min}^2}{\chi_{\exp}^2}$, as explained in section 6.3.1. In this procedure, the fit quality improves with every step, since higher order exponential corrections that are not modeled by the fit decay with the increasing temporal distance between source and sink. When the lower bound is increased too much, the fit parameters of the exponential term cannot be determined any more.

At the point where all but the lowest exponential correction have decayed, we expect $\frac{\chi_{\min}^2}{\chi_{\exp}^2} \approx 1$ and indeed this point is found for all effective masses under consideration. In most cases, we find a shallow valley around a local minimum for the values of $\frac{\chi_{\min}^2}{\chi_{\exp}^2}$ determined from different fit ranges. Based on this observation, we are able to choose the value of t_{\min} such that the fit quality is optimized and we are convinced that the exponential correction is modeled correctly.

The strength of the approach to use χ_{\exp}^2 for the determination of the fit quality becomes apparent in this fitting procedure. Since the correlations between the effective masses on different time slices are strong and fit ranges with a length of $O(40a)$ are considered, the inversion of the correlation matrix in a correlated fit fails in most cases. If the naive ratio $\frac{\chi_{\min}^2}{\text{NDF}}$ is used to monitor the quality of the uncorrelated fit, values smaller than unity are reached well before the optimal fit is found. In this case, it is not clear by the value of the ratio when the optimal fit is found.

In the case of the pion and the kaon, the signal to noise ratio is constant and we are able to resolve the contributions from the

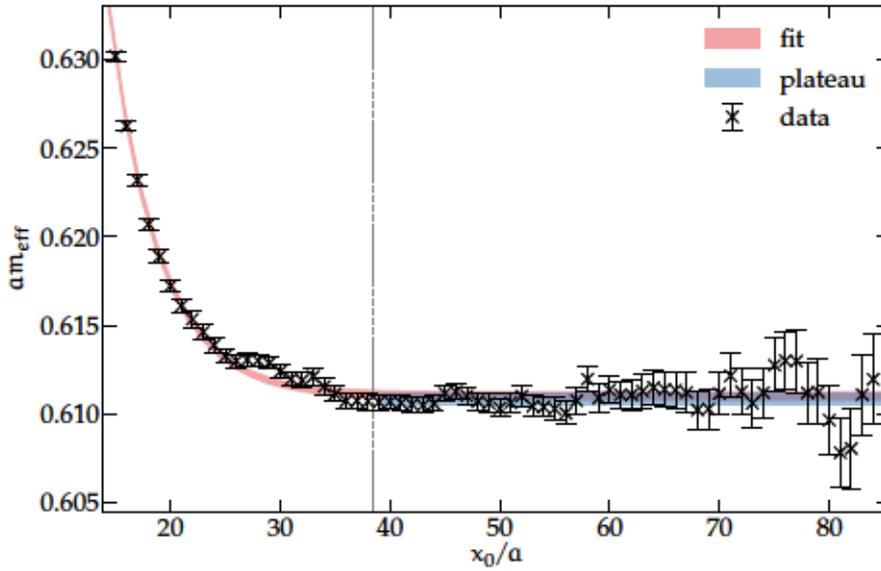


Figure 15.3: Effective mass of a heavy-strange pseudoscalar meson on N200. Data points up to t_{\max} are shown in black. The beginning of the fit interval at $t_{\min} = 11a$ is outside the plot region. The result of the fit is shown by the red curve. The beginning of the plateau region is indicated by the dashed line and the result of the plateau fit is depicted by the blue band.

boundaries. The same mesons have been considered in [142] and there the functional form

$$am_{\text{eff}}(x_0) = am_{\text{PS}}(1 + c_1 e^{-\Delta_1 x_0} + c_2 e^{-E_{2\text{PS}}(T-x_0)} + \dots) \quad (15.10)$$

with $E_{2\text{PS}} \approx 2m_{\text{PS}}$ where m_{PS} is the mass of the pseudoscalar meson, has been chosen for the fits, based on the derivation in [59, 143]. The deviation of $E_{2\text{PS}}$ from $2m_{\text{PS}}$ has its origin in finite-volume effects.

We determine the point where the exponential corrections due to the boundary at T have vanished with the same methods as we have used them for the determination of t_{\min} , i.e., we monitor the fit quality when decreasing the upper limit of the fit range. Since we perform the determination of t_{\max} before we determine t_{\min} , we fix the lower bound to $\approx T/3$ in this procedure. The energy $E_{2\text{PS}}$ is a free parameter in our fit.

When we have determined the optimal fit range, we set t_{\max} to the time slice where the contribution of the exponential correction from the boundary at $x_0 = T$ is smaller than 0.25 times the statistical error of the effective mass. With this procedure, we maximize the fit range and use as many time slices as possible for our determination of the effective meson mass.

An exemplary result of this procedure on the ensemble N200 is depicted in figure 15.5, where we show the data points for the effective mass of the kaon together with the fit to the boundary states. In this case, we obtain $\frac{a}{2}E_{2\text{PS}} = 0.0919(62)$ for the energy of the first

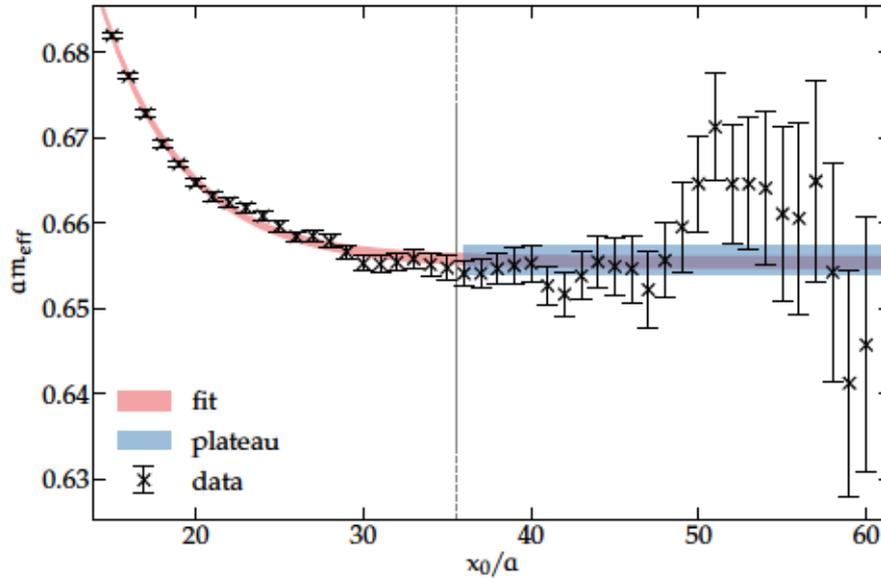


Figure 15.4: Effective mass of a heavy-strange vector meson on N200. Data points up to t_{\max} are shown in black. The beginning of the fit interval at $t_{\min} = 13a$ is outside the plot region. The results of the fit is shown by the red curve. The beginning of the plateau region is indicated by the dashed line and the result of the plateau fit is shown by the blue band.

correction from the boundary source which coincides with the pion mass determined from the full fit, $am_{\pi} = 0.0924(3)$.

Although this is a reassuring observation, the correspondence of E_{2PS} and m_{π} is not crucial for the correctness of our procedure, since we only use it to set the upper bound for the final fit. In all cases we have checked manually that the fit procedure leads to an upper bound where all corrections from the boundary have vanished.

15.2.2 Extraction of the effective mass

Having determined the interval $[t_{\min}, t_{\max}]$, where the data is described by the functional dependence

$$am_{\text{eff}}(x_0) = am_{\text{eff}} + ce^{-\Delta_1 x_0}, \quad (15.11)$$

we have now two possibilities to determine the value of am_{eff} . In the first approach, we take the fit parameter determined by the fit in the chosen fit range. In the second approach, we extract the effective mass from a weighted plateau fit in $[t_{\min}^P, t_{\max}]$. We define the lower end of the plateau range t_{\min}^P to be the time slice where the size of the exponential correction amounts to less than one quarter of the statistical error on $am_{\text{eff}}(x_0)$.

We show the result of the fit on N200 in the window $[t_{\min}, t_{\max}]$ for a heavy-strange pseudoscalar meson in figure 15.3 and for a heavy-s-

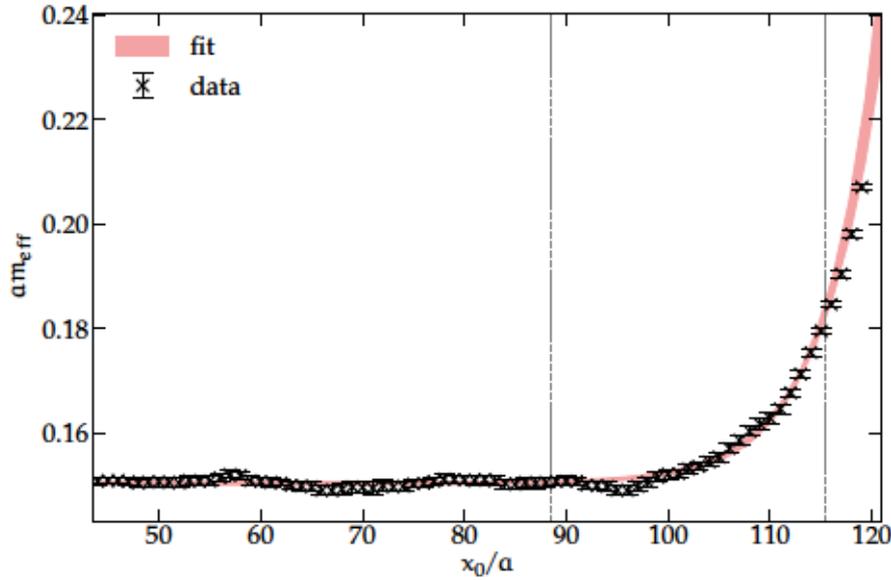


Figure 15.5: Effective mass of the kaon on N200. Data points are shown in black. The fit interval starts at the right dashed line and its result is shown by the red curve. The upper end of the plateau region is indicated by the left dashed line.

trange vector meson in figure 15.4. In both cases, we decide to show a region around the plateau to allow for a comparison of both fit methods. The black data points are shown up to t_{\max} to exclude strongly fluctuating points.

From the figures it can be seen that we have obtained an excellent description of the data within the fit range. The plateau region, where the exponential correction has decayed beyond statistical significance, starts at the dashed line and the result of the weighted average in the plateau region is shown by the blue band.

In most cases the result from the exponential fit and from the plateau average overlap and the statistical errors of the weighted average are larger than those of the full fit. In the case of pseudoscalar meson this difference is only small, since we have a long plateau region with mildly fluctuating data points. In the case of vector mesons, e.g., in fig. 15.4 the statistical error of the plateau fit is significantly larger. This situation would be improved by a more restrictive bound on the size of the statistical error of the primary data points in the determination of t_{\max} . Unfortunately, this would lead to a situation where no plateau region is found for mesons including light quarks on ensembles with small statistics or small light quark masses.

To allow for a better comparison of the results based on both approaches, we show the result from the plateau fit together with the result from the exponential fit for different choices of t_{\min} in figure 15.6. In this exemplary case the fit with the best quality is reached at $t_{\min}/a = 11$. We can see that a small variation of the lower limit

We choose N200 as exemplary ensemble since it is in an intermediate region both in the lattice spacing and the pion mass.

Since the fit parameters are unchanged, t_{\min}^p stays constant for these fits.

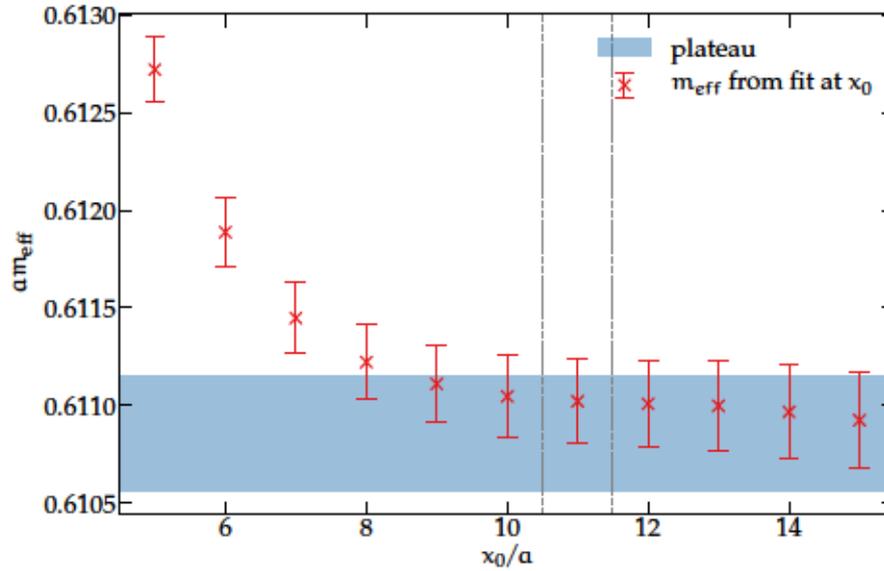


Figure 15.6: Graphical representation of the optimization of the fit region for a heavy-strange pseudoscalar meson on N200. The red crosses show the effective mass obtained from a fit starting at x_0/a . The dashed lines indicate that the best fit was found at $x_0/a = 11$. The blue curve shows the corresponding plateau fit.

does not lead to a change in the fit result. Starting at significantly smaller times leads to incorrect results, since higher states are not modeled correctly. Increasing the lower bound leads to an increase of the statistical error of the fit result, since less data points are taken into account. At some point, the exponential correction cannot be modeled anymore.

In some cases we see a $\approx 1\sigma$ deviation between the results from both methods. In these cases we observe small plateau ranges where local fluctuation of the data have an enhanced influence on the mean value of the plateau fit. Since the exponential fit includes more data points with very small statistical errors, it is not as much affected by these fluctuations as the plateau average.

To rule out systematic effects in the determination of meson masses from the choice of the fit form, we will compare the final results for the RGI charm quark mass after the chiral-continuum extrapolation based on both methods. Since there is no deviation on most ensembles, we expect the difference in the continuum limit to be vanishingly small.

We summarize our results for the light meson masses and the averaged heavy meson masses obtained from the full fit for both heavy quarks in table 15.2. The corresponding values for the plateau averages can be found in tab. B.9. Since the measurements on the J501 ensemble are especially expensive, we were not able to determine the vector correlation functions in this case. Therefore, we cannot give a value for the spin-flavor averaged mass on this ensemble.

From tab. 15.2 it is apparent that the dependence of the heavy meson masses on the light and strange quark masses is very mild. Since we consider flavor averaged quantities and the sum of light and strange quark masses remains constant, this behavior is expected from the construction of the chiral trajectory. The masses of all heavy mesons are listed in tab. B.10 and tab. B.11. From these tables we can see that a decrease in the light quark mass leads to a decrease of the D and D* meson masses and the increase of the strange quark mass leads to an increase of the D_s and D_s* masses.

However, the effects of the light and strange quark masses on the heavy mesons is very mild. This is different for the light mesons, which are influenced by the explicit breaking of the spontaneously broken chiral symmetry.

Table 15.2: Overview of the effective masses used in the calibration of the chiral-continuum extrapolation. The subscript of the flavor averages \overline{D} and spin-flavor averages $\overline{\overline{D}}$ correspond to those of the hopping parameters in table 14.1. No vector correlation functions have been determined for the J501 ensemble. Values are determined from the full fit. The corresponding values from a plateau fit are shown in B.9.

id	am_π	am_K	$am_{\overline{D}_1}$	$am_{\overline{D}_2}$	$am_{\overline{\overline{D}}_1}$	$am_{\overline{\overline{D}}_2}$
H101	0.18280(55)	0.18280(55)	0.85118(52)	0.82029(50)	0.8880(19)	0.8590(19)
H102	0.15416(75)	0.19114(62)	0.84997(59)	0.81898(59)	0.8896(17)	0.8605(18)
H105	0.12142(155)	0.20194(95)	0.84975(74)	0.81877(73)	0.8867(30)	0.8576(29)
C101	0.09706(77)	0.20624(46)	0.84947(49)	0.81847(47)	0.8905(15)	0.8615(15)
H400	0.16328(53)	0.16328(53)	0.74502(51)	0.69355(51)	0.7804(22)	0.7316(24)
H200	0.13651(58)	0.13651(58)	0.64697(61)	0.60109(56)	0.6778(12)	0.6346(12)
N202	0.13424(30)	0.13424(30)	0.64554(57)	0.59955(55)	0.6758(12)	0.6327(9)
N203	0.11229(27)	0.14395(21)	0.64379(62)	0.59796(47)	0.6761(19)	0.6329(19)
N200	0.09242(32)	0.15074(26)	0.64449(52)	0.59862(47)	0.6782(10)	0.6350(11)
D200	0.06473(47)	0.15635(22)	0.64353(44)	0.59774(39)	0.6763(17)	0.6331(16)
N300	0.10561(63)	0.10561(63)	0.49351(83)	0.47571(83)	0.5195(17)	0.5028(17)
N302	0.08750(35)	0.11442(61)	0.49273(67)	0.47499(59)	0.5161(33)	0.4994(33)
J303	0.06484(21)	0.11975(20)	0.49264(40)	0.47478(39)	0.5206(11)	0.5038(11)
J500	0.08085(38)	0.08085(38)	0.37099(78)	0.35515(67)	0.3935(8)	0.3786(8)
J501	0.06632(29)	0.08820(25)	0.37228(41)	0.35636(38)		

15.3 CURRENT QUARK MASSES

As described in section 6.1.2, we expect $O(ag_0^2)$ boundary effect in the PCAC masses, since the boundary improvement coefficients c_G and c_F have not been determined non-perturbatively for open boundary conditions. Additional cut-off effects are introduced from contact terms at small source-sink separations. If the source is placed close to the boundary, both effects overlap. In the bulk of the lattice, we expect the masses $m_{rs}(x_0)$ to form a plateau. In principle, any value within this plateau, e.g., $m_{rs}(T/2)$ could be chosen as estimator for the bare quark mass and all differences based on different choices are expected to scale vanish in the continuum limit.

In the simulations, c_G and c_F are set to tree-level, e.g., to 1.

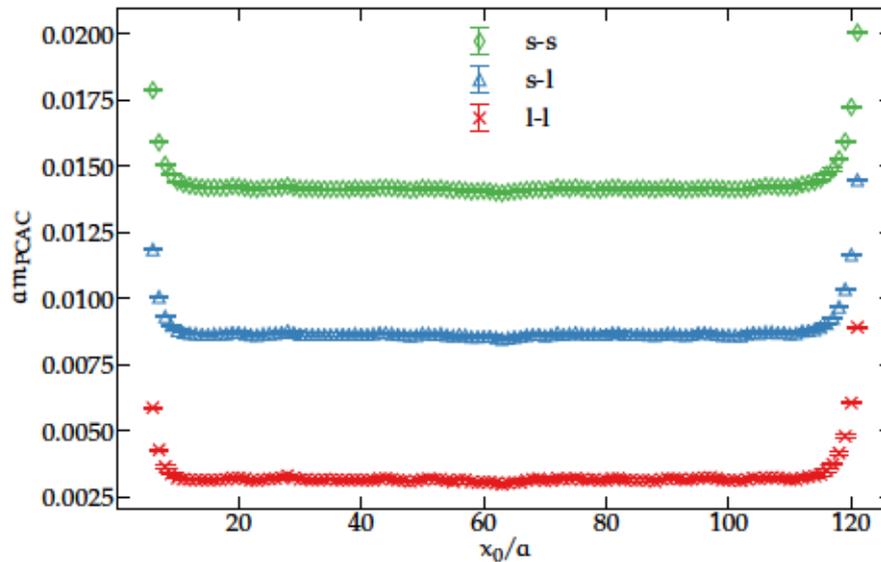


Figure 15.7: Current quark masses based on light and strange propagators determined from standard derivatives.

To use as much information as possible, we like to determine the range, where the quark masses form a plateau and perform a plateau average. In our investigations in the Schrödinger functional, this was mainly done to smooth out fluctuations. In large volume and for quark masses based on heavy-light propagators, we also face a decreasing signal-to-noise ratio with increased source-sink separation. Therefore we are able to decrease the statistical error of our mass determination significantly, if the plateau range is maximized.

As in our studies in the Schrödinger functional, we are free to choose any definition of the discretized derivatives in the determination of the PCAC masses according to eq. (15.9). We will investigate both definitions of the discretized derivatives in app. A.3 and study the cut-off effects. Since the mass-dependent cut-off effects can become sizable for quarks as heavy as the charm quark, we expect significant deviations between both definitions at finite lattice spacing.

Before we describe the procedure to maximize the plateau region, we have a look at the time dependence of the current quark masses under investigation. In figure 15.7 we show the three quark masses $m_{l,l'}$, $m_{s,l}$ and $m_{s,s'}$ which have already been considered in [77, 142]. All three quark masses form long plateaus and show the expected boundary effects. Since barely any difference is visible between the two choices of the derivatives in this case, we show the quark masses from the standard definition.

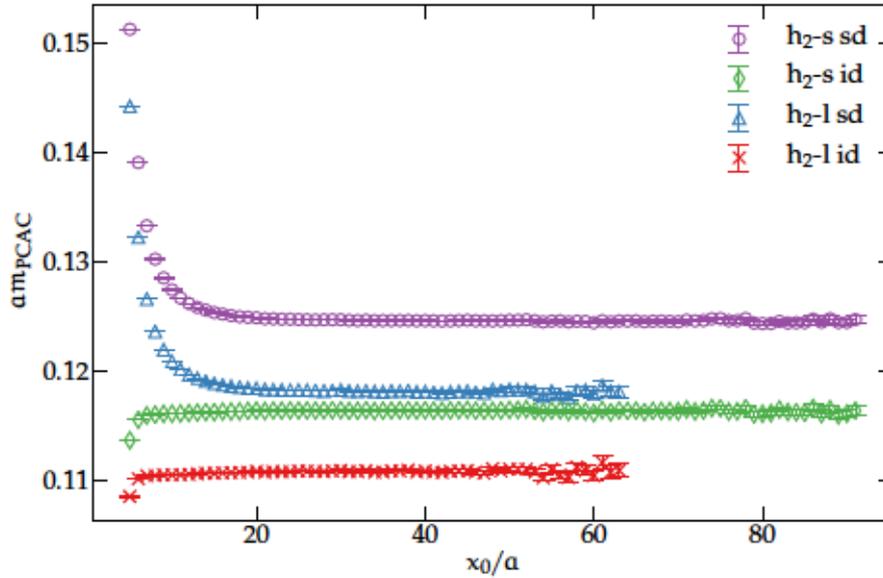


Figure 15.8: Current quark masses based on heavy and light propagators for standard derivatives (*sd*) and improved derivatives (*id*). Only data points with a relative error below 3% are shown. The hopping parameter for h_2 is chosen according to tab. 14.1.

In fig. 15.8 we show the partially quenched heavy-light quark masses. For these masses, the difference in the mass-dependent cut-off effects of standard and improved derivatives is immediately visible. Masses from the standard definition are significantly larger than from the improved derivatives. There are also considerable differences in the form of the cut-off effects for small source-sink separations. In general, the plateau region is reached faster if the improved definition is chosen. In this case, the plateau is approached from below. In the case of the standard derivative, the plateau is approached from above. Since the signal deteriorates with increasing source-sink separation we only show data points up to a relative error of 3%.

To complete the picture, we show the quark mass $m_{h_2 h_2'}$ in figure 15.9. Since the signal does not deteriorate for large source-sink separations, we are able to see that the cut-off effects due to the boundary seem to be similar for both definitions. We draw the conclusion that the cut-off effects for small source-sink separations are entirely different between both choices. The absolute difference between the

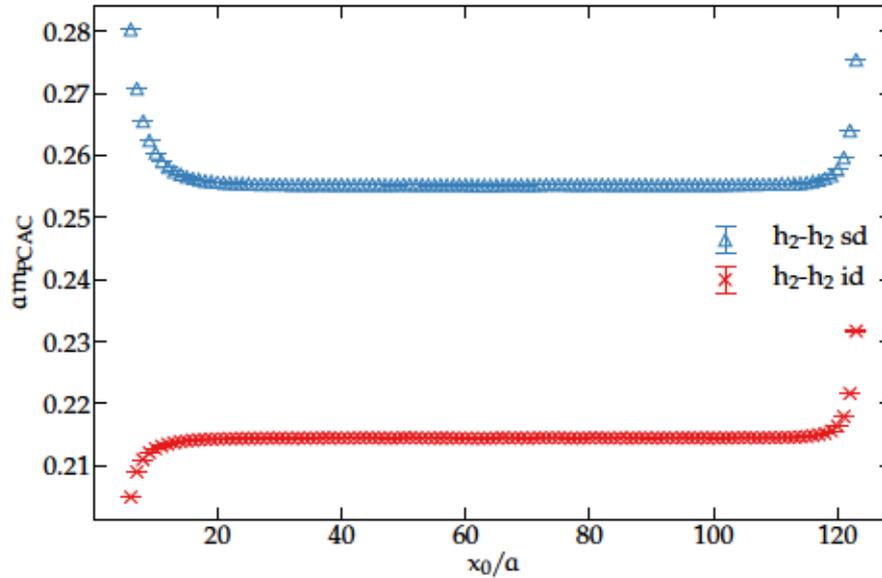


Figure 15.9: Current quark masses based on heavy propagators for standard derivatives (*sd*) and improved derivatives (*id*). The hopping parameter for h_2 is chosen according to tab. 14.1.

plateaus increases with the quark mass and becomes sizable at the scale of the charm quark.

In general, we do not know which derivative leads to a better approach to the continuum in the final extrapolation. From the absolute size of the quark masses, we can guess that the slope will be smaller for the definition based on improved derivatives. With the improvement coefficients from chapter 11 we hope to cancel $O(a)$ effects and we will investigate whether we are able to suppress higher order mass-dependent cut-off effects using LCP-1.

15.3.1 Fit ranges

Since we have to determine nine quark masses on 15 ensembles, we like to automatize the determination of the plateau region. The functional dependence of the cut-off effects is not known a priori and it is not clear how to model the deviations from a plateau in a fit. In [142], the same methods as for the effective masses, i.e., fits to exponential corrections, have been used to determine the plateau region.

Based on the experiences of [142], we employ our fit routines in the determination of the plateau region for the PCAC masses. The final determination of the current quark mass is done via a weighted fit in the plateau region, since our only theoretical expectation on the functional form of the data is a constant in the bulk of the lattice.

For the vast majority of quark masses considered in our study, the methods used for the fits of effective masses also work to determine

the start and end point of the plateau region. Since we do not expect a perfect fit result, we are not as strict concerning the bound on the ratio $\frac{\chi^2}{\chi_{\text{exp}}^2}$ and it suffices to require that a value ≈ 2 is reached to obtain a save determination of the plateau.

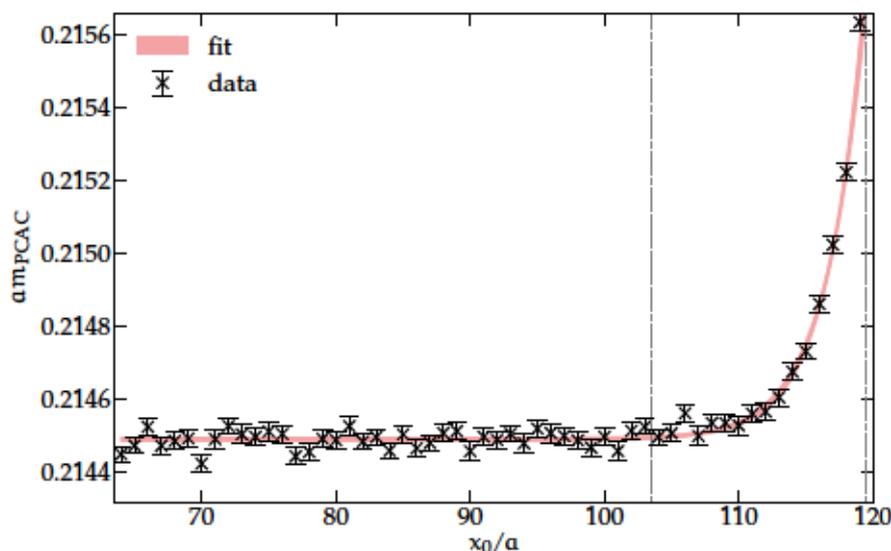


Figure 15.10: Fit to determine the upper range of the plateau for the current quark mass based on h_2 - h_2 using improved derivatives on N200. The dashed lines depict the start of the fit range and the plateau range respectively. The hopping parameter for h_2 is chosen according to tab. 14.1.

In figures 15.10 and 15.11 we show exemplary fits for the determination of the plateau range for the mass $m_{h_2h_2}$ using improved derivatives on N200. From these figures it can be seen that the boundary effects can be modeled appropriately. We define the start of the plateau range at the time slice where the effect of the exponential corrections is smaller than 0.25 times the statistical error of the data point. In figure 15.11, the best fit starts at $x_0/a = 9$ and for better visibility of the plateau region, we do not show this point in the plot. The plateau range determined from the fit overlaps with the determination from the weighted average but we discard the former one.

In the case under display in fig. 15.11, the plateau region is approached from below. As apparent from fig. 15.8 and 15.9, this behavior depends on the chosen discretization of the derivatives and the flavor combination. In figure 15.12, we show the current quark mass $am_{h_1,1}$ for the five ensembles on the symmetric line of the $\text{Tr}[M_q] = \text{const}$ CLS ensembles. For the chosen flavor combination, we can see that the form of the cut-off effects depends on the lattice spacing.

From measurements with sources in the bulk of the lattice, it is known that the cut-off effects due to small source-sink separations is roughly constant in lattice units and the effect is visible for separations

On the symmetric line, three mass degenerate sea quarks have been simulated.

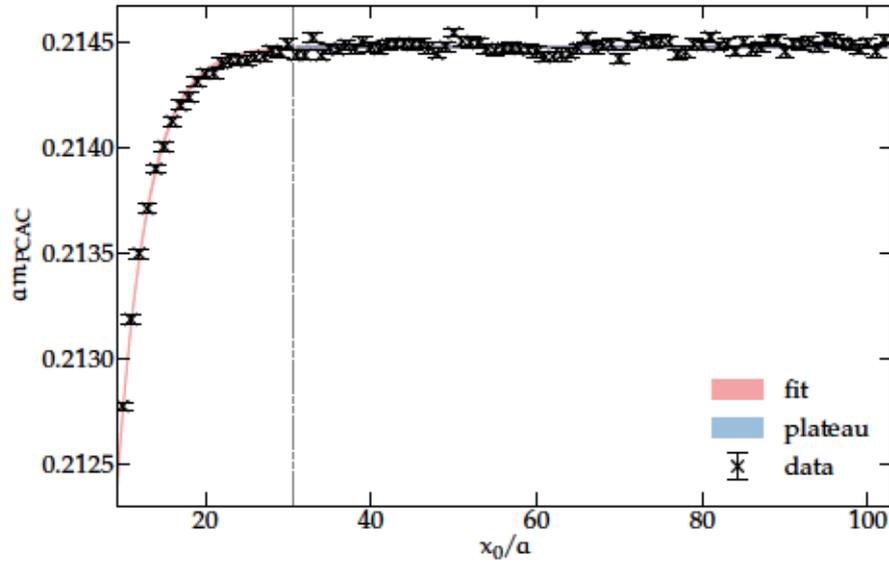


Figure 15.11: Fit to determine the lower range of the plateau for the current quark mass based on h_2 - h_2 using improved derivatives on N200. The dashed line depicts the start of the plateau range. The hopping parameter for h_2 is chosen according to tab. 14.1.

smaller than $10a$ [269]. The end of this region is marked with vertical lines in the subplots of figure 15.12. For coarser lattice spacings and improved derivatives, the effect of a short source-sink separation seems to dominate the cut-off effects. The boundary effects dominate at finer lattice spacings.

On J500, the plateau for am_{h_1} from improved derivatives is reached from above. On N300, the plateau is reached from above but a local minimum forms before the plateau is reached. This behavior is also visible for N302 and J303 have been generated with the same bare coupling. On these two ensembles, the fit procedure does not work since no exponential-like behavior can be seen in the data. Therefore, we have to set the lower end of the plateau region by hand for heavy-light and heavy-strange quark masses determined from improved derivatives on J303 and N302. Since the boundary cut-off effects reach further into the bulk if standard derivatives are used, the fit works without problems.

We collect all quark masses determined in our study in table 15.3 for standard derivatives and in tab. B.12 for improved derivatives. It can be seen that the relative statistical errors of heavy quark masses are smaller, if improved derivatives are used. The comparison of the bare light quark masses with the results in table 2 of [142] shows an excellent agreement. The quark masses are ready to be used in a chiral-continuum extrapolation. We expect both choices of derivatives to lead to the same continuum limit. It has to be tested, if a combined extrapolation of both definitions with a common continuum limit and differing cut-off effects leads an improved result at the physical point.

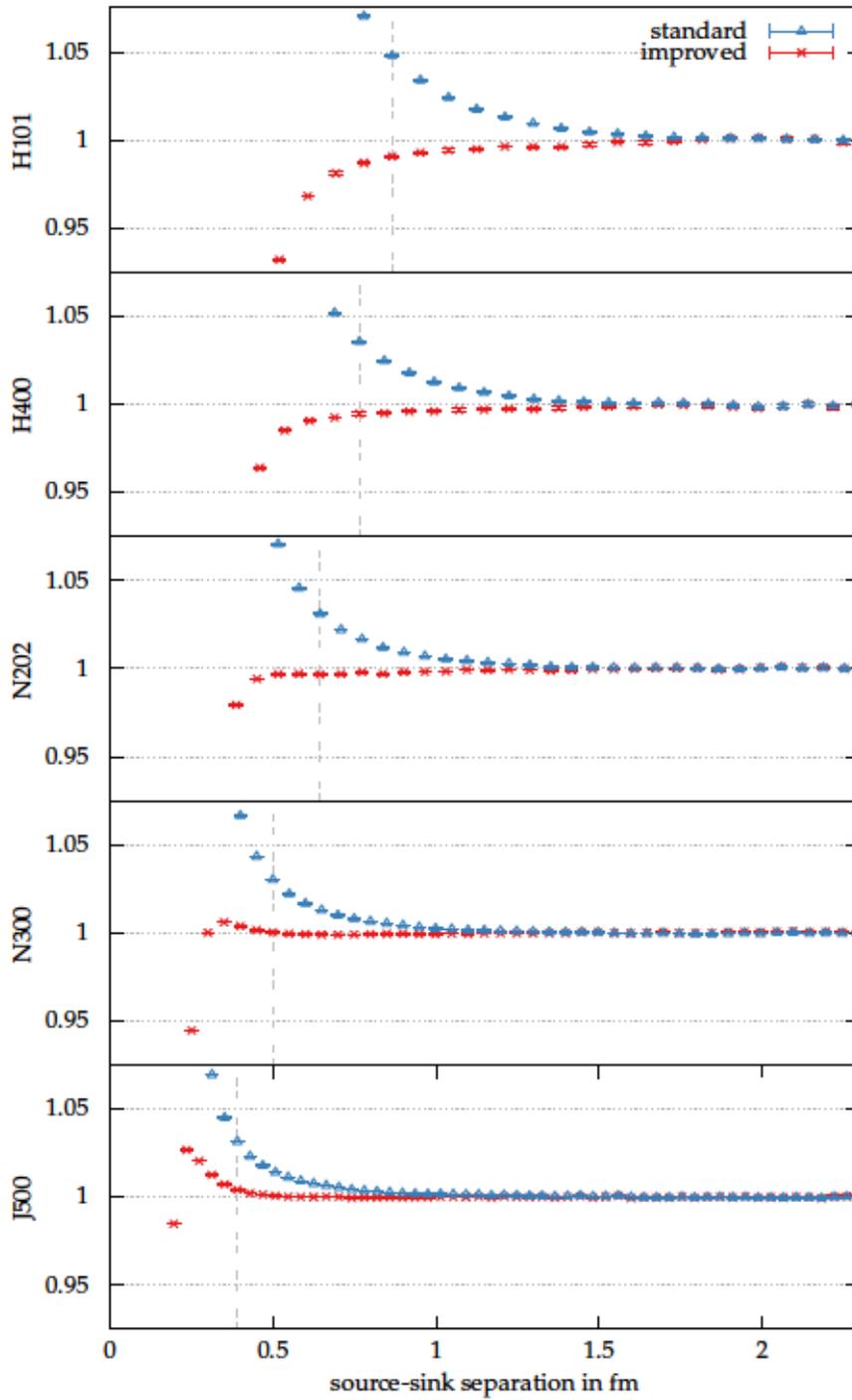


Figure 15.12: Current quark mass $am_{h_2,1}$ determined from standard and improved derivatives along the symmetric line of the CLS ensembles. To allow for a better comparison, all masses have been normalized to their plateau averages. The lattice spacing decreases from top to bottom. Vertical lines marks $x_0/a = 10$. The hopping parameter for h_2 has been chosen according to tab. 14.1.

In any case, the comparison of both extrapolations can be employed to quantify systematic effects.

15.4 DISTANCE PRECONDITIONING

We have motivated the need for a distance preconditioning of heavy quark propagators in section 5.3.6. In this section we will show the impact of this procedure on the heavy propagators in our measurements and explain how we tuned the preconditioning parameter α .

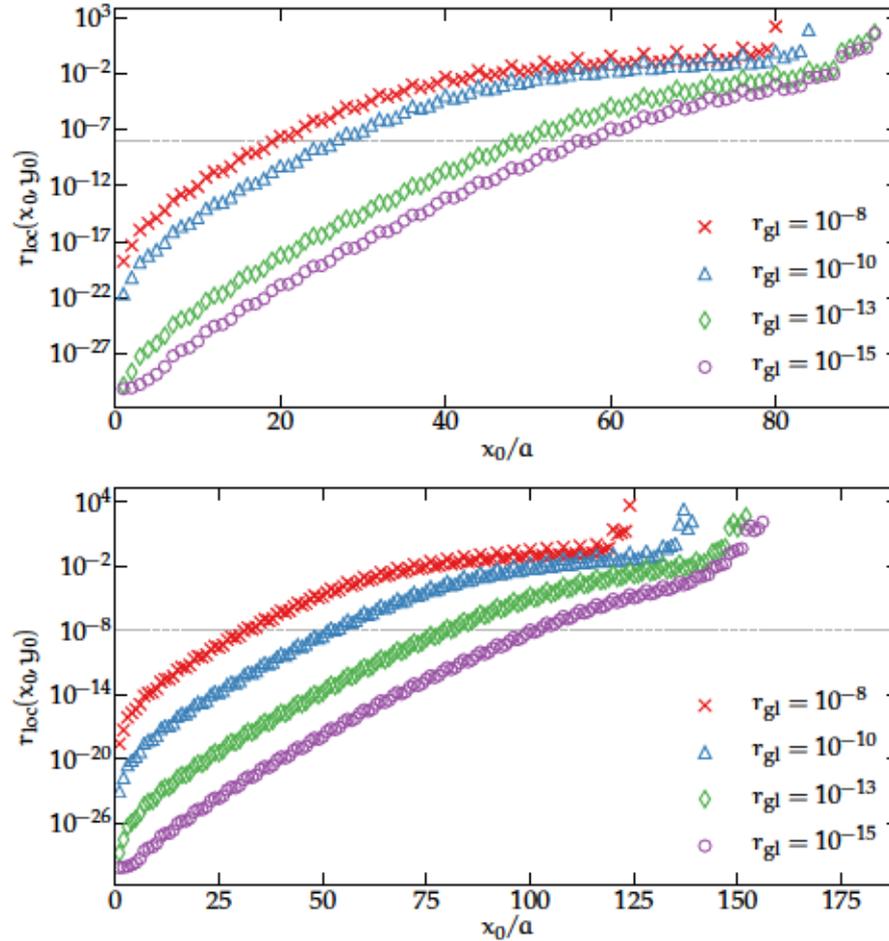


Figure 15.13: Local residuum $r_{\text{loc}}(x_0, y_0)$ for $y_0 = a$ with different global residua for the heavy propagator with the hopping parameter set to κ_{h_2} from tab. 14.1. *Top:* H400. *Bottom:* J501. The targeted precision is depicted by the dashed line.

To illustrate the need for the preconditioning, we show the local residuum $r_{\text{loc}}(x_0, y_0)$ for heavy propagators on H400 and J501 where the source is placed at $y_0 = a$ and no preconditioning has been used in figure 15.13. It can be seen that the size of r_{loc} increases exponentially and reaches values far beyond the target precision. If the global residuum is set to $r_{\text{gl}} = 10^{-8}$, the local residuum violates this bound

at source-sink separations of $O(20)$. If the global residuum is reduced, the range of time slices where the local residuum is smaller than 10^{-8} can be extended with a moderate increase of the cost.

The lower bound on possible values of r_{gl} is given by the machine precision, which is $O(10^{-16})$ in the case of double precision. Therefore, we cannot decrease the global residuum further than it is done in fig. 15.13. Despite the small residuum, the fraction that is unusable for the extraction of physical quantities grows from 40% on H400 to 50% on J500.

15.4.1 Tuning procedure

An increase in the solver precision at late time slices has to be paid with a higher iteration count and therefore with an increased need for computing time. We like to tune the parameter α such that the required precision is reached with minimal cost.

We perform this tuning on one single configuration per ensemble. On this configuration, we scan a range of values of α by inverting a heavy valence propagator with the hopping parameter chosen according to table 14.1. For each propagator, we determine r_{loc} as defined in eq. (5.75) for all source-sink separations.

We set the global residuum to $r_{\text{gl}} = 10^{-8}$ and require $r_{\text{loc}}(\frac{7}{8}T, \alpha) < r_{\text{gl}}$. With this criterion we are able to use the full plateau range of the heavy PCAC mass and risk numerical instabilities only in the region where boundary effects dominate.

In figure 15.14 we show the local residuum at $x_0 = \frac{7}{8}T$ and the number of iterations of the distance preconditioned SAP_GCR solver depending on the parameter α . We see an exponential increase of the number of iterations needed to reach the global precision given by r_{gl} when α is increased.

Starting at $\alpha \approx 0.35$, we can observe a fast decrease of the local residuum at $x_0 = \frac{7}{8}T$ when α is increased. For $\alpha > 0.42$ the local residuum stays below the global one. If the global solver residuum r_{gl} is decreased, the value of α that is needed to stay below 10^{-8} decreases as well. Nevertheless, the cost for one inversion is nearly independent of the pair (r_{gl}, α) if the same criterion for r_{loc} is used.

The value for α is nearly independent of the light quark masses but varies with the lattice spacing. In any case, the cost of the tuning procedure is negligible and it can be performed on every ensemble before the measurements are performed.

15.4.2 Impact on physical observables

To illustrate the impact of distance preconditioning on the physical observables, we repeat the measurements for one heavy quark mass on the ensemble H400 using the standard solver without preconditioning.

There is a one-to-one correspondence between the number of iterations and the computational effort.

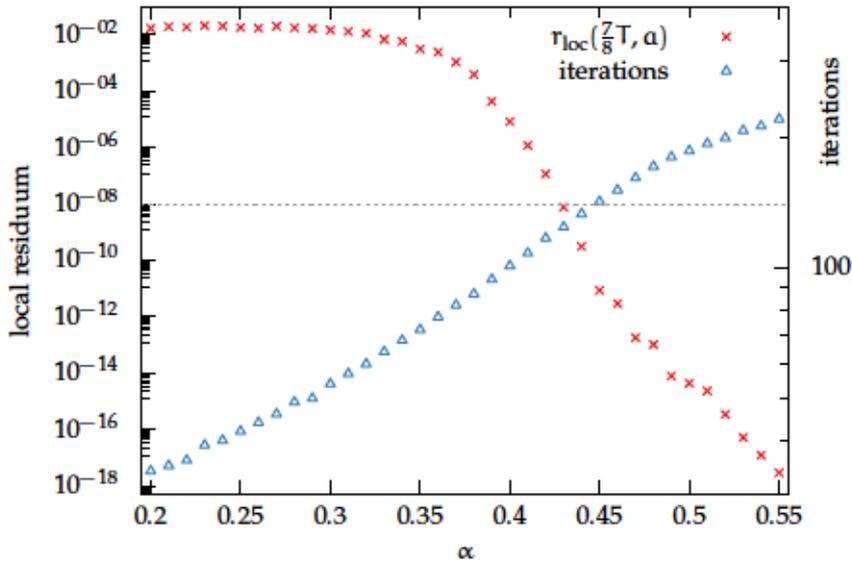


Figure 15.14: Local residuum $r_{\text{loc}}(\frac{7}{8}T, \alpha)$ on H400 for the heavy propagator with the hopping parameter set to κ_{h_2} from tab. 14.1 together with the number of iterations of the preconditioned SAP_GCR solver. The global residuum is depicted by the dashed line. Reproduced from [264].

Using these measurements, we are able to compare preconditioned data with unmodified data at the same level of statistics.

We show two exemplary observables, namely the effective mass of the pseudoscalar heavy-light meson and the current quark mass m_{hl} in figure 15.15. The lower end of the plateau region $t_{\text{min}}^{\text{P}}$ as determined from the fit with the preconditioned data is depicted by the dashed lines. We can see significant deviations of the data without preconditioning from the true result at source-sink separations which are smaller than $t_{\text{min}}^{\text{P}}$. A reliable determination of the physical observables is not possible from the data without preconditioning. If such setup is used, the local residuum has to be lowered significantly to reach the required precision at distances where excited state contributions and cut-off effects have vanished.

We did not test the influence of a smaller value of r_{gl} on observables at late time slices. The efficiency of this approach compared to distance preconditioning certainly also depends on the observable under investigation. For am_{hh} we obtain very clean plateaus which span over the whole bulk of the lattice. This would probably be impossible without distance preconditioning. The signal of heavy-light observables on the other hand gets lost in the noise at some point and it is not necessary to solve the Dirac equation properly at late time slices.

In any case, a test similar to the one shown in figure 15.13 should be performed if no distance preconditioning is used. Without this explicit check it is not possible to properly estimate the systematic

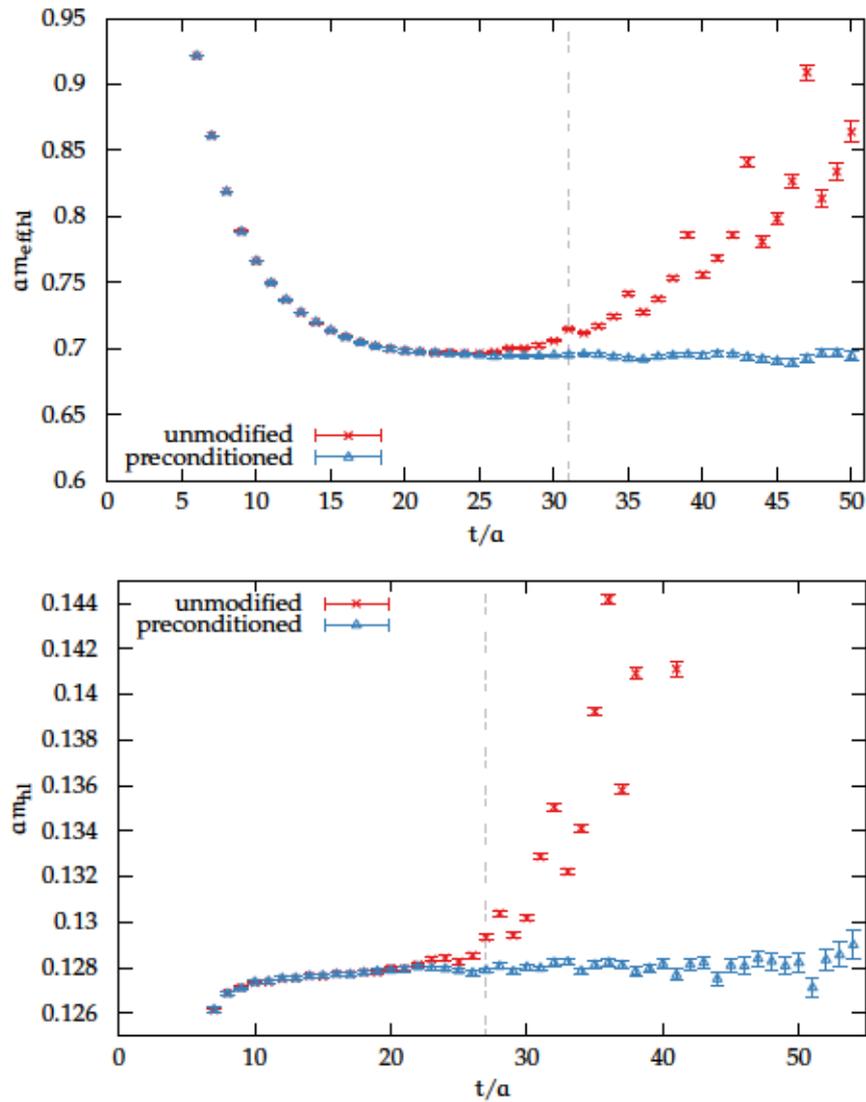


Figure 15.15: Effective mass of the pseudoscalar heavy-light meson and current quark mass m_{h_l} determined with a global residuum $r_{gl} = 10^{-8}$ with and without preconditioning depending on the source-sink separation $t = x_0 - y_0$. The lower end of the plateau region for the preconditioned data is shown by the dashed line.

effects introduced by the lack of precision at large source-sink separations. Deviations from the true values could be interpreted as local fluctuation in the data and introduce a bias in the final observable. In our opinion, the upper end of any fit interval or plateau range has to be dictated by a bound on the local residuum.

15.5 STOCHASTIC SOURCES

Since we chose to fix the source positions on the time slices a and $T - a$, we generated a number of random sources on each time slice to reduce the statistical error of our observables. It would have been possible to include sources on other time slices to avoid correlations between random sources on the same time slice. However, this approach would have resulted in a more complicated handling of the boundary effects. The use of different source positions on gauge configurations with open boundary conditions is of importance, if one is interested in entangling boundary and excited state effects, e.g., in the computation of decay constants [134].

We like to test whether correlations between random sources on the same time slice limit the gain in statistical precision which is achieved by an increase of the number of random sources. To compare the statistical precision of physical observables based on a different number of random sources, we determine the plateau range $[t_{\min}^P, t_{\max}]$ obtained from the analysis using the full statistics. In a second step we perform a weighted average in this region for different numbers of random sources. With this approach we are able to focus on the effect of the number of sources on the statistical error and we are able to neglect further complications based on systematic effects, e.g., when excited state contributions are not modeled properly by a fit based on low statistics.

In figures 15.16 and 15.17 we show the results of our analysis on the ensemble J303. We choose this ensemble because the measurements on J303 are among the most expensive ones and the optimization of the number of sources is especially important in this case. Nevertheless, the results are qualitatively the same on all ensembles. In all cases, we use time reversal symmetry, i.e., we average the same number of sources at $y_0 = a$ and $y_0 = T - a$.

On the left hand side of the figures we show the statistical error of the observable depending on the number of sources N normalized to the statistical error for $N = 1$. If no significant correlations are present, we expect this normalized error to decrease with $1/\sqrt{N}$ and for comparison, we show this curve in the figures. On the right hand side we show the evolution of the mean value and its error with the number of sources. Since we are interested in the general trend rather than the absolute numbers, we show the observables in arbitrary units.

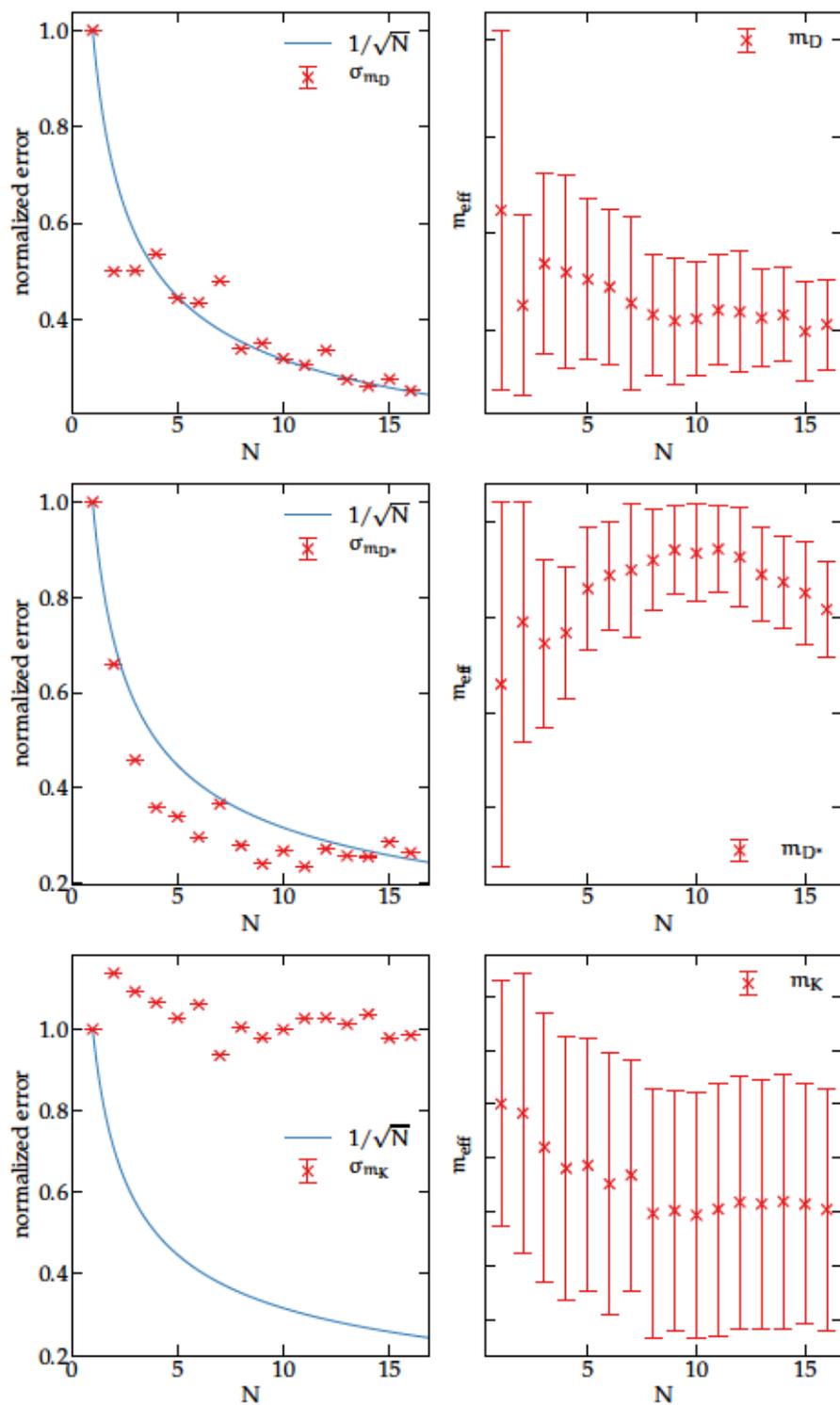


Figure 15.16: Statistical error of the plateau averages for the effective masses m_D , m_{D^*} and m_K depending on the number of random sources per source position N on J303. *Left:* Size of the statistical error normalized to the error at $N = 1$. *Right:* Mean values and statistical errors in arbitrary units.

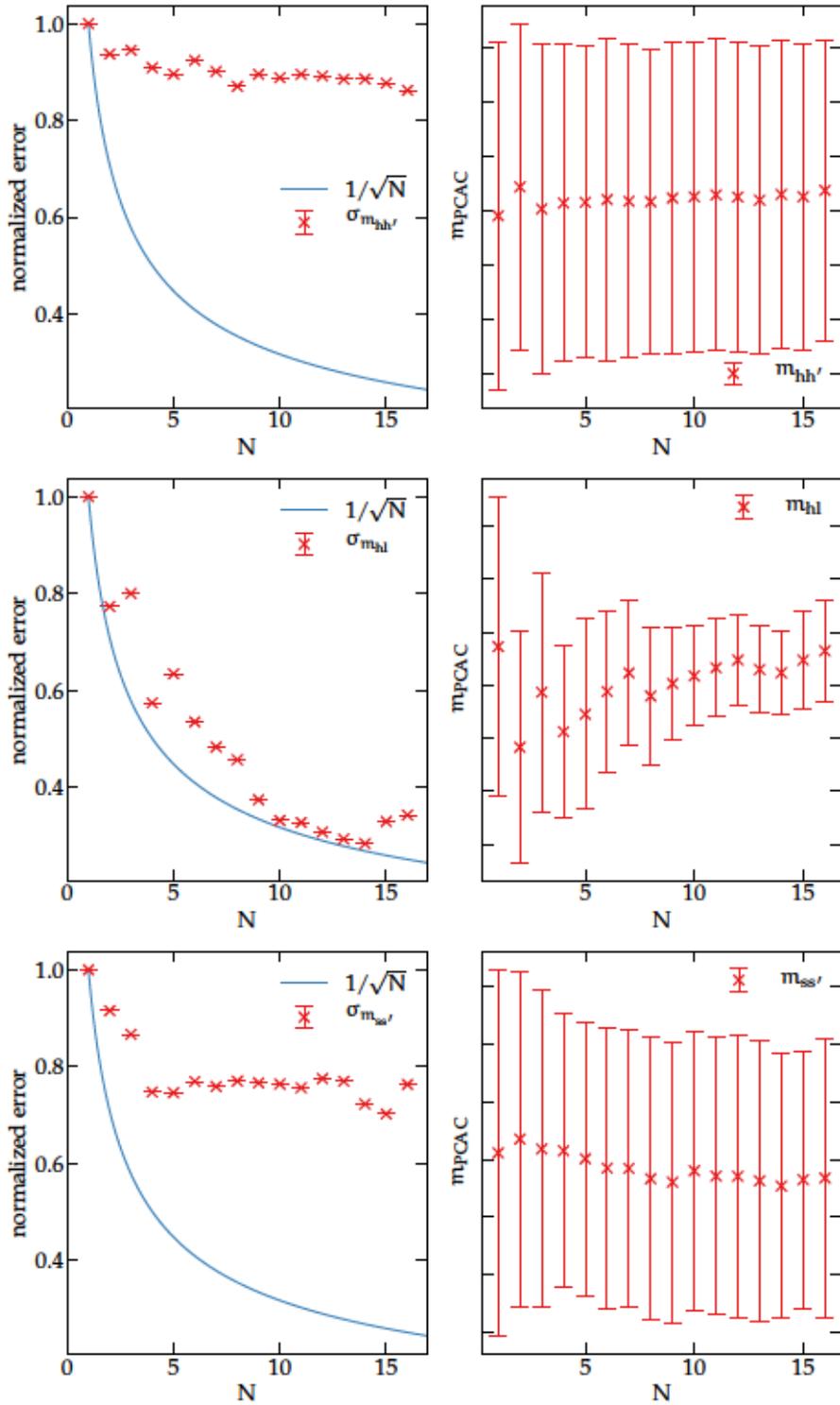


Figure 15.17: Statistical error of the plateau average for the current quark masses $m_{hh'}$, m_{hs} and $m_{ss'}$ depending on the number of random sources per source position N on J303. *Left:* Size of the statistical error normalized to the error at $N = 1$. *Right:* Mean values and statistical errors in arbitrary units.

From fig. 15.16 and 15.17 we can see that we achieve a perfect scaling of the error for all observables based on heavy and light quarks. Since we employ the heavy-light mesons in the determination of κ_c , the increase in statistics directly translates in a more precise estimation of the hopping parameter.

For the heavy-heavy observable $m_{hh'}$, we see only a small improvement by the use of a higher number of sources. In contrast to the heavy-light observables, where the signal is lost to the noise at some point, the signal-to-noise ratio for the heavy-heavy quark mass remains constant.

In all observables based on sea quarks we see only marginal effects from an increase of N . In the light observables the signal-to-noise ratio stays constant and we expect significantly larger correlations. Both effects cannot be disentangled from this analysis. In general we can judge on the size of the correlations using the size of τ_{int} for the observables at hand. The exact connection between autocorrelation in Monte Carlo time and correlation between random sources on single or multiple time slices would be an interesting field of study.

15.6 CHIRAL-CONTINUUM EXTRAPOLATION

The results for the meson and PCAC masses obtained in sections 15.2 and 15.3 enable us to perform a combined chiral-continuum extrapolation to determine M_c at the physical point. As outlined in section 14.2, we have a wide variety of different observables at our disposal to perform the extrapolation and estimate possible systematic effects. A thorough investigation will be described in [265, 266]. We conclude this chapter with a first investigation of the general aspects of the extrapolation and quote a preliminary result for the RGI charm quark mass.

To arrive at the physical point, we have to take several dependencies of the data into account. Since we have determined the meson and quark masses for two different valence quark masses that differ from the physical charm quark mass, we have to interpolate in $m_{\overline{D}}$ or $m_{\overline{D}}$ respectively to arrive at the physical charm quark mass. Since we have not determined the vector correlation functions on J501 and the spin-flavor average shows smaller statistical uncertainties, we will use the flavor average $m_{\overline{D}}$ in the following.

Since we simulate at unphysically large sea quark masses, we have to extrapolate along the chiral trajectory to physical light and strange quark masses, as outlined in section 13.1. As soon as the mass-shifts have been applied, we can parametrize the dependence of M_c on the light quark masses via ϕ_2 .

We employ the results of chapter 11 to perform the non-perturbative improvement of the PCAC masses. This enables us to perform the continuum extrapolation in a^2 rather than a . Since mass-dependent

cut-off effects are enhanced for the charm quark mass in contrast to light quark observables, e.g. in [77], we possibly have to take higher order cut-off effects into account when we include ensembles at coarse lattice spacing.

In the following we perform the chiral-continuum extrapolation for the current quark masses based on standard derivatives as they are tabulated in tab. 15.3. The running and renormalization are based on the information of [64, 66] and for the mass-dependent improvement we use the estimators of LCP-1 provided in table 11.4. We employ the data from tab. 15.2 to interpolate to the physical value of the flavor averaged D meson mass.

Since the measurements of the derivative $\langle \frac{\partial S}{\partial m_l} \rangle$ are not yet available for some ensembles, we do not perform the mass shifts described in section 13.2 at this point. However, no dependence of the charm quark mass on the light quark masses can be resolved in our chiral-continuum extrapolation. Therefore, we expect the impact of the shifts to be small.

One single parameter is sufficient to describe the dependence of the charm quark mass on the flavor averaged meson mass across all ensembles. What remains to be investigated is the lattice spacing dependence of the charm quark mass. Based on the above discussion, we choose

$$M_c \left(\sqrt{8t_0} m_{\overline{D}}, \frac{a^2}{8t_0} \right) = c_0 + c_1 \sqrt{8t_0} m_{\overline{D}} + c_2 \frac{a^2}{8t_0} \quad (15.12)$$

as basis for the chiral-continuum extrapolation. If we take only the results on for the three finest lattice spacings into account, we are able to describe the data using eq. (15.12). If we include one of the two terms

$$+ c_3 \left(\frac{a^2}{8t_0} \right)^{\frac{3}{2}}, \quad + c_4 \left(\frac{a^2}{8t_0} \right)^2, \quad (15.13)$$

to incorporate higher order cut-off effects, we are able to perform an extrapolation using all ensembles considered in our study.

In figure 15.18 we show $M_c(m_{\overline{D}}, a)$ determined from $m_{R,c}^{(c)}$, $\overline{m}_{R,c}$ and $m_{R,c}^{(rd)}$ defined in eqs. (14.5–14.8) for all ensembles together with combined chiral-continuum fits based on eq. (15.12) with an additional term of $O(a^4)$. We show the quark masses in dependence of $a^2/8t_0$ and $\sqrt{8t_0} m_{\overline{D}}$ and to allow for two-dimensional representations of the data, all other dependencies have been projected out in these plots, based on the fit coefficients.

The dependence of the charm quark mass on the mass of the flavor averaged pseudoscalar meson is shown in the upper part of fig. 15.18. The plot illustrates that the initial choices for the heavy quark hopping parameters of tab. 14.1 were tuned well enough such that a linear scaling in the region around the physical meson mass can be observed.

We list the values for t_0/a^2 in tab. B.8.

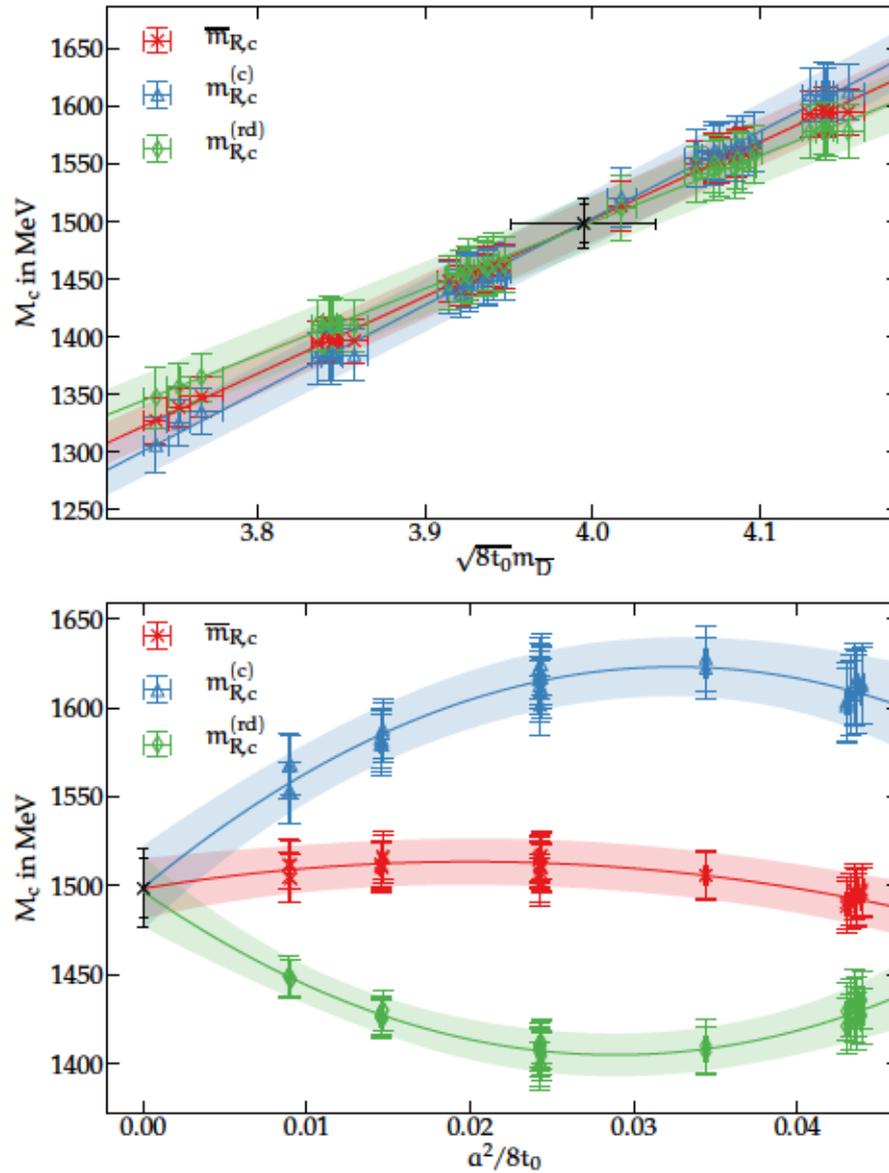


Figure 15.18: Chiral-continuum extrapolation of the charm quark mass. *Top:* Dependence of the charm quark mass on m_D . *Bottom:* Dependence of the charm quark mass on the lattice spacing. To obtain two-dimensional representations of the data, other dependencies have been projected based on the fit coefficients.

The dependence of the charm quark mass on the lattice spacing is shown in the lower part of fig. 15.18. As anticipated, the three definitions of the charm quark mass show entirely different cut-off effects. These effects are most suppressed for $\overline{m}_{R,c}$, where we observe an almost flat behavior in the squared lattice spacing. The inclusion of the term of $O(a^4)$ leads to coinciding results at the physical point. If a term of $O(a^3)$ is chosen instead, the fit quality is equally good but the mean values of the continuum results differ, albeit well within their statistical errors.

From the interpolations in fig. 15.18, we can conclude that the ensembles at $\beta = 3.70, 3.85$ are vital for the determination of the correct continuum limit. Their inclusion constrains the extrapolation close to the continuum limit and correspondingly the deviations between different definitions of the quark mass and various sets of improvement coefficients turn out to be small.

15.7 PRELIMINARY RESULT AND DISCUSSION

Since the definition of the charm quark mass based on $\overline{m}_{R,c}$ exhibits the smallest statistical errors, we use it to determine our preliminary result

$$M_c = 1499(22) \text{ MeV} \quad (15.14)$$

for the RGI charm quark mass at the physical point. The dominant contributions to the error are based on the $\approx 1\%$ relative errors of the physical value of t_0 from [142] and the running factor from [64]. Since the error of the latter is only added after the chiral-continuum extrapolation, we indicate the size of the error before the addition by the inner pair of markers on the black cross in fig. 15.18.

The detailed analysis of systematic effects in the chiral-continuum extrapolation will lead to the addition of a systematic error to our final result. First investigations show that these effects are small. At the same time, we could significantly decrease the statistical error, if the external quantities t_0^{phys} and $M/m_{R,\text{had}}$ were determined to higher precision.

Stochastic sources and distance preconditioning allow us to determine meson and quark masses to high precision and the computation of the renormalization constants and improvement coefficients of chapter 11 allows us to cancel the leading cut-off effects, thereby reducing the systematic uncertainties of the chiral-continuum extrapolation. Systematic effects due to the setup of the chiral trajectory will be canceled by the introduction of mass shifts. However, since no dependence of the charm quark mass on the light quark masses is resolvable, we believe that the effect of this procedure will be negligible.

Our preliminary result is compatible with the $2 + 1$ flavor results of [210, 254, 255] and the global average [209]. Given our statistical error

Without the error of the running factor, we obtain $M_c = 1499(17)$ MeV.

Currently, there is an effort to increase the statistical precision of t_0^{phys} by the RQCD collaboration.

and preliminary results for the systematic uncertainties, we will be able to reach a similar precision as the other studies in the $2 + 1$ flavor theory.

Part V

HQET IN LARGE VOLUME

HEAVY QUARK EFFECTIVE THEORY WITH OPEN BOUNDARY CONDITIONS

In chapters 8 and 12 we have described how we set up the line of constant physics for the non-perturbative matching of QCD and HQET. This matching is needed to renormalize the effective theory which is then used for the extraction of physical observables on large-volume ensembles.

Such computations in the effective theory have already been performed by the ALPHA collaboration in the quenched [148, 271, 272] and the two-flavor theory [211, 239, 273, 274]. For our computations in the theory with $N_f = 2 + 1$ quark flavors, we will use the CLS ensembles described in chapter 14. Since open boundary conditions in the time direction instead of periodic ones are used for this set of ensembles, the strategy for the computations has to be adapted accordingly.

To perform the calculations of the static and light quark propagators in the presence of open boundary conditions, a measurement program based on the solvers of the openQCD package has been developed. All types of smearing that have been introduced in section 6.5 are implemented in openHQET and we will employ them in the following.

In this part of our thesis we will discuss some of the experiences and test that we have made on our way to set up the calculation of heavy quark observables such as the b quark mass, the decay constants of the B and B_s mesons and the form factors of semi-leptonic B meson decays. For all of these observables, the input from the matching procedure is necessary to obtain physical values in the continuum limit.

Later on we will also discuss preliminary results for the computation of the static $B^*B\pi$ coupling, which is defined in the static theory and can be calculated without input from the matching. We will work closely along the lines of the determination in the quenched and two-flavor theories [275].

16.1 SETUP OF THE MEASUREMENTS

We use the $2 + 1$ flavor CLS ensembles on the $\text{Tr}[M_q] = \text{const.}$ trajectory for our measurements of matrix elements in the effective theory. Since no relativistic heavy valence quarks are simulated, the discretization effects for heavy-light observables in HQET are known to be rather mild. Therefore, we do not take into account the two ensembles at $\beta = 3.85$. With the remaining four different lattice spacings we

An overview of the ensembles is given in fig. 13.1.

are able to perform a controlled continuum extrapolation and to study systematic effects by performing cuts on the lattice spacing.

In exchange, heavy-light observables have a relatively strong dependence on the mass of the sea quarks [211, 239, 275]. To control the chiral extrapolation, we take into account ensembles with light quarks as light as the physical ones, i.e., $m_l \geq 134$ MeV. Since the trace of the quark mass matrix is held constant, the strange quark also approaches its physical value in the extrapolation.

On this set of ensembles, we determine heavy-light correlation functions in the effective theory. To improve the signal-to-noise ratio of the observables, we use HYP1 and HYP2 smearing for the construction of the static propagator. Since no inversion is needed, the overhead of considering both smearings consists of a doubled number of contractions and an enlarged memory consumption. The overhead in computing time is negligible. If one of both variants is superior in the reduction of the statistical error, we will use it to obtain the final results. Since observables determined from HYP1 and HYP2 differ in their cut-off effects, a comparison between both smearings offers a test for the absolute size of discretization effects.

Sea quark propagators are determined using the deflated solver of the openQCD package and only the use of deflation allows us to consider sea quarks with physical quark masses in the first place. We use time-diluted random $U(1)$ sources for construction of quark propagators. On each time slice in the bulk of the lattice, we use N_τ random sources to improve the statistical precision of our observables. N_τ is adapted such that the optimal precision of the final result is achieved using the available computational resources. The region where boundary effects are absent has to be explicitly determined and we investigate boundary effects in section 16.3.

For the construction of the variational basis of the GEVP, we consider three different types of smearing, namely Gaussian smearing and smearing using three-dimensional spinor and scalar fields, as described in section 6.5. For the covariant derivatives used in the smearing procedures, we use gauge links that have been triply APE smeared in the spatial directions. From each smearing procedure, we construct four smeared interpolators. The smearing parameters, i.e., the number of iterations in the case of the Gauss smearing and the three-dimensional hopping parameters for spinor and scalar smearing have been tuned such that the same smearing radii are obtained for all three types of smearing.

16.2 OPTIMIZATION OF THE VARIATIONAL BASIS

We have introduced the effective energies E^{stat} , E^{kin} and E^{spin} in section 7.4 and motivated their extraction with the determination of the bottom quark mass. Since the signal-to-noise ratio of heavy-light

HYP smearing has been introduced in section 6.5.2.

We use $\alpha_{\text{APE}} \approx 0.62$.

observables is exponentially decreasing, it is of great importance to be able to determine the energies from plateau fits that start at small source sink separations. We have introduced the use of the GEVP to achieve an effective suppression of systematic effects from excited state contributions.

To construct a matrix of correlation functions based on different interpolators, we use the smearing procedures described in section 6.5. The computational cost for the smearing is not negligible compared to the cost for the solution of the Dirac matrix and it scales with the spatial extent of the smearing radius. In the case of Gaussian smearing, we have to apply $O(100)$ iterations to reach $r \approx 0.45$ fm at the coarsest lattice spacing used in our study. In the case of smearing with three-dimensional spinor and scalar fields the cost of the inversion of the three-dimensional equations scales with the smearing radius.

The construction of correlation functions of smeared interpolators is less expensive in HQET as it is in QCD, since an additional inversion per smeared interpolator is necessary in the full theory. This is not the case in the effective theory since the static propagator in the heavy-light correlation functions is determined from the gauge field via eq. (7.37) without an inversion of the Dirac matrix.

From four smeared interpolators per smearing procedure and one local interpolator we can build a 13×13 correlation matrix for all correlation functions introduced in section 7.4. The inversion of such correlation matrix in the GEVP at finite t_0 likely fails due to the statistical uncertainties in the data. Therefore, we have to build a subset of operators that is well suited for the extraction of the energies. In the studies in the two-flavor sector only Gaussian smearing has been used to construct the correlation matrix. Therefore we have to test the performance of the smearing using three-dimensional scalar and spinor fields. In [188] the stability of the GEVP using these smearings has been checked.

We investigate the performance of the different smearing types and smearing radii on the ensemble H101, where the measurements are cheap. The ensemble is on the symmetric line, i.e., light and strange quark masses are degenerate, and at the coarsest lattice spacing at $\beta = 3.4$. Based on the experiences made on this ensemble, we set up the measurements on the other ensembles. We have checked that the results obtained in this section also hold on the ensemble H200, which features a smaller lattice spacing.

16.2.1 Definition of a suitable basis

The definition of a suitable basis for the GEVP is ambiguous. In principle, the generalized eigenvalue method is used to project the variational basis onto the physical states. If the basis operators are linearly dependent, the quality of the projection deteriorates. Large

statistical fluctuations also have a negative impact on the projection. Therefore, the quality of the results depends on the set of basis vectors.

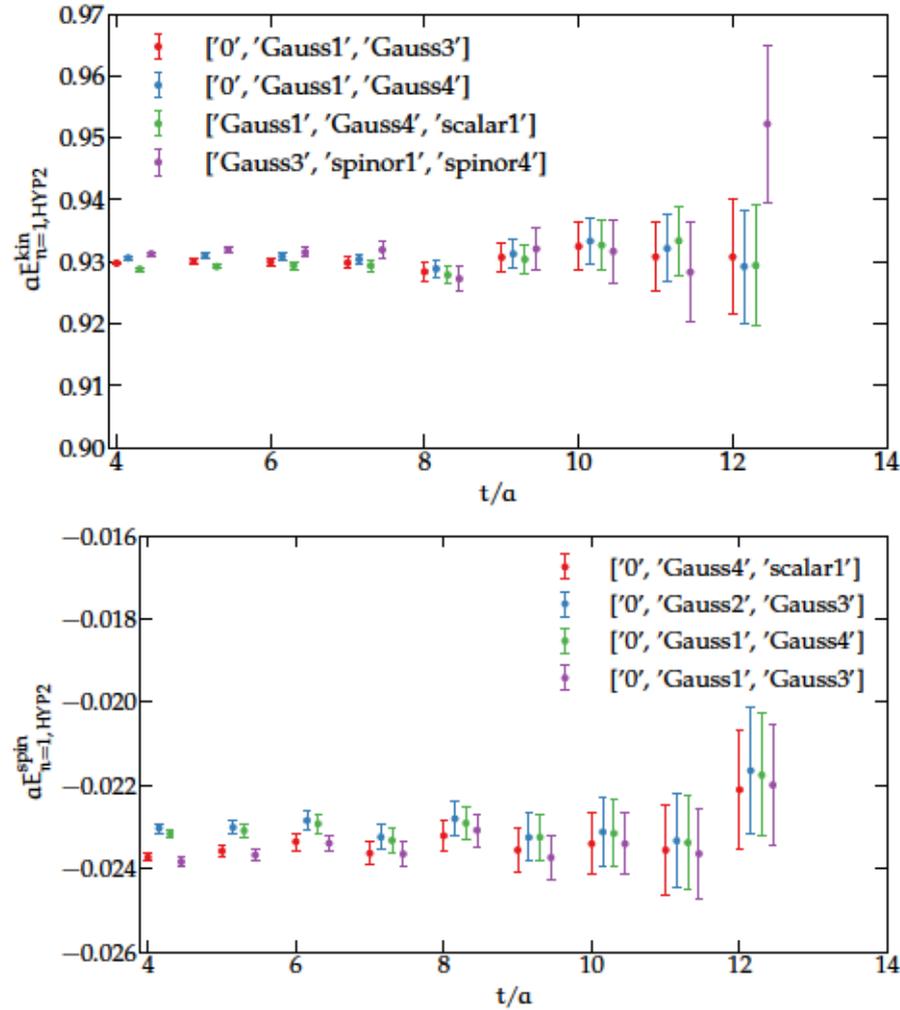


Figure 16.1: E_1^{kin} and E_1^{spin} from the variational method using the four combinations that yield the best plateaus. Data points have been shifted horizontally to improve visibility.

To judge on the quality of different sets, we determine the energies $E_n^{\text{stat}}, E_n^{\text{kin}}, E_n^{\text{spin}}$ for $n = 1, 2$ from all possible 3×3 correlation matrices that can be built as submatrices of the 13×13 correlation matrix. To quantify the quality of one set of operators, we determine the maximal fluctuation of the mean values of $E_n(t)$ in a range $[t_{\text{min}}, t_{\text{max}}]$ which is adapted to the overall precision in the determination of the different energy levels. We define the best combinations as the ones where the fluctuations are minimal. For these combinations the deviations from a plateau behavior due to excited state contributions are small and fluctuations at larger source-sink separations are suppressed.

For simplicity we choose $t_0 = 4$ for all of the following tests. This allows us to judge on the pre-plateau behavior and at the same time

We have tested different criteria to quantify the quality of the plateau.

to suppress systematic effects up to intermediate values of t/a , where we test the quality of the plateaus. We choose HYP2 for the static propagator.

In the following, we will show some representative results of our investigations. In our plots, we label the 13 different interpolating operators as follows.

- The local operator θ
- Gaussian smeared operators: Gauss, $i = 1 \dots 4$
- Operators constructed from 3D scalar fields: scalar, $i = 1 \dots 4$
- Operators constructed from 3D spinor fields: spinor, $i = 1 \dots 4$.

Interpolating operators with the same value of i are tuned to have the same smearing radius $\langle r^2 \rangle$ and $i = 4$ denotes the most smeared operators.

The results for the best combinations are somewhat ambiguous since they are not independent of t_{\min} and t_{\max} . However, qualitatively our conclusions remain the same when the plateau region is varied. For the ground states of all energies under investigation we obtain clean plateaus. The situation is more difficult for the excited states and more work has to be invested to determine a reliable plateau average.

In figure 16.1 we show the time dependence of the ground states of the effective spin and kinetic energies obtained from the four best combinations of operators. Depending on the set of operators, different contributions of excited state contamination seem to be present in the data. When these have decayed, the results of all combinations overlap. This behavior could be used for a first estimation of the point, where systematic effects are smaller than the statistical error.

For the two energies in fig. 16.1, the combination of Gaussian smeared interpolators together with the local interpolator seems to be superior. In contrast, spinor and scalar smearing perform better for the static energy. In all cases, a mixture of two types of smearing or the inclusion of the local operator lead to the best result.

If $2t_0 \geq t$ is chosen, the size of the systematic contributions can be quantified analytically.

16.2.2 Reduction of the number of smearing procedures

Since measurements including all three types of smearing are costly, we investigate whether we can neglect one of the three smearings to save computing time. The reduction to a 9×9 correlation matrix also saves a factor of two in memory consumption. We investigate the best results obtained from the three different permutations of two types of smearings and compare them with the best result from all three combinations.

We indicate the different sets by their color, according to

- Gauss, scalar, spinor

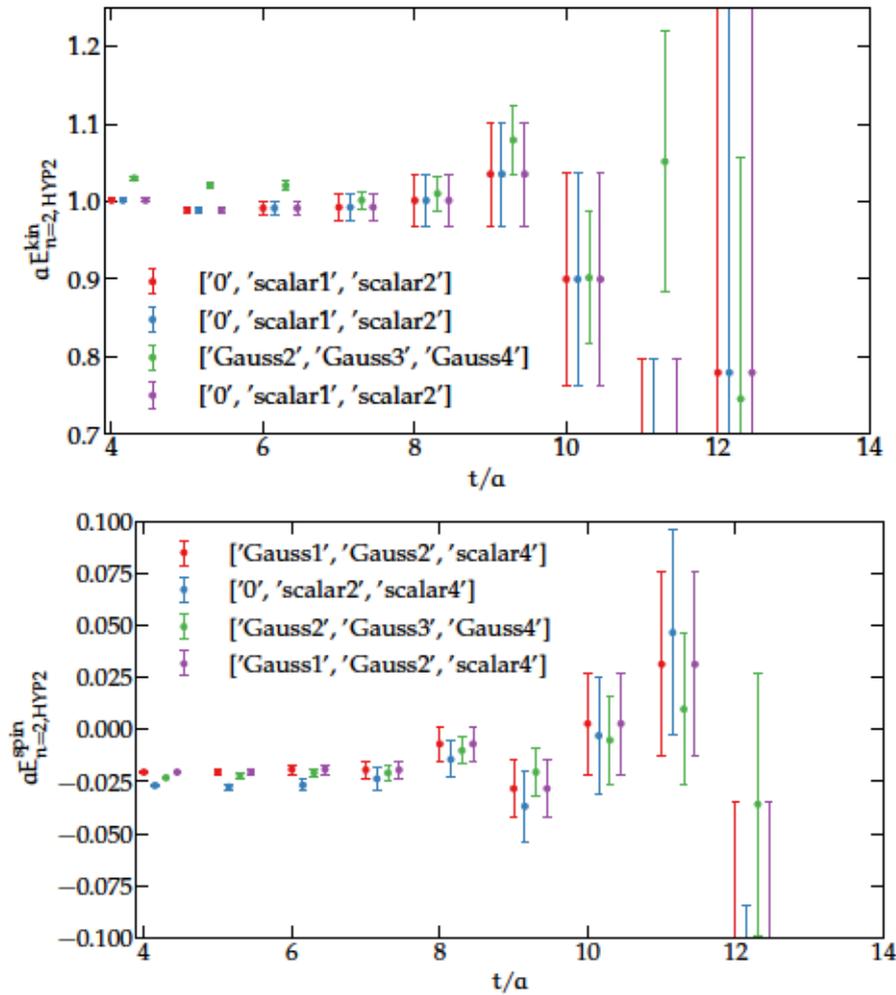


Figure 16.2: E_2^{kin} and E_2^{spin} from the variational method using the four combinations that yield the best plateaus using only two different types of smearing. Data points have been shifted horizontally to improve visibility.

- scalar, spinor
- Gauss, spinor
- Gauss, scalar

and in all of these cases, the inclusion of the local operator is allowed. We show the best combinations within these categories for the excited states of spin and kinetic energy in figure 16.2. We note in passing that for both excited states, the signal is lost to noise relatively soon. For both energies, the spinor smearing does not seem to play a role and it is not included in any combination.

Of course, this suppression can happen by chance.

In the case of E_2^{kin} the inclusion of scalar smearing suppresses best the systematic effects but the statistical errors seem to be smaller if Gaussian smearing is used. For E_2^{spin} we cannot see any qualitative difference between scalar and Gaussian smearing. We do not show the

static energy, where the inclusion of spinor smearing leads to good results, because the signal for the static energy is good for a large number of combinations. We conclude that, although it has good renormalization properties, spinor smearing has the worst performance among the three different methods under investigation.

16.2.3 Variational basis from smearing method

In the last part of our study of the operator basis, we investigate the extraction of effective energies without a mixture of smearings in the correlation matrix to analyze whether the signal is improved when several smearing methods are used. We look at the best combinations using one single type of smearing and compare them with the best combination including all operators:

- Gauss, scalar, spinor
- Gauss
- scalar
- spinor

We show the results from the best combinations for the ground state and the first excited state of the static energy and the ground state of the spin energy in figure 16.3. We see that the effective energies built from spinor smeared operators alone show a slow convergence towards the plateau region and large statistical uncertainties. The data points obtained from Gaussian smearing have the smallest statistical uncertainties. In all three cases, the best combination includes two or more smearing procedures.

We can conclude that it is beneficial to construct the correlation matrix from two different types of smeared interpolating operators for the extraction of the energies. Since smearing from three-dimensional spinor fields performs worst, we will use Gaussian smearing and the smearing from three-dimensional scalar fields for our measurements on ensembles with larger volumes, finer resolutions and smaller quark masses.

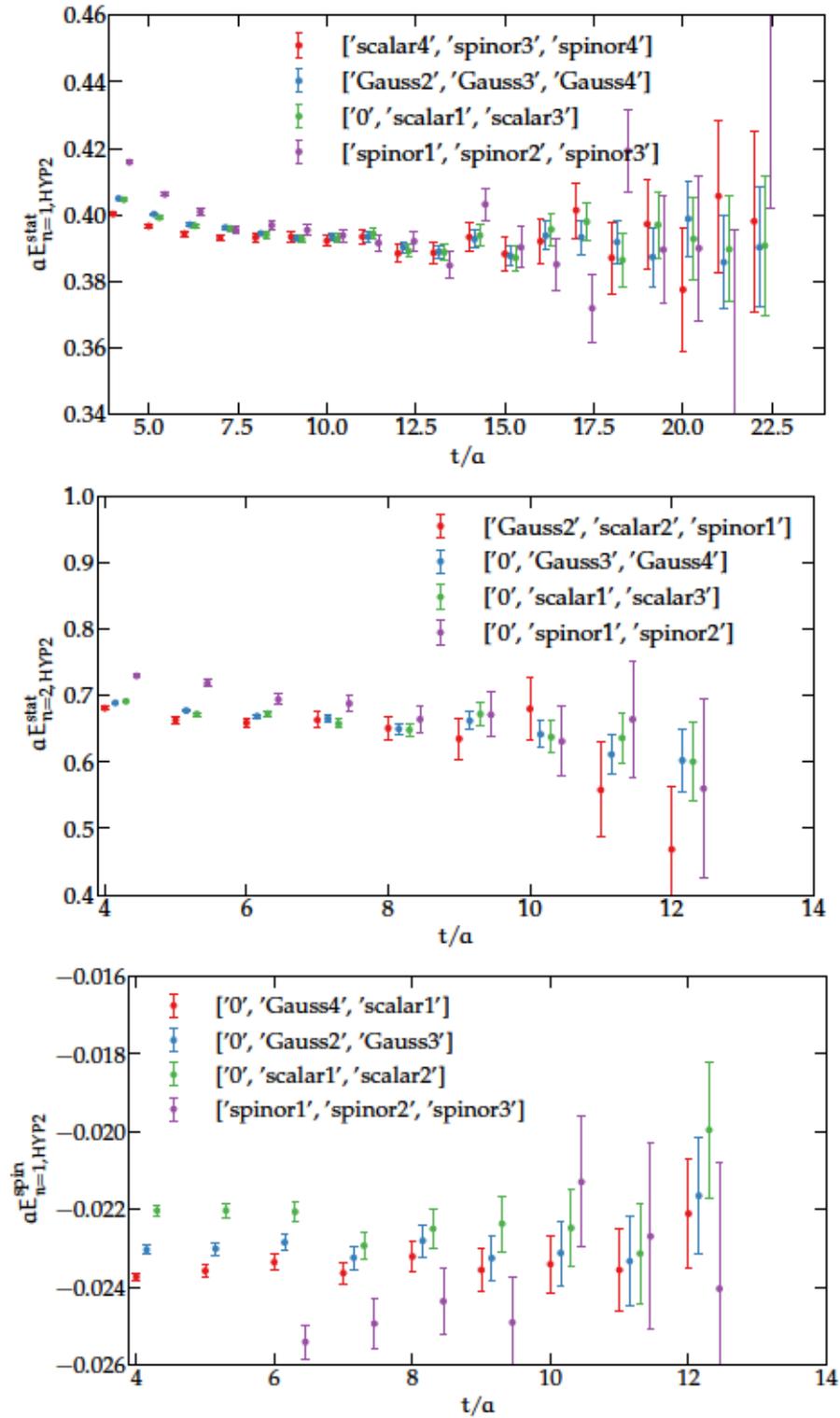


Figure 16.3: E_1^{stat} , E_2^{stat} and E_1^{spin} from the variational method using the four combinations that yield the best plateaus using only one type of smearing. Data points have been shifted horizontally to improve visibility.

16.3 BOUNDARY EFFECTS

In the quenched and two-flavor studies, translational invariance in time direction has been used to improve the statistical precision by summing over the results obtained from random sources on all time slices. The final results therefore have been determined on $T/a \times N_T$ random sources. Since we work on ensembles with open boundary conditions, we cannot take into account all time slices for two reasons.

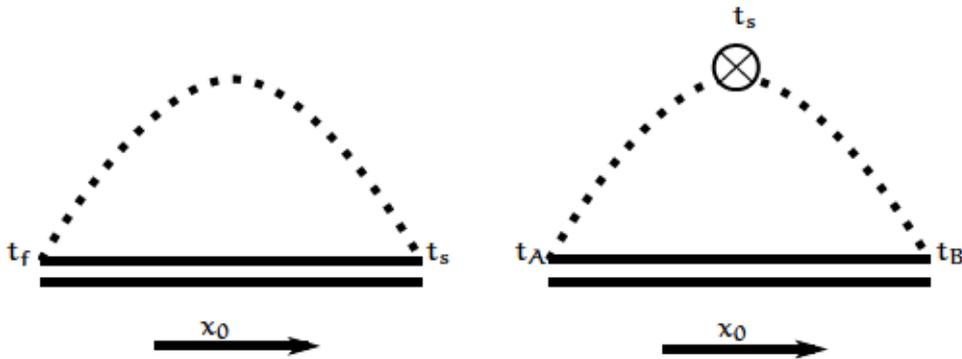


Figure 16.4: Two- and three-point correlation functions determined in this work. Relativistic quark lines are dashed and the heavy quark propagator is indicated by the double line. The source position is denoted by t_s and the positive time direction is indicated by the arrow.

On the left side of figure 16.4 we show the schematic contraction for the heavy-light two-point functions determined in the effective theory. The heavy quark propagator is depicted by the double line and the heavy quark propagates only in positive time direction. Due to this property, the correlation function can only be evaluated on one side of the source. In our case, as shown in 16.4, the source-sink separation is given by $t = t_s - t_f$. Due to the boundary at $x_0 = 0$, the maximal source-sink separation is given by t_s . Therefore, the boundaries introduce a technical constraint on the number of time slices.

Another constraint is introduced by the physical properties of the boundaries, which have the quantum numbers of the vacuum. A contamination due to boundary states can pose a problem in the GEVP since these are present in the eigenspace of the correlation matrix. If the GEVP is solved on each time slice, the correct states are obtained. But since the statistical fluctuations are large in this case, the inversion of the correlation matrix fails at relatively small values of t_0 . If we average the correlation functions over all time slices to stabilize the GEVP, the correct projection is not guaranteed anymore. Therefore we have to ensure that all contributions from the vacuum have decayed at t_s and t_f .

Since we believe that results from noise sources on different time slices are less correlated than from two random sources on the same time slice, we like to maximize the number of time slices included in the measurements. In this section we will try to determine the region where boundary effects are negligible. In the final analysis, we will introduce additional cuts on the number of time slices which are included in the summation to ensure that no systematic effects due to boundary contributions are present.

16.3.1 Boundary states

We can consider the boundary as source with the quantum numbers of the vacuum. Therefore, the spectral decomposition of the heavy-light two-point function is given by [143, 276]

$$f_{\text{hl}}(t_s, t_f) = A \exp(-E_{\text{hl}}(t_s - t_f)) [1 + B \exp(-E_b(T - t_s)) + \dots], \quad (16.1)$$

when the source position t_s is close to the boundary and the sink position is at t_f . E_{hl} is the energy of the heavy-light bound state and E_b is the lowest state with the quantum numbers of the vacuum. Keeping $t \equiv t_s - t_f$ constant but varying the source position, we arrive at

$$f_{\text{hl}}(t_s) \approx A \exp(-E_{\text{hl}}(t)) [1 + B \exp(-E_b(T - t_s))] \quad (16.2)$$

$$\equiv C + D \exp(-E_b(T - t_s)), \quad (16.3)$$

where the dependence on the fixed value of t is only implicit. We can investigate the functional behavior for different values of $t = t_s - t_f$. When we fix the ground state energy of the vacuum to $E_b = 2m_\pi$, we can perform a two parameter fit, provided that we are in a region where all excited boundary state contributions have vanished.

In figure 16.5 we show $f_{\text{hl}}(t_s)$ constructed from the local operator together with a fit to eq. (16.3), where we have chosen $f_{\text{hl}} = C_{\text{AA}}^{\text{stat}}$. The upper range of the fit region and t have been tuned such that a good fit quality is obtained for a one-state fit. The pion mass has been fixed to the value given in table 15.2. The dashed line denotes the time slice where the relative contribution of the exponential to $f_{\text{hl}}(t_s)$ is smaller than one per mille.

The distance of this timeslice from the boundary at $x_0 = T$ is 1.5 fm for H101 and 1.4 fm for H200. We conclude that we have to restrict our measurements to the region $x_0 \in [1.5 \text{ fm}, T - 1.5 \text{ fm}]$. This translates directly to a cut on t_s at the upper end of this range. For the lower end, the minimal source position depends on the length of the plateau region. Since the signal for the quantities under investigation is lost relatively soon, these plateaus are rather short. We choose $t_{s,\text{min}} \approx 2.5 \text{ fm}$ for our measurements and perform cuts in the final data to ensure that no boundary effects are present.

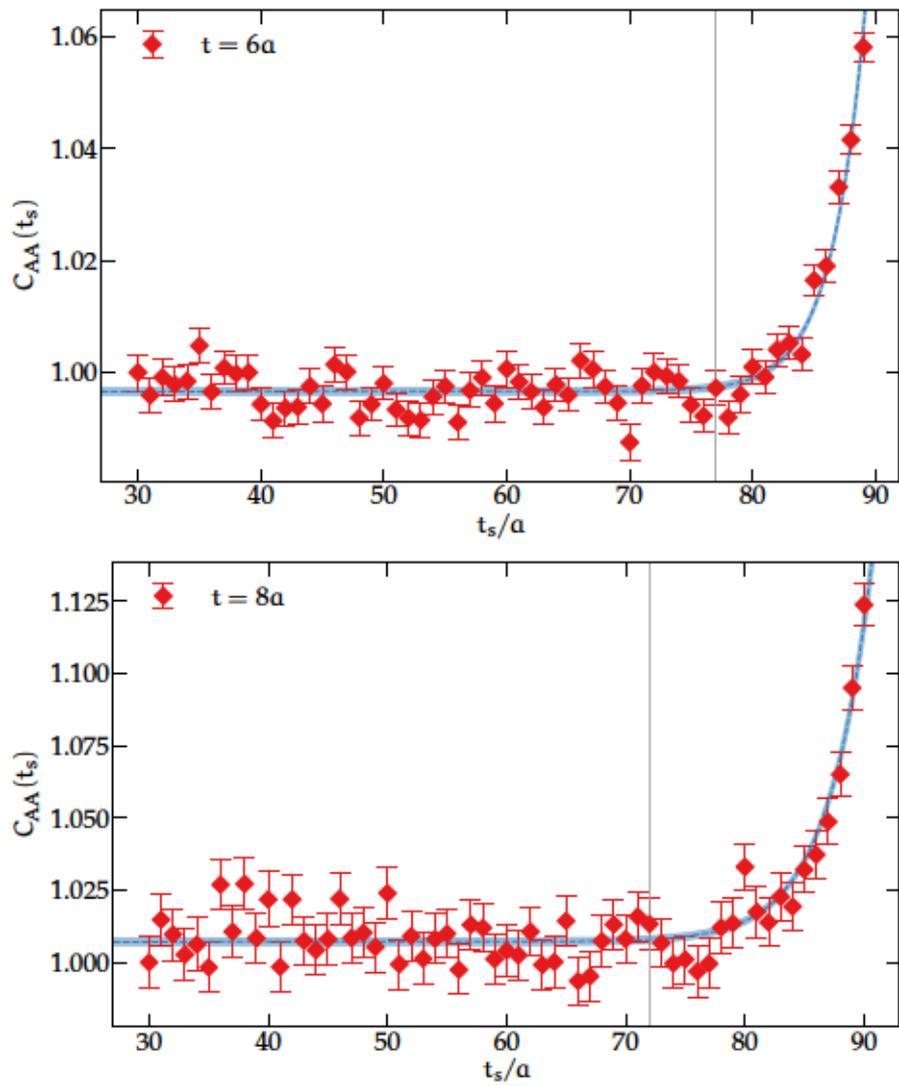


Figure 16.5: Correlation function $C_{AA}(t, t_s)$ for fixed value of t normalized by $C_{AA}(t, t_s = 30a)$ for H101 and H200 together with the fit to eq. (16.3). The dashed line denotes the time slice where the exponential corrections have decayed.

16.3.2 Asymmetry in the correlation matrix

We expect an asymmetry in the off-diagonal entries of the correlation matrix $C_{ij}(t)$ when one end is affected by the boundaries, since additional states contribute. Therefore we try to estimate the effect of the boundaries on the extraction of quantities with the GEVP from the asymmetry

$$\Delta_{ij}(t_s, t) \equiv C_{ij}(t_s, t) - C_{ji}(t_s, t). \quad (16.4)$$

The result does not change for other smearing levels.

In figure 16.6, we show Δ_{ij} for the correlation function C_{AA}^{stat} built from the local operator and the Gaussian smeared operator with the largest smearing radius. The vertical line denotes the time slice where $\Delta_{ij}(t)$ is zero within errors.

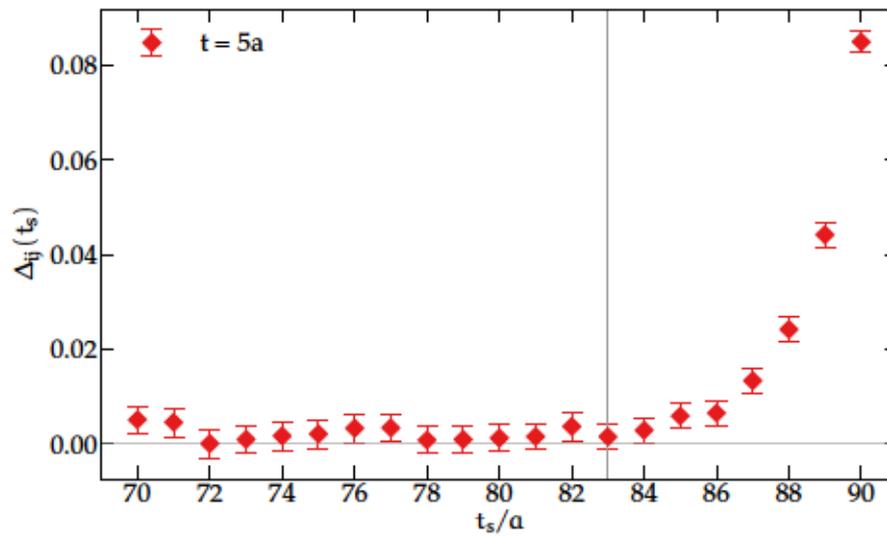


Figure 16.6: Asymmetry Δ_{ij} for C_{AA}^{stat} and $i = \theta$ and $j = \text{Gauss4}$ on H101. The vertical line denotes the source position where Δ_{ij} is zero within errors and the source-sink separation is $t = 5a$.

Close to the boundary, we see significant deviations between the corresponding correlation functions which quickly vanish when the distance to the boundary is increased. As far as we can judge with the statistical errors present in the data, the range of the effect of the boundaries on the asymmetry in the correlation matrix might be smaller than the region determined in the last section.

In conclusion, this criterion does not seem to be as conservative as the fit to the boundary states, performed above. Nevertheless, this is an encouraging result since the effect of the boundary on the extraction of physical quantities seems to vanish fast, when the distance to the boundaries is enlarged.

COMPUTATION OF THE $B^*B\pi$ COUPLING

All physical quantities that have been defined in the effective theory need to be renormalized via the matching of HQET to QCD before the continuum limit can be taken. Therefore, we are not able to extract physical results from the large-volume computations before the matching is finalized.

However, the measurement program `openHQET` that has been developed to calculate two- and three-point correlation functions can be employed to extract the correlation functions that are necessary to determine the static $B^*B\pi$ coupling which is of phenomenological importance for the chiral extrapolation of heavy-light observables.

In this chapter, we will briefly introduce the $B^*B\pi$ coupling, its use and its definition in our lattice setup. Afterwards we will present results that have been obtained on a subset of CLS ensembles and perform an extrapolation to the chiral limit.

As outlined in the introduction of this thesis, lattice calculations of B meson form factors are valuable inputs to constrain CKM matrix elements which are used to test the flavor sector of QCD. A major systematic uncertainty of these determinations comes from the difficulties to simulate at physical light quark masses and therefore, most results have to be extrapolated to the physical point. Theoretical insight on the functional form of this extrapolation helps to reduce systematic uncertainties.

The $B^*B\pi$ coupling serves as external input for such extrapolation guided by Heavy Meson Chiral Perturbation Theory ($\text{HM}\chi\text{PT}$). The coupling \hat{g} which is closely related to the $B^*B\pi$ coupling is a leading order coefficient in the expansion of observables in $\text{HM}\chi\text{PT}$. Its knowledge with high precision can improve chiral extrapolations of physical quantities as the ones we will determine in our project on HQET.

The determination of \hat{g} can be performed experimentally from $D^* \rightarrow D\pi$ decays [277]. The theoretically more appealing decay $B^* \rightarrow B\pi$ is not accessible experimentally, since it is kinematically forbidden. Lattice calculations have been used to determine the coupling with relativistic charm and bottom quarks [278, 279] and static heavy quarks [275, 280–283]. The relative precision of the latter is limited by the statistical uncertainties of static observables and the extrapolation of the results to the chiral limit of massless pions leads to systematic uncertainties for all lattice determinations.

17.1 HEAVY MESON CHIRAL PERTURBATION THEORY

HM χ PT [284–287] is constructed as effective theory for quantities including light and heavy quarks from a combination of HQET and Chiral Perturbation Theory (χ PT). Whereas HQET describes QCD observables in the limit of one infinitely heavy quark, χ PT describes the symmetries of QCD in the limit of massless quarks. Chiral symmetry is only softly broken in nature by the small masses of up, down and strange quark and has a significant impact on phenomenological results.

We will give a short introduction to HM χ PT and the coupling \hat{g} based on the reviews given in [280, 288]. Since pseudoscalar and vector mesons are mass-degenerate in the static limit of HQET, cf. eq. (7.28), we can express them in terms of the effective meson fields

$$H = \frac{1+\not{v}}{2} [B_\mu^* \gamma^\mu - B \gamma_5], \quad \bar{H} = \gamma_0 H^\dagger \gamma_0 \quad (17.1)$$

with the γ -matrices defined in appendix A, the velocity of the heavy meson v and the annihilation operators B and B^* for particles containing a bottom quark in the initial state.

For an introduction
to χ PT see, e.g.,
[289].

The Lagrangian of SU(3) χ PT contains an octet of pseudogoldstone bosons. In matrix notation, this octet can be expressed in the exponentiated form

$$\xi = \exp(i\mathcal{M}/f) \quad (17.2)$$

At tree level, $f = f_\pi$.

where f is the meson decay constant and the matrix \mathcal{M} is given by

$$\mathcal{M} = \begin{pmatrix} \frac{1}{\sqrt{2}}\pi^0 + \frac{1}{\sqrt{6}}\eta & \pi^+ & K^+ \\ \pi^- & -\frac{1}{\sqrt{2}}\pi^0 + \frac{1}{\sqrt{6}}\eta & K^0 \\ K^- & \bar{K}^0 & -\sqrt{\frac{2}{3}}\eta \end{pmatrix}. \quad (17.3)$$

Since the strange quark is about 30 times heavier than the averaged light quark mass, we will treat it as non-chiral and consider SU(2) χ PT where the up and down quarks are the dynamical degrees of freedom. The matrix \mathcal{M} is modified accordingly.

The strong interactions of B and B^* mesons with pions at low momentum are described by the effective Lagrangian of HM χ PT which is constructed by a combination of HQET and χ PT. It is built from a joint expansion in powers of the light quark mass and the inverse of the heavy quark mass. The interaction term to lowest order is given by [286, 287]

$$\mathcal{L}_{\text{HM}\chi\text{PT}}^{\text{int}} = \hat{g} \text{Tr} [\bar{H}_a H_b \mathcal{A}_\mu^{b_a} \gamma^\mu \gamma^5] \quad (17.4)$$

where the trace is taken in the space of 4×4 Dirac matrices and the indices a, b denote the SU(3) flavor components. The light degrees of freedom are contained in the factor

$$\mathcal{A}_\mu = \frac{i}{2} (\xi^\dagger \partial^\mu \xi - \xi \partial^\mu \xi^\dagger), \quad (17.5)$$

which can be expanded in the pion fields as

$$\mathcal{A}_\mu = -\frac{1}{f} \partial_\mu \mathcal{M} + \dots \quad (17.6)$$

As apparent from eq. (17.4), the coupling \hat{g} is the only leading order constant. This constant can be related to the $B^*B\pi$ coupling which is defined via [280]

$$\langle B^0(p)\pi^+(q)|B^{*+}(p') \rangle = -g_{B^*B\pi}(q^2)q_\mu\eta^\mu(p')(2\pi)^4\delta(p'-p-q) \quad (17.7)$$

where η^μ is the polarization vector of the B^* and the states are relativistically normalized via

$$\langle B(p)|B(p') \rangle = 2p^0(2\pi)^3\delta^{(3)}(p-p'). \quad (17.8)$$

The matrix element of eq. (17.7) at tree-level of HM χ PT is given by

$$\langle B^0(p)\pi^+(q)|B^{*+}(p') \rangle = -\frac{2m_B}{f_\pi}\hat{g}q_\mu\eta^\mu(p')(2\pi)^4\delta(p'-p-q) \quad (17.9)$$

and from the comparison of eqs. (17.7) and (17.9) we derive

$$g_{B^*B\pi} = \frac{2m_B}{f_\pi}\hat{g} \quad (17.10)$$

to tree-level. The physical coupling is defined for on-shell pions,

$$g_{B^*B\pi} \equiv \lim_{q^2 \rightarrow m_\pi^2} g_{B^*B\pi}(q^2). \quad (17.11)$$

Beyond leading order of HQET, the factor m_B is replaced by $\sqrt{m_B m_{B^}}$.*

17.2 LATTICE COMPUTATION

The determination of \hat{g} via a lattice computation is achieved by relating the form factor of the axial current between B and B^* states with the $B^*B\pi$ coupling via Lehmann–Symanzik–Zimmermann reduction of the pion [280]. The matrix element of the axial current is split into three form factors

$$\begin{aligned} \langle B^0(p)|A_\mu(0)|B^{*+}(p+q) \rangle &= \eta_\mu F_1(q^2) + (\eta \cdot q)(2p+q)_\mu F_2(q^2) \\ &\quad + (\eta \cdot q)q_\mu F_3(q^2) \end{aligned} \quad (17.12)$$

and with the PCAC relation

$$\pi(x) = \frac{1}{m_\pi^2 f_\pi} \partial^\mu A_\mu(x) \quad (17.13)$$

the $B^*B\pi$ coupling can be expressed via

$$\begin{aligned} g_{B^*B\pi}(q^2) &= -\frac{1}{f_\pi} \frac{m_\pi^2 - q^2}{m_\pi^2} [F_1(q^2) + (m_{B^*}^2 - m_B^2)F_2(q^2) \\ &\quad + q^2 F_3(q^2)]. \end{aligned} \quad (17.14)$$

If the computation is performed in the static limit, the second term vanishes and the analytical continuation to $q^2 \rightarrow 0$ leads to [280]

$$g_{B^*B\pi}(0) = -\frac{1}{f_\pi} F_1(0). \quad (17.15)$$

In the static limit for the heavy quark and the chiral limit for the light quarks, the determination of the coupling is therefore possible from the zero-momentum form factor, when

$$g_{B^*B\pi}(m_\pi^2) \approx g_{B^*B\pi}(0) \quad (17.16)$$

is assumed.

The determination of form factors from lattice computations involves the computation of two- and three-point correlation functions and the schematic form of the functions used in our work is given in fig. 16.4. Since another insertion is made, compared to two-point functions, two time scales are present in three-point functions. Excited state contributions are minimized if both time scales are maximized. Since the signal-to-noise ratio of heavy-light correlation functions decays exponentially, we cannot extract a signal at large time separations.

In [290], different approaches have been developed to minimize the systematic and statistic uncertainties in the determination of matrix elements and we will present two of them for the problem at hand. We aim to compute the matrix element [275]

$$\hat{g} = \frac{1}{2} \langle B^0(0) | (\hat{A}_k)_R(0) | B_k^{*+}(0) \rangle, \quad \hat{A} = \bar{\psi}_d(x) \gamma_\mu \gamma_5 \psi_u(x) \quad (17.17)$$

where we do not sum over the index $k \in \{1, 2, 3\}$ and ψ_u and ψ_d annihilate up and down quark. The normalization is such that

$$\langle B^0(p) | B^0(p) \rangle = \langle B^*(p) | B^*(p) \rangle = 2L^3 \quad (17.18)$$

We define the correlation matrix of summed three-point functions

$$D_{ij}^{3pt}(t) = a^3 \sum_y \langle (B_i^*)^\dagger(t) (A_k)_R(y) B_j(0) \rangle \quad (17.19)$$

with $t = t_B - t_A$ according to fig. 16.4. $B_i^{*\dagger}$ and B_j are interpolating fields for the corresponding mesons and the subsets ij denote different definitions of interpolating operators. The generic form of these interpolators, neglecting any smearing that is applied to them, reads

$$B(x_0) = a^3 \sum_x \bar{\psi}_b(x) \gamma_5 \psi_d(x) \quad (17.20)$$

$$B_k^*(x_0) = a^3 \sum_x \bar{\psi}_b(x) \gamma_k \psi_u(x) \quad (17.21)$$

$$A_k(x) = \bar{\psi}_d(x) \gamma_k \gamma_5 \psi_u(x). \quad (17.22)$$

In eq. (17.19), we sum over all time slices y_0 . The three-point function is independent of the index k . From now on, we will work in the

The renormalization of the axial current is specified below.

We sum over all spatial lattice points to project onto zero spatial momentum.

static limit, i.e., B and B^* are mass-degenerate. The renormalized axial current according to eq. (4.7) is given by

$$(A_k)_R = Z_A(g_0^2)(1 + b_A(g_0^2)am_q)(A_k + ac_A(g_0^2)\partial_k P). \quad (17.23)$$

At zero momentum transfer, the term proportional to c_A does not contribute [275, 291].

In combination with the two-point function

$$C_{ij}^{2pt}(t) = \langle B_i^\dagger(t)B_j(0) \rangle \quad (17.24)$$

we can define the effective matrix element via

$$M^{\text{ratio}}(t) = -\partial_t a \frac{D_{ij}^{3pt}(t)}{\sqrt{C_{ii}^{2pt}(t)C_{jj}^{2pt}(t)}} = \hat{g} + O(t\Delta e^{-t\Delta}) \quad (17.25)$$

$$\text{with } \Delta = E_2 - E_1,$$

where $\partial_t f(t) \equiv \frac{1}{a}[f(t+a) - f(t)]$. The systematic corrections to this quantity scale with the energy gap between ground state and excited state of the B mesons.

These effects can be further suppressed if the GEVP is used. We determine eigenvalues and eigenvectors from the solution of

$$C^{2pt}(t)v_n(t, t_0) = \lambda_n(t, t_0)C^{2pt}(t_0)v_n(t, t_0) \quad (17.26)$$

where we use the correlation matrix C_{ij}^{2pt} of size $N \times N$ and define the effective matrix element

$$M_n^{\text{GEVP}}(t, t_0) = -\frac{1}{2}\partial_t \frac{(v_n(t, t_0), [D^{3pt}(t)\lambda_n^{-1}(t, t_0) - D^{3pt}(t_0)] v_n(t, t_0))}{(v_n(t, t_0), C^{2pt}(t_0)v_n(t, t_0))} \\ = \hat{g}_{n,n} + O(e^{-\Delta_{N,n}t}), \quad (17.27)$$

$$\Delta_{N,n} = E_{N+1} - E_n. \quad (17.28)$$

In this case, the leading systematic correction is given by the energy gap between the ground state and the first state that is not included in the GEVP, provided that $2t_0 \geq t$ holds. We are therefore able to effectively suppress excited state contributions and expect plateaus already at comparably small source-sink separations. In the following, we will consider the most conservative choice for t_0 which most effectively suppresses systematic effects and define

$$M_n^{\text{GEVP}}(t) \equiv M_n^{\text{GEVP}}(t, t-a). \quad (17.29)$$

In eq. (17.27) we generalized the definition of the effective matrix element to matrix elements between degenerate excited states. In this notation we set $\hat{g} \equiv \hat{g}_{11}$ and the definition of the matrix element in eq. (17.17) changes accordingly for $n > 2$.

In [275], the argument of C^{2pt} in the denominator is t . We corrected this typo to t_0 in this formula.

Table 17.1: Overview of the HQET measurements performed on the large-volume CLS ensembles. m_π is taken from [270]. τ_{exp} has been determined according to eq. (13.1) and is given in units of the separation between two configurations. N_t gives the number of time slices included in the analysis and N_r the number of random sources per time slice.

id	m_π	N_{cfg}	τ_{exp}	N_t	N_r
H101	416	2000	10.0	47	2
H102	354	1997	10.0	47	2
N101	282	1457	10.1	80	2
C101	221	2000	10.2	47	2
S400	351	800	12.9	64	2
H200	419	2000	18.1	47	1
N203	345	755	18.0	61	1
N200	282	1306	18.1	61	1

17.3 SETUP

We use the setup which is described in section 16.1. To stabilize the GEVP and to decrease the statistical error, we sum over N_r random sources per time slice and N_t time slices per configuration. The range of time slices that is used for this summation is dictated by the findings of section 16.3. In contrast to what was done in [275], we do not sum over all polarizations k but randomly choose a value of $k \in \{1, 2, 3\}$ for each random source. Therefore, we perform $N_t \times N_r \times 2$ inversions of the Dirac equation per configuration. We give an overview of the status of the measurements in table 17.1.

The GEVP is based on a 3×3 correlation matrix. An increase of the number of interpolators in the matrix leads to numerical instabilities in the inversion of $C(t_0)$ since we have to consider rather large values of t_0 due to our choice $t_0 = t - a$. An analysis of the variational basis for the matrix elements in the spirit of section 16.2 has revealed that the suppression of systematic effects is similar for a large number of combinations. We will consider a combination of the local interpolator, a Gaussian smeared interpolator with intermediate smearing radius and a Gaussian smeared interpolator with the maximal smearing radius in the following.

To renormalize the matrix elements, we use the values for $Z_{A,\text{sub}}^1$ from the L_1 -LCP given in table 6 of [66]. The one-loop perturbative formula for b_A is given by [234]

$$b_A = 1 + 0.0881 C_F g_0^2, \quad C_F = \frac{4}{3}. \quad (17.30)$$

The estimates for κ_{cr} at the coupling values of the CLS simulations for the computation of the subtracted quark mass are taken from table III of [68].

To determine the value of the bare matrix element from $M_n^{\text{GEVP}}(t)$, we perform a plateau fit at large values of t . We determine the beginning of this plateau denoted by t_{min} by the following procedure. We start from $t \approx 0.5 \text{ fm}$ and increase t until the condition

$$|M_n^{\text{GEVP}}(t) - M_n^{\text{GEVP}}(t - \delta t)| < \sigma(t) \quad \text{with} \quad \delta t = \frac{1}{\Delta_{N,n}} \quad (17.31)$$

is fulfilled. $\sigma(t)$ is the statistical error on $M_n^{\text{eff}}(t)$. For the matrix element from the GEVP we use $\Delta_{4,n}$, whereas for the matrix element from the summed ratio in eq. (17.25) we consider $\Delta_{2,1}$. From this condition it is ensured that the statistical errors exceed the systematic ones by a factor of $e - 1$ at $t = t_{\text{min}}$. The upper end of the plateau range denoted by t_{max} is given by the first time slice where the relative error of the matrix element exceeds 30%. Since we consider plateaus of length $O(5a)$, we are able to perform a correlated plateau fit. The differences to uncorrelated fits are small, compared to the statistical errors.

At the level of statistics at our disposal, the determination of $\Delta_{4,1}$ is rather an estimation than a reliable calculation, since the static energy for levels $n > 3$ has significant statistical uncertainties. We obtain values of $O(1.2 \text{ GeV})$ which is in line with the findings in the two-flavor theory [275].

We use a 6×6 correlation matrix to compute E_n^{stat} for $n = 1, \dots, 4$.

17.4 RESULTS

With the current status of the measurements as detailed in table 17.1, we have gathered statistic that is almost compatible to the computation in the two-flavor theory [275], albeit at slightly larger lattice spacings. We have performed measurements at three lattice spacings and light quark masses down to 221 MeV. The extension to a fourth lattice spacing and physical quark masses is planned, to further reduce systematic uncertainties.

17.4.1 Plateaus

In figure 17.1 we show the results for \hat{g}_{11} together with the corresponding plateau fits for exemplary ensembles. To illustrate the qualitative

differences between the determinations based on the summed ratio and the GEVP, we show the results for both methods. In the case of the summed ratio, we have used the local and the maximally smeared interpolator.

It is apparent that t_{\min} has to be set to significantly larger values in the case of the summed ratio due to the smaller gap in the energy spectrum. At fixed source-sink separation, the determination from the GEVP shows larger statistical errors and at large source-sink separations, the inversion of $C(t_0)$ is likely to fail. However, both plateau averages are compatible and in most cases the statistical error on the plateau average from the GEVP is significantly smaller than the error from the summed ratio. This nicely illustrates the superiority of the GEVP. We will only consider the values obtained from the GEVP for the chiral-continuum extrapolation.

Table 17.2: Renormalized couplings \hat{g}_{11} and \hat{g}_{22} evaluated with HYP1 and HYP2 static quark actions.

id	$\hat{g}_{11}^{\text{HYP1}}$	$\hat{g}_{11}^{\text{HYP2}}$	$\hat{g}_{22}^{\text{HYP1}}$	$\hat{g}_{22}^{\text{HYP2}}$
H101	0.534(2)	0.539(2)	0.444(9)	0.450(9)
H102	0.520(6)	0.524(5)	0.433(28)	0.453(12)
N101	0.512(4)	0.517(3)	0.423(10)	0.426(10)
C101	0.503(4)	0.507(4)	0.406(14)	0.415(14)
S400	0.519(5)	0.522(5)	0.420(29)	0.427(27)
H200	0.532(6)	0.534(9)	0.448(31)	0.460(30)
N203	0.530(8)	0.530(7)	0.416(49)	0.413(47)
N200	0.523(8)	0.521(8)	0.431(33)	0.429(31)

The results from the summed ratio are tabulated in tab. B.13.

An overview of the results obtained from both variants of the HYP smearing is provided in table 17.2. The small difference between both sets of smearing parameters decreases towards the continuum limit. In general, observables defined from HYP2 have slightly smaller statistical errors. Therefore, we will consider the HYP2 smeared observables in the chiral-continuum extrapolation.

17.4.2 Chiral-continuum extrapolation

The approach of \hat{g} to the chiral limit including the next-to-leading order of $\text{HM}\chi\text{PT}$ has been derived in [283, 292]. Logarithmic modifications of the leading behavior due to the next-to-leading order terms have been found. In the two-flavor study [275] a simple linear approach has been investigated in addition to the analytical one, due

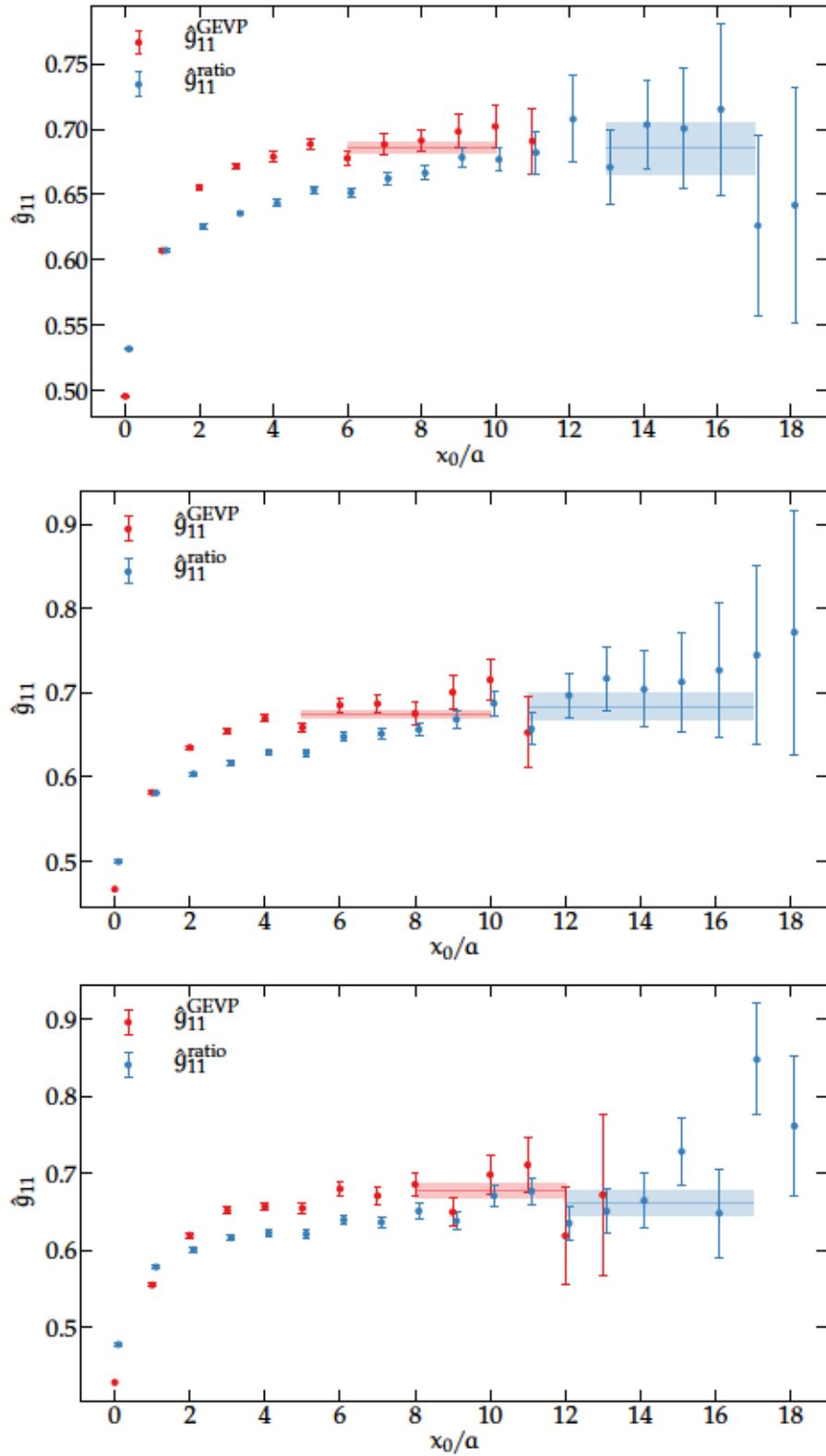


Figure 17.1: Plateaus of the bare matrix element \hat{g}_{11} computed from the GEVP and the summed ratio for the ensembles (top to bottom) N101, C101 and N200. The data from the ratio method has been shifted horizontally to improve visibility.

to the linearity in the data in the considered region of pion masses. The two forms read

$$\hat{g}_{11}^{\text{lin}}(y, a) = \hat{g}_x + By + Ca^2, \quad (17.32)$$

$$\hat{g}_{11}^{\text{nlo}}(y, a) = \hat{g}_x \left[1 - (1 + 2(\hat{g}_x)^2 y \log y) \right] + By + Ca^2, \quad (17.33)$$

where the pion mass dependence has been parametrized through the variable

$$y = \frac{m_\pi^2}{8\pi^2 f_\pi^2}. \quad (17.34)$$

For am_π and af_π we use the values given in [270]. In both ansätze \hat{g}_x , B and C are free parameters. The term Ca^2 is included to model the leading cut-off effects.

We show our results for \hat{g} depending on y based on three different lattice spacings in fig. 17.2 together with fits to eqs. (17.32–17.33). Irrespective of the fit ansatz, the constant C is zero within errors. We therefore fix $C = 0$ in our fits. Terms of $O(y^2)$ are not resolvable, as well.

A good quality of the fit is found for both ansätze and within the currently available range of pion masses, we are not able to rule out one of the two fit forms. Since pion masses of $O(400 \text{ MeV})$ are far from being chiral, we perform a cut at $y = 0.1$ such that we take only ensembles with pion masses $< 355 \text{ MeV}$ into account. The linear fit is compatible with both data points outside the fit range. For $\hat{g}_{11}^{\text{NLO}}$, the fit and the data point for H101 are compatible within 2σ .

In the chiral limit we obtain

$$\hat{g}_x^{\text{lin}} = 0.493(7) \quad \hat{g}_x^{\text{nlo}} = 0.456(6) \quad (17.35)$$

and to take into account the systematic uncertainty due to the two different fit forms, we combine both values and their errors to the result

$$\hat{g}_x = 0.475(24). \quad (17.36)$$

The uncertainty of this result is dominated by the systematic uncertainty and changes in the fits, such as allowing for a non-zero value of C or taking into account all pion masses, do not lead to significant variations of this result.

17.5 DISCUSSION

Our current result has the same level of accuracy as the result $\hat{g} = 0.492(29)$ of the two-flavor study [275], which is the most precise determination of \hat{g} available by now. Both results are compatible with each other and in both cases, no dependence on the lattice spacing can be resolved.

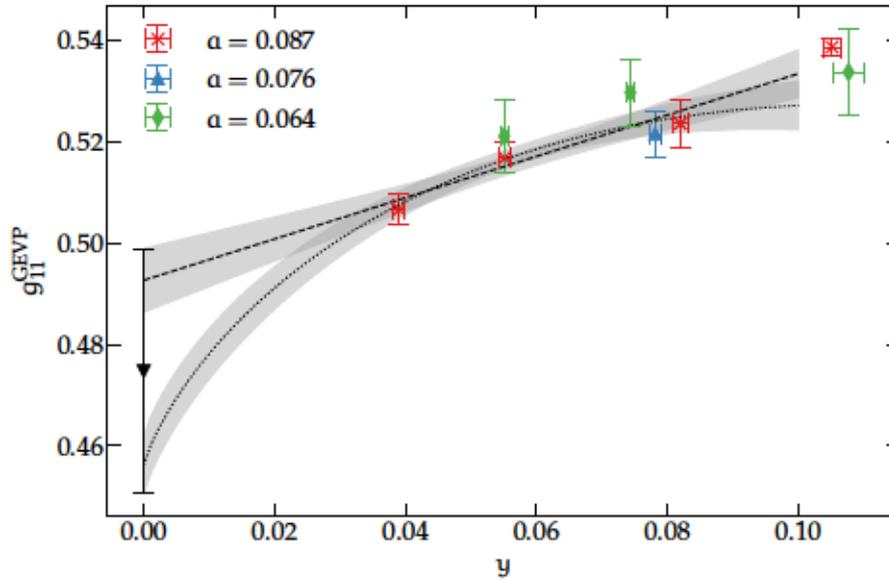


Figure 17.2: Overview of the results for \hat{g}_{11} at three lattice spacings together with the chiral extrapolations according to eqs. (17.32) and (17.33). The error of the result at the chiral point covers both extrapolation formulas.

We plan to improve our result in the near future by taking into account ensembles with considerably smaller pions masses and a finer lattice spacing. The ensembles D200, E300 and E250 will allow us to constrain the fits in a region down to $y = 0.0135$ and help us to decide between the two fit forms in eqs. (17.32–17.33) or to tightly constrain the chiral log. This will result in a drastic reduction of the uncertainties.

As soon as we reach a relative precision of $O(2\%)$, we have to take another source for systematic uncertainties into account that has its origin in our approach to the chiral limit. On the chiral trajectory where $\text{Tr}[M_q]$ is held constant, the strange quark mass approaches its physical value from below when the light quark masses are lowered towards their physical values. In the chiral extrapolation of \hat{g} we do not extrapolate to the physical point but to the chiral point, i.e., the point where the light quark masses vanish.

At this point of our chiral trajectory, the strange quark mass is larger than physical. The chiral $SU(2)$ symmetry is preserved and the definition of \hat{g} in the limit of $SU(2)$ HM χ PT is valid. Nevertheless, if this low-energy constant is used for extrapolations to the physical point, we have to think about the impact of the unphysically large strange quark mass on \hat{g} . From the value of ϕ_4^{phys} in eq. (13.5) we obtain

$$m_\pi^\chi \equiv 0, \quad m_K^\chi \approx 503 \text{ MeV} \quad (17.37)$$

An overview of the ensembles is shown in fig. 13.1.

at the chiral point whereas the isospin and QED corrected corrected physical value of the Kaon is $m_K^{\text{phys}} = 494.2(3)$ MeV.

In [279, 283] a similar deviation of the strange quark mass from its physical value due to a miss-tuning has been translated to an additional 1.5% systematic uncertainty on the final value of \hat{g} based on the size of the effect of quark mass reweighting on axial couplings determined on the same set of ensembles. A similar error has to be added to our result, when the systematic uncertainty due to the fit forms has been reduced. Alternatively, the inclusion of ensembles on the $m_s = \text{phys}$ trajectory of the CLS ensembles could help to map out the dependence of \hat{g} on m_s .

To conclude our discussion, we also quote the value of \hat{g} at physical quark masses. From our fits we obtain

$$\hat{g}^{\text{lin}}(m_\pi^{\text{phys}}) = 0.498(5) \quad \hat{g}^{\text{nlo}}(m_\pi^{\text{phys}}) = 0.483(5) \quad (17.38)$$

and to take into account the systematic uncertainty due to the two different fit forms, we combine both values and their errors to the result

$$\hat{g}(m_\pi^{\text{phys}}) = 0.491(12). \quad (17.39)$$

which is evaluated at the physical strange quark mass.

17.6 THE MATRIX ELEMENT OF THE FIRST RADIAL EXCITATIONS

The GEVP allows us to determine the matrix element of the first radial excitations

$$\hat{g}_{22} = \frac{1}{2} \langle B^0(0), 2 | (\hat{A}_k)_R(0) | B_k^{*+}(0), 2 \rangle. \quad (17.40)$$

We show exemplary plateaus for both variants of the HYP smearing in fig. 17.3. The systematic effects due to excited state contributions appear to be small in comparison to the statistical uncertainties. Although the latter are rather large, we are able to determine a plateau region according to the criterion in eq. (17.29) using an estimate for $\Delta_{4,2}$. We tabulate the results for all ensembles and both variants of the HYP smearing in tab. 17.2. Although the statistical errors are a little bit smaller for HYP2 smearing, significant differences between HYP1 and HYP2 smearing are not visible.

Since the functional form of \hat{g}_{22} towards the chiral limit including chiral logs is not known, we use eq. (17.32) to extrapolate the matrix element to the chiral-continuum limit. As for \hat{g}_{11} we cannot resolve a dependence on the lattice spacing and set $C = 0$. From the chiral fit, we obtain

$$\hat{g}_{22} = 0.398(15). \quad (17.41)$$

We use the convention where $f_\pi \approx 130$ MeV in the chiral limit.

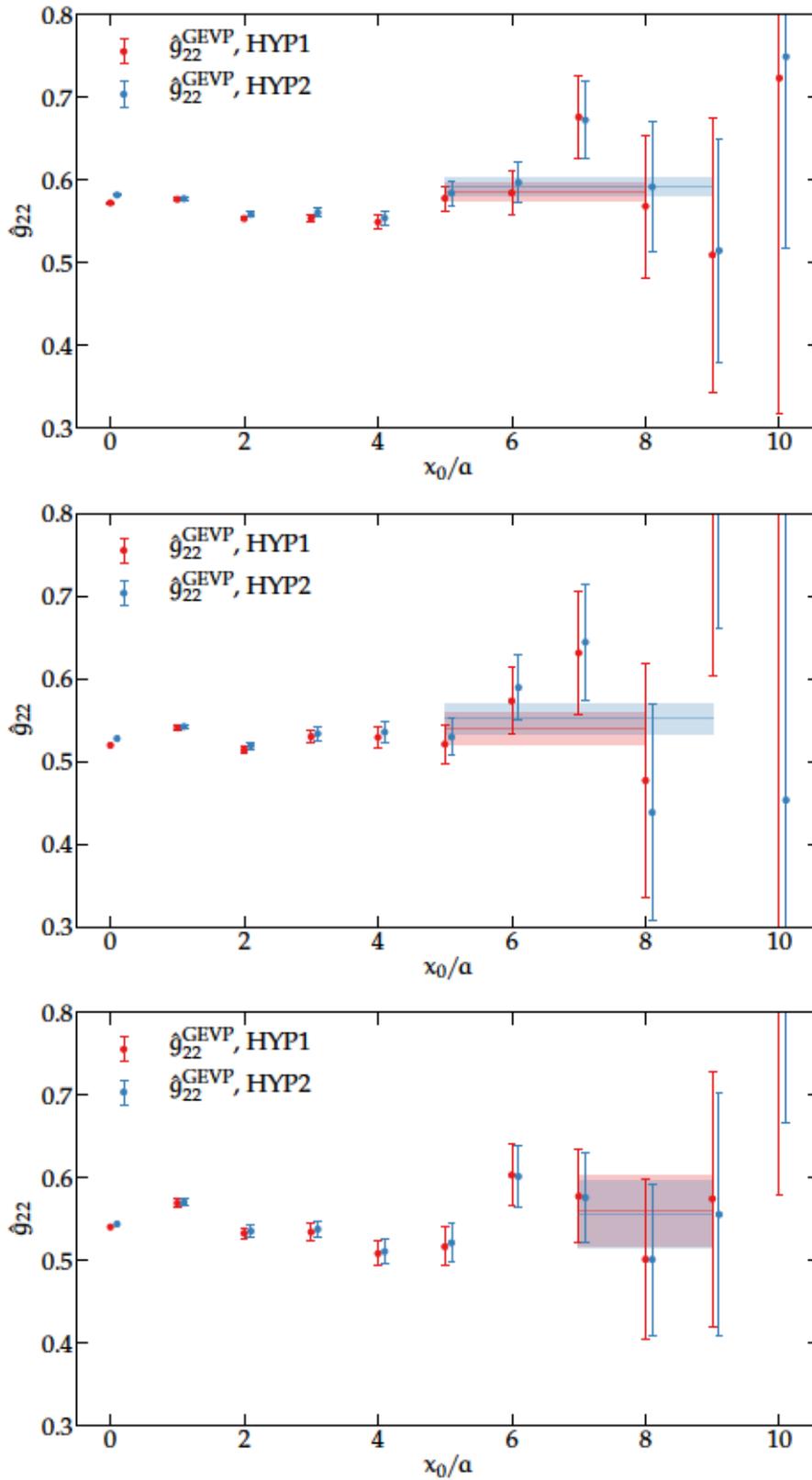


Figure 17.3: Plateaus of the bare matrix element \hat{g}_{22} based on HYP1 and HYP2 smearing for the ensembles (top to bottom) H101, C101 and N200. The data from HYP2 smearing has been shifted horizontally to improve visibility.

This value is compatible with the result $\hat{g}_{22} = 0.425(70)$ of [275], where the spread of the two fits with and without an a^2 term has been used to estimate the uncertainty of the continuum extrapolation, and the result $\hat{g}_{22} = 0.38(4)$ of [293]. We are confident that we are able to constrain our result more tightly and reduce systematic effects, that we have not yet quantified, by the addition of further ensembles.

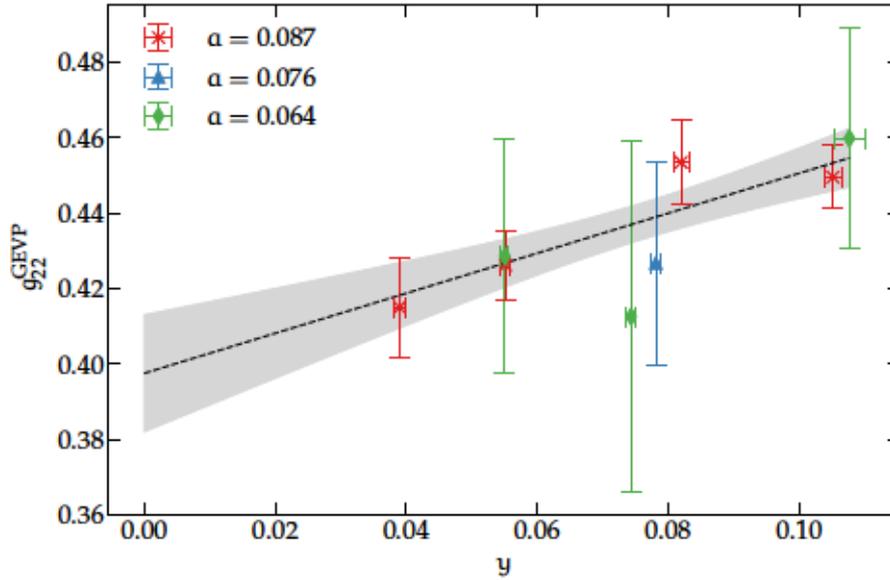


Figure 17.4: Overview of the results for \hat{g}_{22} at three lattice spacings together with the chiral extrapolation according to eq. (17.32).

Part VI

CONCLUSIONS

CONCLUSIONS

In this work, we have describe our efforts to extract Standard Model parameters in the heavy quark sector of three-flavor lattice QCD. We have set up the line of constant physics along which we will perform the non-perturbative matching of QCD and HQET. The resulting matching parameters will be combined with measurements in the effective theory on large-volume ensembles to compute observables including bottom quarks. We have described our measurements in the $2+1$ flavor theory on these ensembles and calculated the coupling \hat{g} of $\text{HM}\chi\text{PT}$. Furthermore we have performed the determination of the charm quark mass on the same set of ensembles based on $\mathcal{O}(a)$ improved current quark masses.

The ensemble generation for the non-perturbative matching is based on an extensive tuning procedure of the physical and algorithmic parameters. We have described our tuning towards constant renormalized couplings and massless sea quarks based on preexisting information and our own experiences. The maximal error of the gradient flow couplings in the tuning volume translates to a relative error of 0.7% in the physical lattice extent which is about a factor of three more precise than in the corresponding project in the two-flavor theory. This improvement in the precision will propagate into physical observables via the matching parameters.

We have performed simulations at five different lattice spacings at all stages of the matching procedure. This will allow us to perform the continuum extrapolations of the matching parameters to high precision and with small systematic uncertainties. The simulation of the most demanding ensembles with 48 and 64 lattice sites in all four dimensions required an extensive tuning of the algorithmic parameters of the `openQCD` package. Since the deflated and SAP preconditioned solvers had not been used in prior simulations in the Schrödinger functional, we have carefully investigated the parameter space based on experiences from the generation of the large-volume CLS simulations. The documentation of an exemplary parameter scan may allow to speed up the tuning for future simulations and measurements in the Schrödinger functional. With the ensembles that have been generated in our work, we are in the position to perform the measurements for the calculation of the matching parameters at all stages of our strategy.

The second prerequisite for the determination of the matching parameters is the extraction of valence quark hopping parameters for fixed renormalized quark masses on the matching ensembles. As outlined in our introduction to HQET on the lattice, Symanzik im-

provement of $O(a)$ is necessary to disentangle discretization effects from terms of next-to-leading order in the HQET expansion. We have computed the coefficients that are necessary to improve the renormalized quark mass from heavy-light and heavy-heavy correlation functions in a broad range of valence quark masses. Scaling tests confirm our assumption that the computation of these coefficients at heavy valence quark masses improves the scaling of heavy quark observables in the matching volume. This improvement is achieved by the reduction of mass-dependent cut-off effects of higher orders. To eliminate the systematic uncertainties inherent in the interpolating formulas of the renormalization constant Z_A determined on different lines of constant physics, we have performed the determination of this constant on the matching LCP. Together with our determination of Z_P we have gathered all ingredients for the computation of renormalized quark masses.

By combining our results with the running factor from the matching scale to RGI values, we have performed the determination of hopping parameters at fixed RGI quark masses and demonstrated their use in scaling tests. Based on this calculation we are now able to fix the hopping parameters for the determination of the matching coefficients in a range of fixed heavy quark masses.

As soon as the matching coefficients are determined, we will use them to renormalize large-volume observables in the effective theory. The computation of the corresponding correlation functions has been started using a newly developed code and we have given an overview over the current status of the measurements and the extraction of effective energies using the GEVP. We have tested different basis interpolators at coarse lattice spacing to optimize the measurements on the computationally demanding ensembles.

The correlation functions that are currently being measured can be employed to determine the mass of the bottom quark and the decay constants of B and B_s mesons. Based on our optimization of the variational basis, we are able to extract the corresponding energies and matrix elements in the effective theory. For the computation of form factors of semi-leptonic B decays, we will perform measurements of the appropriate three-point functions with momentum insertions to map out the momentum dependence of the form factors.

Whereas these computations will need the input from the matching, we have used the current status of the large-volume measurements to perform a chiral-continuum extrapolation of the coupling \hat{g} of $\text{HM}\chi\text{PT}$. We are able to reach the precision of the two-flavor result and at the same time eliminated the systematic error due to the use of two flavors. Since the impact of heavy quarks on low-energy quantities is strongly suppressed, the error due to missing charm and bottom quark loops is small. We have not yet been able to eliminate the systematic error stemming from different approaches to the chiral point. However, the

inclusion of measurements on ensembles with significantly smaller pion masses into our analysis will allow us to tightly constrain the extrapolation. At this point, the effect of an unphysically large strange quark mass at the $SU(2)$ chiral point may have to be taken into account.

Considering the currently existing determinations of \hat{g} , our final result will most likely be the most precise one. It will be used to guide chiral extrapolations of B meson quantities on the lattice and for phenomenological applications of $HM\chi PT$.

In addition to our investigations in the matching regime, we have applied our strategy to determine the improvement coefficients for renormalized quark masses in the coupling region of the large-volume CLS ensembles. The detailed investigation of ambiguities and systematic uncertainties and the comparison with already existing but less precise results obtained from other methods is reassuring. Our results have been employed for the determination of light and strange quark masses on the CLS ensembles, where their impact has been small compared to the size of the statistical error.

In our determination of the charm quark mass, the removal of these cut-off effects is vital to perform the continuum extrapolation with leading cut-off effects of $O(\alpha^2)$. We have detailed the extraction of meson and quark masses on the CLS ensembles with a focus on the boundary and cut-off effects in the presence of open boundary conditions and investigated the impact of distance preconditioning and noise sources on single time slices. Our results are ready to be employed in chiral-continuum extrapolations based on several different definitions of the charm quark mass and various improvement conditions. Whereas a final investigation of the systematic uncertainties in this extrapolation is in progress and shows little deviation between different setups, we have quoted a preliminary result that is compatible in its mean and relative error with the currently available three-flavor results.

CONVENTIONS

A.1 THE SPECIAL UNITARY GROUP

In this work, we consider the so-called special unitary groups of degree N , which we abbreviate $SU(N)$. These are defined as the Lie groups of $N \times N$ unitary matrices with determinant 1. The corresponding Lie algebra $\mathfrak{su}(N)$ is the space of complex $N \times N$ anti-hermitian and traceless matrices.

Following the notation of [31], we choose a basis T^a , $a = 1, 2, \dots, N^2 - 1$ in this space such that

$$\text{Tr}[T^a T^b] = -\frac{1}{2} \delta^{ab}. \quad (\text{A.1})$$

The structure constants f^{abc} are then defined by

$$[T^a, T^b] = if^{abc} T^c \quad (\text{A.2})$$

and they are real and totally anti-symmetric under permutation of the indices. For the group $SU(2)$, the generators are defined from the three Pauli matrices

$$\tau^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (\text{A.3})$$

by

$$T^a = \frac{\tau^a}{2}. \quad (\text{A.4})$$

The structure constants are given by the Levi-Civita symbol ϵ^{abc} . Furthermore, we define the ladder operators

$$\tau^\pm = \tau^1 \pm i\tau^2, \quad \tau^+ = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}, \quad \tau^- = \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}, \quad (\text{A.5})$$

For the gauge group $SU(3)$, the generators are defined from the Gell-Mann matrices λ^a via

$$T^a = \frac{\lambda_a}{2} \quad (\text{A.6})$$

and the explicit form of the Gell-Mann matrices reads

$$\begin{aligned}
 \lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda_2 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\
 \lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, & \lambda_5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, & \lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\
 \lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, & \lambda_8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. & & (A.7)
 \end{aligned}$$

A.2 DIRAC MATRICES

The Dirac matrices in Euclidean space and in a chiral representation are defined by

We choose the conventions of [31].

$$\gamma_\mu = \begin{pmatrix} 0 & e_\mu \\ e_\mu^\dagger & 0 \end{pmatrix} \quad (A.8)$$

from the 2×2 matrices

$$e_0 = -\mathbf{1}, \quad e_k = -i\tau_k \quad (A.9)$$

with the Pauli matrices defined in eq. (A.3). From this definition, the relation

$$\gamma_\mu^\dagger = \gamma_\mu, \quad \{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu} \quad (A.10)$$

can be derived and we define the matrix

$$\gamma_5 \equiv \gamma_0 \gamma_1 \gamma_2 \gamma_3 = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \quad (A.11)$$

which has the properties

$$\gamma_5 = \gamma_5^\dagger, \quad \gamma_5^2 = \mathbf{1}. \quad (A.12)$$

The Dirac matrices are used to define the hermitian matrices

$$\sigma_{\mu\nu} = \frac{i}{2} [\gamma_\mu, \gamma_\nu] \quad (A.13)$$

and their explicit form reads

$$\sigma_{0k} = \begin{pmatrix} \tau_k & 0 \\ 0 & -\tau_k \end{pmatrix}, \quad \sigma_{ij} = -\epsilon_{ijk} \begin{pmatrix} \tau_k & 0 \\ 0 & \tau_k \end{pmatrix}, \quad (A.14)$$

where ϵ_{ijk} is the totally anti-symmetric tensor with $\epsilon_{123} = 1$.

A.3 LATTICE DERIVATIVES

In this appendix, we define the discretized derivatives used in this work. We define them as finite-difference operators. The forward and backward lattice derivatives acting on function $\psi(\mathbf{n})$ are given by

$$\partial_{\mu}\psi(\mathbf{n}) = \frac{1}{a} [\psi(\mathbf{n} + a\hat{\mu}) - \psi(\mathbf{n})] , \quad (\text{A.15})$$

$$\partial_{\mu}^*\psi(\mathbf{n}) = \frac{1}{a} [\psi(\mathbf{n}) - \psi(\mathbf{n} - a\hat{\mu})] , \quad (\text{A.16})$$

and we can also formulate them via

$$\partial_{\mu} = \frac{1}{a} [\delta_{\mathbf{n}+a\hat{\mu},\mathbf{n}} - \delta_{\mathbf{n},\mathbf{n}}] , \quad (\text{A.17})$$

$$\partial_{\mu}^* = \frac{1}{a} [\delta_{\mathbf{n},\mathbf{n}} - \delta_{\mathbf{n}-a\hat{\mu},\mathbf{n}}] . \quad (\text{A.18})$$

Based on these definitions, the gauge covariant derivatives are given by

$$\nabla_{\mu}\psi(\mathbf{n}) = \frac{1}{a} [U_{\mu}(\mathbf{n})\psi(\mathbf{n} + a\hat{\mu}) - \psi(\mathbf{n})] , \quad (\text{A.19})$$

$$\nabla_{\mu}^*\psi(\mathbf{n}) = \frac{1}{a} [\psi(\mathbf{n}) - U_{\mu}(\mathbf{n} - a\hat{\mu})^{-1}\psi(\mathbf{n} - a\hat{\mu})] , \quad (\text{A.20})$$

$$\bar{\psi}(\mathbf{n}) \overleftarrow{\nabla}_{\mu} = \frac{1}{a} [\bar{\psi}(\mathbf{n} + a\hat{\mu})U_{\mu}(\mathbf{n})^{-1} - \bar{\psi}(\mathbf{n})] , \quad (\text{A.21})$$

$$\bar{\psi}(\mathbf{n}) \overleftarrow{\nabla}_{\mu} = \frac{1}{a} [\bar{\psi}(\mathbf{n}) - \psi(\mathbf{n} - a\hat{\mu})U_{\mu}(\mathbf{n} - a\hat{\mu})] . \quad (\text{A.22})$$

The symmetric derivative is defined by

$$\delta_{\mu}\psi(\mathbf{n}) \equiv \frac{1}{2}(\partial_{\mu} + \partial_{\mu}^*)\psi(\mathbf{n}) = \frac{1}{2a} [\psi(\mathbf{n} + a\hat{\mu}) - \psi(\mathbf{n} - a\hat{\mu})] , \quad (\text{A.23})$$

and the covariant counterpart is given by

$$\bar{\nabla}\psi(\mathbf{n}) = \frac{1}{2a} [U_{\mu}(\mathbf{n})\psi(\mathbf{n} + a\hat{\mu}) - U_{\mu}(\mathbf{n} - a\hat{\mu})^{-1}\psi(\mathbf{n} - a\hat{\mu})] . \quad (\text{A.24})$$

The second derivative is defined by

$$\partial_{\mu}^*\partial_{\mu}\psi(\mathbf{x}) = \frac{1}{a^2} [\psi(\mathbf{n} + a\hat{\mu}) - 2\psi(\mathbf{n}) + \psi(\mathbf{n} - a\hat{\mu})] \quad (\text{A.25})$$

and from this definition the Laplacian in k dimensions is defined by

$$\begin{aligned} \Delta\psi(\mathbf{x}) &= -\nabla_{\mu}\nabla_{\mu}^*\psi(\mathbf{n}) \\ &= \frac{1}{a^2} \sum_{\mu=1}^k [U_{\mu}(\mathbf{n})\psi(\mathbf{n} + a\hat{\mu}) - 2\psi(\mathbf{n}) + U_{\mu}(\mathbf{n} - a\hat{\mu})^{-1}\psi(\mathbf{n} - a\hat{\mu})] . \end{aligned} \quad (\text{A.26})$$

In the determination of PCAC masses, we also consider symmetric improved derivatives. These are constructed by an investigation of the Taylor expansion of the derivative in orders of the lattice spacing. Standard forward and backward derivatives are the continuum ones up to terms of $O(a)$ and the symmetric derivative approximates the continuum one up to terms of $O(a^2)$. The improved symmetric derivatives are exact up to terms of $O(a^4)$, when acting on smooth functions.

The improved symmetric derivative of first order is defined by

$$\begin{aligned}\delta_\mu\psi(n) &\rightarrow \delta_\mu\left(1 - \frac{1}{6}a^2\partial_\mu^*\partial_\mu\right)\psi(n) \\ &= \frac{1}{12a}\left(\psi(n-2a\hat{\mu}) - 8\psi(n-a\hat{\mu}) \right. \\ &\quad \left. + 8\psi(n+a\hat{\mu}) - \psi(n+2a\hat{\mu})\right),\end{aligned}\quad (\text{A.27})$$

and the improved symmetric derivative of second order is given by

$$\begin{aligned}\partial_\mu^*\partial_\mu\psi(x) &\rightarrow \partial_\mu^*\partial_\mu\left(1 - \frac{1}{12}a^2\partial_\mu^*\partial_\mu\right)\psi(x) \\ &= -\frac{1}{12a^2}\left(\psi(n-2a\hat{\mu}) - 16\psi(n-a\hat{\mu}) \right. \\ &\quad \left. + 30\psi(n) - 16\psi(n+a\hat{\mu}) + \psi(n+2a\hat{\mu})\right).\end{aligned}\quad (\text{A.28})$$

TABLES

B.1 THE LINE OF CONSTANT PHYSICS FOR A NON-PERTURBATIVE MATCHING OF HQET AND QCD

Table B.1: Tuning runs to fix $\bar{g}_{\text{GF}}^2(L_0) = g_*^2 = 3.949$ and $m_1 = 0$. Boundary conditions and action are chosen exactly as in [45], and we use $T = L$. N_{rep} is the number of replica runs, while N_{ms} gives the number of configurations, separated by τ_{ms} molecular dynamics units. The trajectory length is $\tau = 2$ throughout, except for the first run with $L_0/a = 12$. Values of the sea quark mass $L_0 m_1$ marked by a star have not been determined on the full statistics. The ensembles where the coupling is printed in bold have been used as basis for the matching in L_1 .

L_0/a	β	κ	N_{rep}	$\frac{\tau_{\text{ms}}}{\text{MD}}$	N_{ms}	\bar{g}_{GF}^2	$L_0 m_1$
12	4.297489	0.1360085	5	12.5	7500	3.9596(43)	-0.0012(7)*
	4.296368	0.136011	5	10	8000	3.9674(45)	-0.0004(7)*
	4.3020	0.1359977	7	8	5800	3.9533(59)	-0.00038(45)
	4.3030	0.1359947	7	8	9669	3.9461(41)	-0.00032(36)
16	4.474541	0.1355870	7	10	2457	3.9136(97)	-0.0009(4)*
	4.4624	0.1356163	7	10	2947	3.9469(83)	
	4.46466	0.1356109	7	10	2726	3.9676(93)	
	4.4662	0.1355985	7	10	5887	3.9475(61)	+0.00043(34)
20	4.5997	0.1352889	7	10	3130	3.9648(97)	+0.0005(4)*
	4.6066	0.1352731	7	10	2139	3.928(14)	
	4.6006	0.1352874	7	10	1519	3.982(14)	
	4.6032	0.1352812	18	10	4026	3.9337(87)	+0.00053(42)*
	4.6017	0.1352848	18	10	8478	3.9493(63)	+0.00100(21)
	4.6024	0.1352844	18	10	7620	3.9522(69)	+0.00013(30)*
24	4.70	0.1350531	1	10	1340	3.999(19)	+0.00384(71)
	4.7141	0.1350217	1	16	200	3.943(45)	+0.0023(11)
	4.7141	0.1350230	3	16	2216	3.945(11)	+0.00098(32)
	4.7141	0.1350237	2	16	10081	3.9582(54)	+0.00022(15)
	4.7131	0.1350260	2	16	5096	3.9620(80)	+0.00040(22)
	4.7165	0.1350181	3	16	7303	3.9492(64)	+0.00012(17)
32	4.90	0.1345991	3	20	5014	3.949(11)	+0.00543(34)

Table B.2: Tuning runs to fix $\bar{g}_{\text{GF}}^2(L_1) = \sigma(g_\star^2) = 5.867$ and $m_l = 0$. Boundary conditions and action are chosen exactly as in [45], and we use $T = L$. N_{rep} is the number of replica runs, while N_{ms} gives the number of configurations, separated by τ_{ms} molecular dynamics units. The trajectory length is $\tau = 2$ throughout. The ensembles where the coupling is printed in bold will be used for the matching.

L_1/a	β	κ	N_{rep}	$\frac{\tau_{\text{ms}}}{\text{MD}}$	N_{ms}	\bar{g}_{GF}^2	$L_1 m_l$
8	3.6537	0.1370722	4	4	15000	5.8648(70)	+0.00209(80)
12	3.8349	0.1369654	4	4	15004	5.8697(85)	-0.00112(49)
16	4.0018	0.1366803	4	4	17999	5.865(11)	+0.00036(33)
20	4.1363	0.1364075	20	10	5654	5.892(16)	-0.00858(36)
	4.1408	0.1363825	20	4	21682	5.858(12)	+0.00129(22)
	4.1394	0.1363857	20	4	21101	5.858(12)	+0.00149(21)
24	4.2511	0.1360984	1	10	1692	5.901(33)	+0.02482(49)
	4.2530	0.1361224	2	4	20002	5.876(14)	+0.00216(20)

Table B.3: Tuning runs to fix $\bar{g}_{\text{GF}}^2(L_2) = \sigma(\sigma(g_*^2)) = 11.27$ and $m_1 = 0$. Boundary conditions and action are chosen exactly as in [45], and we use $T = L$. N_{rep} is the number of replica runs, while N_{ms} gives the number of configurations, separated by τ_{ms} molecular dynamics units. $N_{\text{ms}}^{(0)}$ gives the number of configurations in the sector of vanishing topological charge. The trajectory length is $\tau = 2$ throughout. The ensembles where the coupling is printed in bold will be used for the matching.

L_2/a	β	κ	N_{rep}	$\frac{\tau_{\text{ms}}}{\text{MD}}$	N_{ms}	$N_{\text{ms}}^{(0)}$	\bar{g}_{GF}^2	$L_2 m_1$
12	3.40118	0.1368077	4	4	6400	4027	11.360(86)	+0.0114(33)
	3.4028	0.1368830	4	8	4004	2679	11.174(69)	-0.0250(35)
	3.4014	0.1368790	4	8	2504	1742	11.046(79)	-0.0328(36)
	3.4014	0.1368690	4	8	4004	2543	11.165(70)	-0.0255(33)
	3.4014	0.1368390	4	8	4004	2733	11.240(71)	-0.0095(32)
	3.4014	0.1368240	4	8	5004	3319	11.288(60)	-0.0029(25)
16	3.55132	0.1371546	4	4	4204	2974	11.37(14)	-0.0089(25)
	3.5515	0.1371530	4	8	6000	3943	11.228(56)	-0.0098(14)
	3.5515	0.1371430	12	8	18000	11896	11.326(39)	-0.0033(9)
	3.5522	0.1371379	4	8	6000	4014	11.299(66)	+0.0003(13)
20	3.6868	0.1371474	10	4	9799	7434	11.397(70)	+0.0006(9)
	3.6900	0.1371452	10	4	10000	7027	11.25(10)	+0.0013(10)
24	3.7989	0.1370354	2	4	2754	1568	11.62(15)	+0.0075(15)
	3.7996	0.1370405	2	8	7500	4534	11.46(10)	+0.0014(8)
	3.8013	0.1370387	2	8	8000	4991	11.328(90)	-0.0001(7)

B.2 RENORMALIZATION AND IMPROVEMENT FROM FINITE-VOLUME SIMULATIONS

In this appendix, tables (B.4)–(B.7) list the results for all ensembles entering our analysis for the determination of Z , b_m and $b_A - b_P$ in the coupling range of large-volume CLS simulations.

Table B.4: Sea quark PCAC masses and estimators R_{AP} , R_m and R_Z for LCP-0 in the sector of vanishing topological charge. Reproduced from [74].

ID	$m_{11}^{(0)}$	$R_{AP}^{(0)}$	$R_m^{(0)}$	$R_Z^{(0)}$
A1k1	-0.00278(80)	-1.016(119)	1.344(157)	0.7463(83)
A1k3	0.00079(118)	-0.866(156)	1.192(141)	0.7420(96)
A1k4	-0.00110(36)	-0.744(89)	1.395(73)	0.7490(48)
E1k1	0.00262(26)	-0.755(84)	0.313(66)	0.8687(63)
E1k2	-0.00022(22)	-0.816(57)	0.290(57)	0.8768(42)
B1k1	0.00549(21)	-0.352(61)	-0.247(45)	0.9676(48)
B1k2	0.00444(25)	-0.271(90)	-0.272(73)	0.9761(67)
B1k3	0.00107(20)	-0.545(62)	-0.241(59)	0.9701(47)
B1k4	-0.00057(19)	-0.486(68)	-0.316(51)	0.9798(44)
C1k2	0.00600(11)	-0.222(35)	-0.523(40)	1.0489(29)
C1k3	-0.00109(11)	-0.304(54)	-0.698(50)	1.0606(36)
D1k2	0.00079(10)	-0.205(103)	-0.684(59)	1.0849(52)
D1k4	-0.00007(3)	-0.152(20)	-0.743(20)	1.0885(11)

Table B.5: Sea quark PCAC masses and estimators R_{AP} , R_m and R_Z for LCP-0 for all topological sectors. Reproduced from [74].

ID	$m_{11}^{(\text{all})}$	$R_{AP}^{(\text{all})}$	$R_m^{(\text{all})}$	$R_Z^{(\text{all})}$
A1k1	-0.00166(61)	-0.876(104)	1.252(100)	0.7458(55)
A1k3	0.00262(130)	-0.587(133)	1.139(102)	0.7402(78)
A1k4	0.00030(29)	-0.583(61)	1.251(49)	0.7485(36)
E1k1	0.00308(22)	-0.652(89)	0.272(51)	0.8684(58)
E1k2	0.00034(18)	-0.757(46)	0.354(38)	0.8715(32)
B1k1	0.00562(14)	-0.375(37)	-0.286(31)	0.9674(29)
B1k2	0.00481(19)	-0.333(57)	-0.230(45)	0.9677(44)
B1k3	0.00164(16)	-0.575(43)	-0.136(42)	0.9589(35)
B1k4	0.00002(14)	-0.475(44)	-0.215(37)	0.9712(29)
C1k2	0.00619(7)	-0.214(26)	-0.498(30)	1.0461(25)
C1k3	-0.00086(8)	-0.288(42)	-0.594(40)	1.0547(26)
D1k2	0.00084(8)	-0.147(70)	-0.637(79)	1.0837(47)
D1k4	-0.00002(3)	-0.144(17)	-0.702(18)	1.0867(11)

Table B.6: Sea quark PCAC masses and estimators R_{AP} , R_m and R_Z for LCP-1 in the sector of vanishing topological charge. Reproduced from [74].

ID	$m_{PCAC}^{(0)}$	$R_{AP}^{(0)}$	$R_m^{(0)}$	$R_Z^{(0)}$
A1k1	-0.00278(80)	-0.443(25)	0.065(25)	0.7805(52)
A1k3	0.00079(118)	-0.349(34)	0.011(28)	0.7872(60)
A1k4	-0.00110(36)	-0.383(13)	-0.016(10)	0.7889(27)
E1k1	0.00262(26)	-0.329(16)	-0.264(14)	0.8971(36)
E1k2	-0.00022(22)	-0.365(14)	-0.264(13)	0.8994(28)
B1k1	0.00549(21)	-0.197(15)	-0.439(12)	0.9786(30)
B1k2	0.00444(25)	-0.172(21)	-0.457(21)	0.9845(43)
B1k3	0.00107(20)	-0.249(18)	-0.454(17)	0.9819(33)
B1k4	-0.00057(19)	-0.218(16)	-0.476(15)	0.9882(30)
C1k2	0.00600(11)	-0.099(10)	-0.586(11)	1.0541(19)
C1k3	-0.00109(11)	-0.130(17)	-0.653(16)	1.0623(27)
D1k2	0.00079(10)	-0.088(30)	-0.670(20)	1.0862(35)
D1k4	-0.00007(3)	-0.069(7)	-0.685(7)	1.0886(8)

Table B.7: Sea quark PCAC masses and estimators R_{AP} , R_m and R_Z for LCP-1 for all topological sectors. Reproduced from [74].

ID	$m_{PCAC}^{(all)}$	$R_{AP}^{(all)}$	$R_m^{(all)}$	$R_Z^{(all)}$
A1k1	-0.00166(61)	-0.423(19)	0.045(18)	0.7805(34)
A1k3	0.00262(130)	-0.324(25)	0.014(18)	0.7822(44)
A1k4	0.00030(29)	-0.366(9)	-0.013(7)	0.7864(18)
E1k1	0.00308(22)	-0.337(12)	-0.251(10)	0.8929(27)
E1k2	0.00034(18)	-0.360(10)	-0.239(9)	0.8948(21)
B1k1	0.00562(14)	-0.206(9)	-0.445(9)	0.9779(18)
B1k2	0.00481(19)	-0.197(14)	-0.440(13)	0.9782(28)
B1k3	0.00164(16)	-0.267(13)	-0.415(12)	0.9730(25)
B1k4	0.00002(14)	-0.230(10)	-0.439(11)	0.9810(21)
C1k2	0.00619(7)	-0.104(8)	-0.570(8)	1.0512(15)
C1k3	-0.00086(8)	-0.137(13)	-0.613(14)	1.0571(19)
D1k2	0.00084(8)	-0.081(22)	-0.646(22)	1.0847(32)
D1k4	-0.00002(3)	-0.070(6)	-0.669(6)	1.0871(8)

B.3 THE MASS OF THE CHARM QUARK

Table B.8: Overview of the measurements for the determination of the mass of the charm quark. We list the ensemble id and the number of configurations, on which the measurements have been performed. On each configuration we used 16 sources on the two time slices at a and $T - a$. We also quote the value for t_0/a^2 .

β	id	N_{cfg}	t_0/a^2
3.40	H101	2016	2.846(8)
	H102	2005	2.872(12)
	H105	1533	2.891(16)
	C101	1055	2.907(7)
3.46	H400	1045	3.634(13)
3.55	H200	2000	5.149(29)
	N202	899	5.140(21)
	N203	1543	5.143(7)
	N200	1544	5.159(8)
	D200	1191	5.167(10)
3.70	N300	2027	8.564(35)
	N302	2129	8.538(23)
	J303	1073	8.614(20)
3.85	J500	751	14.061(66)
	J501	1579	13.888(53)

Table B.9: Overview of the effective masses used in the calibration of the chiral-continuum extrapolation. The subscript of the flavor averages \overline{D} and spin-flavor averages $\overline{\overline{D}}$ correspond to those of the hopping parameters in table 14.1. No vector correlation functions have been determined for the J501 ensemble. Values are determined from a plateau fit. The corresponding values from the full fit are shown in 15.2.

id	am_π	am_K	$am_{\overline{D}_1}$	$am_{\overline{D}_2}$	$am_{\overline{\overline{D}}_1}$	$am_{\overline{\overline{D}}_2}$
H101	0.18287(65)	0.18287(65)	0.85286(93)	0.82186(84)	0.8876(45)	0.8586(45)
H102	0.15414(82)	0.19125(67)	0.85066(112)	0.81950(102)	0.8833(40)	0.8542(40)
H105	0.12093(164)	0.20201(109)	0.85087(162)	0.81974(155)	0.8843(69)	0.8554(68)
C101	0.09724(84)	0.20601(50)	0.85002(165)	0.81900(157)	0.8931(62)	0.8641(61)
H400	0.16372(66)	0.16372(66)	0.74508(100)	0.69391(97)	0.7795(40)	0.7313(38)
H200	0.13634(60)	0.13634(60)	0.64743(106)	0.60134(92)	0.6784(30)	0.6355(30)
N202	0.13403(36)	0.13403(36)	0.64413(86)	0.59824(68)	0.6725(23)	0.6297(25)
N203	0.11238(28)	0.14390(23)	0.64311(78)	0.59723(63)	0.6770(35)	0.6336(37)
N200	0.09215(38)	0.15053(29)	0.64373(69)	0.59794(56)	0.6777(27)	0.6343(27)
D200	0.06482(45)	0.15649(26)	0.64495(83)	0.59887(72)	0.6766(47)	0.6332(45)
N300	0.10630(55)	0.10630(55)	0.49369(177)	0.47605(157)	0.5198(25)	0.5036(28)
N302	0.08735(47)	0.11395(101)	0.49272(68)	0.47504(55)	0.5169(42)	0.5003(40)
J303	0.06482(24)	0.11971(23)	0.49204(68)	0.47402(68)	0.5208(27)	0.5039(27)
J500	0.08107(45)	0.08107(45)	0.37118(100)	0.35534(85)	0.3908(16)	0.3759(16)
J501	0.06599(42)	0.08794(35)	0.37204(67)	0.35602(62)		

Table B.10: Overview of the effective masses of the heavy mesons considered in this work for the hopping parameters given in tab. 14.1. No vector correlation functions have been determined on J501. Results have been obtained from the full fit

id	am_{D_1}	am_{D_2}	$am_{D_{s,1}}$	$am_{D_{s,2}}$	$am_{D_1^*}$	$am_{D_2^*}$	$am_{D_{s,1}^*}$	$am_{D_{s,2}^*}$
H101	0.85118(52)	0.82029(50)	0.85118(52)	0.82029(50)	0.90021(244)	0.87184(245)	0.90021(244)	0.87184(245)
H102	0.84595(65)	0.81493(63)	0.85801(57)	0.82709(63)	0.89605(242)	0.86759(247)	0.90821(169)	0.87987(172)
H105	0.84168(72)	0.81057(71)	0.86589(45)	0.83517(43)	0.88398(509)	0.85549(495)	0.91292(185)	0.88434(186)
C101	0.83938(80)	0.80828(75)	0.86966(33)	0.83884(32)	0.88613(260)	0.85763(254)	0.92014(78)	0.89176(80)
H400	0.74502(51)	0.69355(51)	0.74502(51)	0.69355(51)	0.79218(270)	0.74422(292)	0.79218(270)	0.74422(292)
H200	0.64697(61)	0.60109(56)	0.64697(61)	0.60109(56)	0.68803(142)	0.64582(147)	0.68803(142)	0.64582(147)
N202	0.64554(57)	0.59955(55)	0.64554(57)	0.59955(55)	0.68595(162)	0.64380(158)	0.68595(162)	0.64380(158)
N203	0.64025(70)	0.59434(63)	0.65088(39)	0.60522(33)	0.68074(296)	0.63827(294)	0.69203(148)	0.64972(151)
N200	0.63843(64)	0.59242(58)	0.65662(32)	0.61102(29)	0.67936(148)	0.63673(159)	0.69762(86)	0.65539(90)
D200	0.63461(59)	0.58867(52)	0.66136(28)	0.61587(27)	0.67136(301)	0.62892(318)	0.70097(68)	0.65873(69)
N300	0.49351(83)	0.47571(83)	0.49351(83)	0.47571(83)	0.52812(193)	0.51176(191)	0.52812(193)	0.51176(191)
N302	0.48961(87)	0.47188(74)	0.49897(32)	0.48121(29)	0.51761(584)	0.50122(582)	0.53011(123)	0.51383(137)
J303	0.48689(47)	0.46898(47)	0.50414(20)	0.48638(20)	0.52018(177)	0.50373(178)	0.53771(67)	0.52131(67)
J500	0.37099(78)	0.35515(67)	0.37099(78)	0.35515(67)	0.40105(99)	0.38648(97)	0.40105(99)	0.38648(97)
J501	0.36998(43)	0.35405(39)	0.37687(54)	0.36097(52)				

Table B.11: Overview of the effective masses of the heavy mesons considered in this work for the hopping parameters given in tab. 14.1. No vector correlation functions have been determined on J501. Results have been obtained from the plateau fit.

id	$a m_{D_1}$	$a m_{D_2}$	$a m_{D_{s,1}}$	$a m_{D_{s,2}}$	$a m_{D_1^*}$	$a m_{D_2^*}$	$a m_{D_{s,1}^*}$	$a m_{D_{s,2}^*}$
H101	0.85286(93)	0.82186(84)	0.85286(93)	0.82186(84)	0.89922(602)	0.87081(592)	0.89922(602)	0.87081(592)
H102	0.84694(135)	0.81570(123)	0.85810(81)	0.82710(77)	0.88630(666)	0.85768(649)	0.90262(348)	0.87446(350)
H105	0.84286(220)	0.81155(213)	0.86690(70)	0.83612(65)	0.88109(1206)	0.85303(1175)	0.90824(442)	0.87940(451)
C101	0.84046(237)	0.80934(225)	0.86915(50)	0.83832(44)	0.89055(1246)	0.86215(1225)	0.92200(246)	0.89360(255)
H400	0.74508(100)	0.69391(97)	0.74508(100)	0.69391(97)	0.79098(534)	0.74373(517)	0.79098(534)	0.74373(517)
H200	0.64743(106)	0.60134(92)	0.64743(106)	0.60134(92)	0.68875(404)	0.64689(394)	0.68875(404)	0.64689(394)
N202	0.64413(86)	0.59824(68)	0.64413(86)	0.59824(68)	0.68192(305)	0.64019(327)	0.68192(305)	0.64019(327)
N203	0.63956(87)	0.59347(71)	0.65022(51)	0.60475(47)	0.68157(544)	0.63891(578)	0.69458(316)	0.65196(320)
N200	0.63738(90)	0.59149(79)	0.65643(46)	0.61085(40)	0.67815(478)	0.63524(467)	0.69801(223)	0.65567(229)
D200	0.63660(125)	0.59026(103)	0.66164(37)	0.61608(37)	0.67108(891)	0.62819(871)	0.70274(202)	0.66045(192)
N300	0.49369(177)	0.47605(157)	0.49369(177)	0.47605(157)	0.52854(367)	0.51281(323)	0.52854(367)	0.51281(323)
N302	0.48958(96)	0.47194(78)	0.49900(48)	0.48124(48)	0.51861(745)	0.50247(713)	0.53153(342)	0.51508(319)
J303	0.48613(94)	0.46796(94)	0.50388(26)	0.48614(23)	0.52084(505)	0.50408(504)	0.53758(145)	0.52114(153)
J500	0.37118(100)	0.35534(85)	0.37118(100)	0.35534(85)	0.39730(216)	0.38276(214)	0.39730(216)	0.38276(214)
J501	0.36980(79)	0.35372(71)	0.37651(91)	0.36063(59)				

Table B.1.2: Bare current quark masses obtained with improved discretized derivatives. Results from standard derivatives are listed in 1.5-3

id	$a\bar{m}_l$	$a\bar{m}_{gl}$	$a\bar{m}_{ss}$	$a\bar{m}_{h_1}$	$a\bar{m}_{h_2}$	$a\bar{m}_{h_1s}$	$a\bar{m}_{h_2s}$	$a\bar{m}_{h_1h_1}$	$a\bar{m}_{h_2h_2}$
H101	0.009153(45)	0.009153(45)	0.009153(45)	0.164695(49)	0.155237(49)	0.164695(49)	0.155237(49)	0.284757(19)	0.274638(14)
H102	0.006466(52)	0.010104(52)	0.013704(50)	0.163358(45)	0.153873(44)	0.167030(41)	0.157584(39)	0.284822(21)	0.274671(14)
H105	0.003896(86)	0.011243(79)	0.018392(83)	0.162020(77)	0.152530(72)	0.169323(44)	0.159911(46)	0.284795(26)	0.274642(15)
C101	0.002495(42)	0.011791(43)	0.020953(64)	0.161367(71)	0.151885(58)	0.170661(35)	0.161241(33)	0.284843(16)	0.274667(10)
H400	0.008190(36)	0.008190(36)	0.008190(36)	0.143921(58)	0.128051(54)	0.143921(58)	0.128051(54)	0.263692(14)	0.238888(14)
H200	0.006835(24)	0.006835(24)	0.006835(24)	0.127113(31)	0.112718(39)	0.127113(31)	0.112718(39)	0.239901(9)	0.214508(10)
N202	0.006826(17)	0.006826(17)	0.006826(17)	0.127129(27)	0.112733(38)	0.127129(27)	0.112733(38)	0.239907(6)	0.214510(6)
N203	0.004735(17)	0.007871(12)	0.010993(12)	0.126094(28)	0.111740(32)	0.129254(22)	0.114886(21)	0.239909(6)	0.214508(5)
N200	0.003147(14)	0.008607(12)	0.014061(15)	0.125244(34)	0.110847(56)	0.130767(18)	0.116393(16)	0.239879(6)	0.214481(6)
D200	0.001529(12)	0.009341(12)	0.017106(10)	0.124424(33)	0.110066(82)	0.132339(11)	0.117978(13)	0.239865(4)	0.214461(4)
N300	0.005491(17)	0.005491(17)	0.005491(17)	0.096306(18)	0.090720(19)	0.096306(18)	0.090720(19)	0.185465(9)	0.174732(8)
N302	0.003718(13)	0.006390(13)	0.009057(12)	0.095425(9)	0.089835(9)	0.098118(9)	0.092530(12)	0.185466(5)	0.174735(4)
J303	0.002049(8)	0.007180(8)	0.012301(6)	0.094594(16)	0.089002(15)	0.099755(8)	0.094167(8)	0.185456(3)	0.174722(3)
J500	0.004205(5)	0.004205(5)	0.004205(5)	0.070512(9)	0.065673(9)	0.070512(9)	0.065673(9)	0.136497(4)	0.126937(3)
J501	0.002737(7)	0.004948(6)	0.007157(6)	0.069775(8)	0.064937(8)	0.072000(5)	0.067163(6)	0.136501(3)	0.126939(3)

B.4 HQET IN LARGE VOLUME

Table B.13: Renormalized couplings \hat{g}_{11} from the summed ratio evaluated with HYP1 and HYP2 static quark actions.

id	$\hat{g}_{11}^{\text{HYP1}}$	$\hat{g}_{11}^{\text{HYP2}}$
H101	0.514(9)	0.519(7)
H102	0.526(9)	0.524(6)
N101	0.514(22)	0.516(15)
C101	0.518(15)	0.514(12)
S400	0.500(9)	0.508(8)
H200	0.530(8)	0.530(7)
N203	0.509(10)	0.509(9)
N200	0.507(15)	0.509(12)

BIBLIOGRAPHY

- [1] K. G. Wilson. "Confinement of Quarks." In: *Phys. Rev. D* 10 (1974), pp. 2445–2459. DOI: 10.1103/PhysRevD.10.2445.
- [2] M. Tanabashi et al. "Review of Particle Physics." In: *Phys. Rev. D* 98.3 (2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001.
- [3] E. Waheed et al. "Measurement of the CKM matrix element $|V_{cb}|$ from $B^0 \rightarrow D^{*-} \ell^+ \nu_\ell$ at Belle." In: *Phys. Rev. D* 100.5 (2019), p. 052007. DOI: 10.1103/PhysRevD.100.052007. arXiv: 1809.03290 [hep-ex].
- [4] K. Symanzik. "Continuum Limit and Improved Action in Lattice Theories. 1. Principles and ϕ^4 Theory." In: *Nucl. Phys. B* 226 (1983), pp. 187–204. DOI: 10.1016/0550-3213(83)90468-6.
- [5] K. Symanzik. "Continuum Limit and Improved Action in Lattice Theories. 2. O(N) Nonlinear Sigma Model in Perturbation Theory." In: *Nucl. Phys. B* 226 (1983), pp. 205–227. DOI: 10.1016/0550-3213(83)90469-8.
- [6] B. Dehnadi, A. H. Hoang, and V. Mateu. "Bottom and Charm Mass Determinations with a Convergence Test." In: *JHEP* 08 (2015). Ed. by A. A. Petrov, p. 155. DOI: 10.1007/JHEP08(2015)155. arXiv: 1504.07638 [hep-ph].
- [7] M. E. Peskin and D. V. Schroeder. *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995. ISBN: 978-0-201-50397-5.
- [8] H. B. Nielsen and M. Ninomiya. "Absence of Neutrinos on a Lattice. 1. Proof by Homotopy Theory." In: *Nucl. Phys. B* 185 (1981). [Erratum: *Nucl. Phys. B* 195,541(1982); ,533(1980)], p. 20. DOI: 10.1016/0550-3213(81)90361-8, 10.1016/0550-3213(82)90011-6.
- [9] H. B. Nielsen and M. Ninomiya. "Absence of Neutrinos on a Lattice. 2. Intuitive Topological Proof." In: *Nucl. Phys. B* 193 (1981), pp. 173–194. DOI: 10.1016/0550-3213(81)90524-1.
- [10] H. B. Nielsen and M. Ninomiya. "No Go Theorem for Regularizing Chiral Fermions." In: *Phys. Lett.* 105B (1981), pp. 219–223. DOI: 10.1016/0370-2693(81)91026-1.
- [11] D. B. Kaplan. "Chiral Symmetry and Lattice Fermions." In: *Modern perspectives in lattice QCD: Quantum field theory and high performance computing. Proceedings, International School, 93rd Session, Les Houches, France, August 3-28, 2009*. 2009, pp. 223–272. arXiv: 0912.2560 [hep-lat].

- [12] P. H. Ginsparg and K. G. Wilson. "A Remnant of Chiral Symmetry on the Lattice." In: *Phys. Rev. D* 25 (1982), p. 2649. DOI: 10.1103/PhysRevD.25.2649.
- [13] I. Montvay and G. Münster. *Quantum fields on a lattice*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1997. ISBN: 9780521599177, 9780511879197. DOI: 10.1017/CB09780511470783.
- [14] C. Gattringer and C. B. Lang. "Quantum chromodynamics on the lattice." In: *Lect. Notes Phys.* 788 (2010), pp. 1–343. DOI: 10.1007/978-3-642-01850-3.
- [15] R. Gupta. "Introduction to lattice QCD: Course." In: *Les Houches Summer School in Theoretical Physics, Session 68: Probing the Standard Model of Particle Interactions*. July 1997, pp. 83–219. arXiv: hep-lat/9807028.
- [16] G. Münster and M. Walzl. "Lattice gauge theory: A Short primer." In: *Zuo Summer School on Phenomenology of Gauge Interactions*. Aug. 2000, pp. 127–160. arXiv: hep-lat/0012005.
- [17] U.-J. Wiese. "An Introduction to Lattice Field Theory." In: *15th Saalburg Summer School* (2009).
- [18] C.-N. Yang and R. L. Mills. "Conservation of Isotopic Spin and Isotopic Gauge Invariance." In: *Phys. Rev.* 96 (1954), pp. 191–195. DOI: 10.1103/PhysRev.96.191.
- [19] M. Lüscher and P. Weisz. "On-Shell Improved Lattice Gauge Theories." In: *Commun. Math. Phys.* 97 (1985). [Erratum: *Commun. Math. Phys.* 98,433(1985)], p. 59. DOI: 10.1007/BF01206178.
- [20] K. Symanzik. "Some topics in quantum field theory." In: *Mathematical Problems in Theoretical Physics*. Ed. by R. Schrader, R. Seiler, and D. A. Uhlenbrock. Berlin, Heidelberg: Springer Berlin Heidelberg, 1982, pp. 47–58. ISBN: 978-3-540-38982-8.
- [21] P. Weisz. "Continuum Limit Improved Lattice Action for Pure Yang-Mills Theory. 1." In: *Nucl. Phys.* B212 (1983), pp. 1–17. DOI: 10.1016/0550-3213(83)90595-3.
- [22] P. Weisz and R. Wohlert. "Continuum Limit Improved Lattice Action for Pure Yang-Mills Theory. 2." In: *Nucl. Phys.* B236 (1984). [Erratum: *Nucl. Phys.* B247,544(1984)], p. 397. DOI: 10.1016/0550-3213(84)90563-7, 10.1016/0550-3213(84)90543-1.
- [23] G. Curci, P. Menotti, and G. Paffuti. "Symanzik's Improved Lagrangian for Lattice Gauge Theory." In: *Phys. Lett.* 130B (1983). [Erratum: *Phys. Lett.* 135B,516(1984)], p. 205. DOI: 10.1016/0370-2693(83)91043-2.

- [24] M. Lüscher and P. Weisz. "Computation of the Action for On-Shell Improved Lattice Gauge Theories at Weak Coupling." In: *Phys. Lett.* 158B (1985), pp. 250–254. DOI: 10.1016/0370-2693(85)90966-9.
- [25] Y. Iwasaki. *Renormalization Group Analysis of Lattice Theories and Improved Lattice Action. II. Four-dimensional non-Abelian SU(N) gauge model.* 1983. arXiv: 1111.7054 [hep-lat].
- [26] M. Lüscher. "Construction of a Selfadjoint, Strictly Positive Transfer Matrix for Euclidean Lattice Gauge Theories." In: *Commun. Math. Phys.* 54 (1977), p. 283. DOI: 10.1007/BF01614090.
- [27] K. Osterwalder and R. Schrader. "Axioms for Euclidean Green's functions." In: *Commun. Math. Phys.* 31 (1973), pp. 83–112. DOI: 10.1007/BF01645738.
- [28] K. Osterwalder and R. Schrader. "Axioms for Euclidean Green's functions II." In: *Commun. Math. Phys.* 42 (1975), p. 281. DOI: 10.1007/BF01608978.
- [29] M. Lüscher and P. Weisz. "Definition and General Properties of the Transfer Matrix in Continuum Limit Improved Lattice Gauge Theories." In: *Nucl. Phys.* B240 (1984), pp. 349–361. DOI: 10.1016/0550-3213(84)90270-0.
- [30] S. Necco. "Universality and scaling behavior of RG gauge actions." In: *Nucl. Phys.* B683 (2004), pp. 137–167. DOI: 10.1016/j.nuclphysb.2004.01.032. arXiv: hep-lat/0309017 [hep-lat].
- [31] M. Lüscher, S. Sint, R. Sommer, and P. Weisz. "Chiral symmetry and O(a) improvement in lattice QCD." In: *Nucl. Phys.* B478 (1996), pp. 365–400. DOI: 10.1016/0550-3213(96)00378-1. arXiv: hep-lat/9605038 [hep-lat].
- [32] M. Lüscher. "Advanced lattice QCD." In: *Probing the standard model of particle interactions. Proceedings, Summer School in Theoretical Physics, NATO Advanced Study Institute, 68th session, Les Houches, France, July 28-September 5, 1997. Pt. 1, 2.* 1998, pp. 229–280. arXiv: hep-lat/9802029 [hep-lat].
- [33] B. Sheikholeslami and R. Wohlert. "Improved Continuum Limit Lattice Action for QCD with Wilson Fermions." In: *Nucl. Phys.* B259 (1985), p. 572. DOI: 10.1016/0550-3213(85)90002-1.
- [34] J. Bulava and S. Schaefer. "Improvement of $N_f = 3$ lattice QCD with Wilson fermions and tree-level improved gauge action." In: *Nucl. Phys.* B874 (2013), pp. 188–197. DOI: 10.1016/j.nuclphysb.2013.05.019. arXiv: 1304.7093 [hep-lat].
- [35] S. Borsanyi et al. "Ab initio calculation of the neutron-proton mass difference." In: *Science* 347 (2015), pp. 1452–1455. DOI: 10.1126/science.1257050. arXiv: 1406.4088 [hep-lat].

- [36] K. Symanzik. "Schrödinger Representation and Casimir Effect in Renormalizable Quantum Field Theory." In: *Nucl. Phys.* B190 (1981), pp. 1–44. DOI: 10.1016/0550-3213(81)90482-X.
- [37] M. Lüscher. "Schrödinger representation in quantum field theory." In: *Nucl. Phys.* B254 (1985), pp. 52–57. DOI: 10.1016/0550-3213(85)90210-X.
- [38] M. Lüscher, R. Narayanan, P. Weisz, and U. Wolff. "The Schrödinger functional: A Renormalizable probe for non-Abelian gauge theories." In: *Nucl. Phys.* B384 (1992), pp. 168–228. DOI: 10.1016/0550-3213(92)90466-0. arXiv: hep-lat/9207009 [hep-lat].
- [39] S. Sint. "On the Schrödinger functional in QCD." In: *Nucl. Phys.* B421 (1994), pp. 135–158. DOI: 10.1016/0550-3213(94)90228-3. arXiv: hep-lat/9312079 [hep-lat].
- [40] R. Sommer. "Nonperturbative renormalization of QCD." In: *Lect. Notes Phys.* 512 (1998), pp. 65–113. DOI: 10.1007/BFb0106893. arXiv: hep-ph/9711243 [hep-ph].
- [41] R. Sommer. "Non-perturbative QCD: Renormalization, O(a)-improvement and matching to Heavy Quark Effective Theory." In: *Workshop on Perspectives in Lattice QCD Nara, Japan, October 31-November 11, 2005*. 2006. arXiv: hep-lat/0611020 [hep-lat].
- [42] S. Sint and P. Weisz. "The Running quark mass in the SF scheme and its two loop anomalous dimension." In: *Nucl. Phys.* B545 (1999), pp. 529–542. DOI: 10.1016/S0550-3213(98)00874-8. arXiv: hep-lat/9808013 [hep-lat].
- [43] S. Sint and R. Sommer. "The Running coupling from the QCD Schrödinger functional: A One loop analysis." In: *Nucl. Phys.* B465 (1996), pp. 71–98. DOI: 10.1016/0550-3213(96)00020-X. arXiv: hep-lat/9508012 [hep-lat].
- [44] S. Aoki, R. Frezzotti, and P. Weisz. "Computation of the improvement coefficient $c(SW)$ to one loop with improved gluon actions." In: *Nucl. Phys.* B540 (1999), pp. 501–519. DOI: 10.1016/S0550-3213(98)00742-1. arXiv: hep-lat/9808007 [hep-lat].
- [45] M. Dalla Brida, P. Fritzsche, T. Korzec, A. Ramos, S. Sint, and R. Sommer. "Slow running of the Gradient Flow coupling from 200 MeV to 4 GeV in $N_f = 3$ QCD." In: *Phys. Rev.* D95.1 (2017), p. 014507. DOI: 10.1103/PhysRevD.95.014507. arXiv: 1607.06423 [hep-lat].
- [46] S. Takeda, S. Aoki, and K. Ide. "A Perturbative determination of O(a) boundary improvement coefficients for the Schrödinger functional coupling at one loop with improved gauge actions." In: *Phys. Rev.* D68 (2003), p. 014505. DOI: 10.1103/PhysRevD.68.014505. arXiv: hep-lat/0304013 [hep-lat].

- [47] M. Lüscher. “The Schrödinger functional in lattice QCD with exact chiral symmetry.” In: *JHEP* 05 (2006), p. 042. DOI: 10.1088/1126-6708/2006/05/042. arXiv: hep-lat/0603029 [hep-lat].
- [48] K. Jansen, C. Liu, M. Lüscher, H. Simma, S. Sint, R. Sommer, P. Weisz, and U. Wolff. “Nonperturbative renormalization of lattice QCD at all scales.” In: *Phys. Lett. B* 372 (1996), pp. 275–282. DOI: 10.1016/0370-2693(96)00075-5. arXiv: hep-lat/9512009 [hep-lat].
- [49] M. Lüscher and P. Weisz. “O(a) improvement of the axial current in lattice QCD to one loop order of perturbation theory.” In: *Nucl. Phys. B* 479 (1996), pp. 429–458. DOI: 10.1016/0550-3213(96)00448-8. arXiv: hep-lat/9606016 [hep-lat].
- [50] F. Joswig, H. Simma, and C. Wittemeier. *Correlation functions in sfcf*. Internal notes of the ALPHA collaboration. 2019.
- [51] M. Lüscher, S. Sint, R. Sommer, and H. Wittig. “Nonperturbative determination of the axial current normalization constant in O(a) improved lattice QCD.” In: *Nucl. Phys. B* 491 (1997), pp. 344–364. DOI: 10.1016/S0550-3213(97)00087-4. arXiv: hep-lat/9611015 [hep-lat].
- [52] C. Wittemeier. “Implementation of a Program for QCD and HQET Correlation Functions in the Schrödinger Functional.” MA thesis. University of Münster, 2012.
- [53] M. Lüscher and S. Schaefer. *openQCD*. <http://luscher.web.cern.ch/luscher/openQCD/>.
- [54] L. Del Debbio, H. Panagopoulos, and E. Vicari. “ θ dependence of SU(N) gauge theories.” In: *JHEP* 08 (2002), p. 044. DOI: 10.1088/1126-6708/2002/08/044. arXiv: hep-th/0204125 [hep-th].
- [55] L. Del Debbio, G. M. Manca, and E. Vicari. “Critical slowing down of topological modes.” In: *Phys. Lett. B* 594 (2004), pp. 315–323. DOI: 10.1016/j.physletb.2004.05.038. arXiv: hep-lat/0403001.
- [56] S. Schaefer, R. Sommer, and F. Virotta. “Critical slowing down and error analysis in lattice QCD simulations.” In: *Nucl. Phys. B* 845 (2011), pp. 93–119. DOI: 10.1016/j.nuclphysb.2010.11.020. arXiv: 1009.5228 [hep-lat].
- [57] M. Lüscher. “Topology, the Wilson flow and the HMC algorithm.” In: *PoS LATTICE2010* (2010), p. 015. DOI: 10.22323/1.105.0015. arXiv: 1009.5877 [hep-lat].
- [58] M. Lüscher and S. Schaefer. “Lattice QCD without topology barriers.” In: *JHEP* 07 (2011), p. 036. DOI: 10.1007/JHEP07(2011)036. arXiv: 1105.4749 [hep-lat].

- [59] M. Lüscher and S. Schaefer. "Lattice QCD with open boundary conditions and twisted-mass reweighting." In: *Comput. Phys. Commun.* 184 (2013), pp. 519–528. DOI: 10.1016/j.cpc.2012.10.003. arXiv: 1206.2809 [hep-lat].
- [60] M. Bruno et al. "Simulation of QCD with $N_f = 2 + 1$ flavors of non-perturbatively improved Wilson fermions." In: *JHEP* 02 (2015), p. 043. DOI: 10.1007/JHEP02(2015)043. arXiv: 1411.3982 [hep-lat].
- [61] G. Heatlie, G. Martinelli, C. Pittori, G. C. Rossi, and C. T. Sachrajda. "The improvement of hadronic matrix elements in lattice QCD." In: *Nucl. Phys.* B352 (1991), pp. 266–288. DOI: 10.1016/0550-3213(91)90137-M.
- [62] M. Lüscher, S. Sint, R. Sommer, P. Weisz, and U. Wolff. "Non-perturbative $O(a)$ improvement of lattice QCD." In: *Nucl. Phys.* B491 (1997), pp. 323–343. DOI: 10.1016/S0550-3213(97)00080-1. arXiv: hep-lat/9609035 [hep-lat].
- [63] T. Bhattacharya, R. Gupta, W. Lee, S. R. Sharpe, and J. M. S. Wu. "Improved bilinears in lattice QCD with non-degenerate quarks." In: *Phys. Rev. D* 73 (2006), p. 034504. DOI: 10.1103/PhysRevD.73.034504. arXiv: hep-lat/0511014 [hep-lat].
- [64] I. Campos, P. Fritzsche, C. Pena, D. Preti, A. Ramos, and A. Vladikas. "Non-perturbative quark mass renormalisation and running in $N_f = 3$ QCD." In: *Eur. Phys. J.* C78.5 (2018), p. 387. DOI: 10.1140/epjc/s10052-018-5870-5. arXiv: 1802.05243 [hep-lat].
- [65] J. Bulava, M. Della Morte, J. Heitger, and C. Wittemeier. "Non-perturbative renormalization of the axial current in $N_f = 3$ lattice QCD with Wilson fermions and a tree-level improved gauge action." In: *Phys. Rev. D* 93.11 (2016), p. 114513. DOI: 10.1103/PhysRevD.93.114513. arXiv: 1604.05827 [hep-lat].
- [66] M. Dalla Brida, T. Korzec, S. Sint, and P. Vilaseca. "High precision renormalization of the flavour non-singlet Noether currents in lattice QCD with Wilson quarks." In: *Eur. Phys. J.* C79.1 (2019), p. 23. DOI: 10.1140/epjc/s10052-018-6514-5. arXiv: 1808.09236 [hep-lat].
- [67] J. Bulava, M. Della Morte, J. Heitger, and C. Wittemeier. "Non-perturbative improvement of the axial current in $N_f=3$ lattice QCD with Wilson fermions and tree-level improved gauge action." In: *Nucl. Phys.* B896 (2015), pp. 555–568. DOI: 10.1016/j.nuclphysb.2015.05.003. arXiv: 1502.04999 [hep-lat].
- [68] A. Gérardin, T. Harris, and H. B. Meyer. "Nonperturbative renormalization and $O(a)$ -improvement of the nonsinglet vector current with $N_f = 2 + 1$ Wilson fermions and tree-level Symanzik improved gauge action." In: *Phys. Rev. D* 99.1 (2019),

- p. 014519. DOI: 10.1103/PhysRevD.99.014519. arXiv: 1811.08209 [hep-lat].
- [69] J. Heitger, F. Joswig, A. Vladikas, and C. Wittemeier. “Non-perturbative determination of c_V, Z_V and Z_S/Z_P in $N_f = 3$ lattice QCD.” In: *EPJ Web Conf.* 175 (2018), p. 10004. DOI: 10.1051/epjconf/201817510004. arXiv: 1711.03924 [hep-lat].
- [70] L. Chimirri, P. Fritzsche, J. Heitger, F. Joswig, M. Panero, C. Pena, and D. Preti. “Non-perturbative renormalization of $O(a)$ improved tensor currents.” In: *PoS LATTICE2019* (2019), p. 212. arXiv: 1910.06759 [hep-lat].
- [71] P. Korcyl and G. S. Bali. “Non-perturbative determination of improvement coefficients using coordinate space correlators in $N_f = 2 + 1$ lattice QCD.” In: *Phys. Rev. D* 95.1 (2017), p. 014505. DOI: 10.1103/PhysRevD.95.014505. arXiv: 1607.07090 [hep-lat].
- [72] P. Korcyl and G. S. Bali. “Non-perturbative determination of improvement coefficients using coordinate space correlators in $N_f = 2 + 1$ lattice QCD.” In: *PoS LATTICE2016* (2016), p. 190. DOI: 10.22323/1.256.0190. arXiv: 1609.09477 [hep-lat].
- [73] P. Fritzsche. “Mass-improvement of the vector current in three-flavor QCD.” In: *JHEP* 06 (2018), p. 015. DOI: 10.1007/JHEP06(2018)015. arXiv: 1805.07401 [hep-lat].
- [74] G. M. de Divitiis, P. Fritzsche, J. Heitger, C. C. Köster, S. Kuberski, and A. Vladikas. “Non-perturbative determination of improvement coefficients b_m and $b_A - b_P$ and normalisation factor $Z_m Z_P/Z_A$ with $N_f = 3$ Wilson fermions.” In: *Eur. Phys. J. C* 79.9 (2019), p. 797. DOI: 10.1140/epjc/s10052-019-7287-1. arXiv: 1906.03445 [hep-lat].
- [75] M. Bochicchio, L. Maiani, G. Martinelli, G. C. Rossi, and M. Testa. “Chiral Symmetry on the Lattice with Wilson Fermions.” In: *Nucl. Phys. B* 262 (1985), p. 331. DOI: 10.1016/0550-3213(85)90290-1.
- [76] L. Maiani and G. Martinelli. “Current Algebra and Quark Masses from a Monte Carlo Simulation with Wilson Fermions.” In: *Phys. Lett. B* 178 (1986), pp. 265–271. DOI: 10.1016/0370-2693(86)91508-X.
- [77] M. Bruno, I. Campos, P. Fritzsche, J. Koponen, C. Pena, D. Preti, A. Ramos, and A. Vladikas. “Light quark masses in $N_f = 2 + 1$ lattice QCD with Wilson fermions.” In: *Eur. Phys. J. C* 80.2 (2020), p. 169. DOI: 10.1140/epjc/s10052-020-7698-z. arXiv: 1911.08025 [hep-lat].

- [78] S. Dürr, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, T. Kurth, L. Lellouch, T. Lippert, K. K. Szabo, and G. Vulvert. "Lattice QCD at the physical point: Simulation and analysis details." In: *JHEP* 08 (2011), p. 148. DOI: 10.1007/JHEP08(2011)148. arXiv: 1011.2711 [hep-lat].
- [79] A. Ukawa. "Computational cost of full QCD simulations experienced by CP-PACS and JLQCD Collaborations." In: *Nucl. Phys. Proc. Suppl.* 106 (2002), pp. 195–196. DOI: 10.1016/S0920-5632(01)01662-0.
- [80] M. Lüscher. "Local coherence and deflation of the low quark modes in lattice QCD." In: *JHEP* 07 (2007), p. 081. DOI: 10.1088/1126-6708/2007/07/081. arXiv: 0706.2298 [hep-lat].
- [81] M. Lüscher. "Deflation acceleration of lattice QCD simulations." In: *JHEP* 12 (2007), p. 011. DOI: 10.1088/1126-6708/2007/12/011. arXiv: 0710.5417 [hep-lat].
- [82] A. Stathopoulos and K. Orginos. "Computing and deflating eigenvalues while solving multiple right hand side linear systems in quantum chromodynamics." In: *SIAM J. Sci. Comput.* 32 (2010), pp. 439–462. DOI: 10.1137/080725532. arXiv: 0707.0131 [hep-lat].
- [83] A. Abdel-Rehim, K. Orginos, and A. Stathopoulos. "Extending the eigCG algorithm to non-symmetric linear systems with multiple right-hand sides." In: *PoS LAT2009* (2009), p. 036. DOI: 10.22323/1.091.0036. arXiv: 0911.2285 [hep-lat].
- [84] R. Babich, J. Brannick, R. C. Brower, M. A. Clark, T. A. Manteuffel, S. F. McCormick, J. C. Osborn, and C. Rebbi. "Adaptive multigrid algorithm for the lattice Wilson-Dirac operator." In: *Phys. Rev. Lett.* 105 (2010), p. 201602. DOI: 10.1103/PhysRevLett.105.201602. arXiv: 1005.3043 [hep-lat].
- [85] A. Frommer, A. Nobile, and P. Zingler. *Deflation and Flexible SAP-Preconditioning of GMRES in Lattice QCD Simulation*. 2012. arXiv: 1204.5463 [hep-lat].
- [86] A. Frommer, K. Kahl, S. Krieg, B. Leder, and M. Rottmann. "Adaptive Aggregation Based Domain Decomposition Multigrid for the Lattice Wilson Dirac Operator." In: *SIAM J. Sci. Comput.* 36 (2014), A1581–A1608. DOI: 10.1137/130919507. arXiv: 1303.1377 [hep-lat].
- [87] M. Lüscher. "Computational Strategies in Lattice QCD." In: *Modern perspectives in lattice QCD: Quantum field theory and high performance computing. Proceedings, International School, 93rd Session, Les Houches, France, August 3-28, 2009*. 2010, pp. 331–399. arXiv: 1002.4232 [hep-lat].

- [88] F. Knechtli, M. Günther, and M. Peardon. *Lattice Quantum Chromodynamics: Practical Essentials*. SpringerBriefs in Physics. Springer, 2017. ISBN: 9789402409970, 9789402409994. DOI: 10.1007/978-94-024-0999-4. URL: http://www.springer.com/0/AVjATAj-VqkG3_-0sK4X.
- [89] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. "Equation of state calculations by fast computing machines." In: *J. Chem. Phys.* 21 (1953), pp. 1087–1092. DOI: 10.1063/1.1699114.
- [90] D. J. E. Callaway and A. Rahman. "Lattice Gauge Theory in Microcanonical Ensemble." In: *Phys. Rev. D* 28 (1983), p. 1506. DOI: 10.1103/PhysRevD.28.1506.
- [91] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth. "Hybrid Monte Carlo." In: *Phys. Lett.* B195 (1987), pp. 216–222. DOI: 10.1016/0370-2693(87)91197-X.
- [92] S. A. Gottlieb, W. Liu, D. Toussaint, R. L. Renken, and R. L. Sugar. "Hybrid Molecular Dynamics Algorithms for the Numerical Simulation of Quantum Chromodynamics." In: *Phys. Rev. D* 35 (1987), pp. 2531–2542. DOI: 10.1103/PhysRevD.35.2531.
- [93] M. Lüscher. *Molecular-dynamics quark forces*. <http://luscher.web.cern.ch/luscher/openQCD/>. 2012.
- [94] H. B. Meyer, H. Simma, R. Sommer, M. Della Morte, O. Witzel, and U. Wolff. "Exploring the HMC trajectory-length dependence of autocorrelation times in lattice QCD." In: *Comput. Phys. Commun.* 176 (2007), pp. 91–97. DOI: 10.1016/j.cpc.2006.08.002. arXiv: hep-lat/0606004 [hep-lat].
- [95] I. Omelyan, I. Mryglod, and R. Folk. "Symplectic analytically integrable decomposition algorithms: classification, derivation, and application to molecular dynamics, quantum and celestial mechanics simulations." In: *Computer Physics Communications* 151.3 (2003), pp. 272–314. ISSN: 0010-4655. DOI: [https://doi.org/10.1016/S0010-4655\(02\)00754-3](https://doi.org/10.1016/S0010-4655(02)00754-3).
- [96] I. P. Omelyan, I. M. Mryglod, and R. Folk. "Optimized Verlet-like algorithms for molecular dynamics simulations." In: *Phys. Rev. E* 65 (5 2002), p. 056706. DOI: 10.1103/PhysRevE.65.056706.
- [97] T. Takaishi and P. de Forcrand. "Testing and tuning new symplectic integrators for hybrid Monte Carlo algorithm in lattice QCD." In: *Phys. Rev. E* 73 (2006), p. 036706. DOI: 10.1103/PhysRevE.73.036706. arXiv: hep-lat/0505020 [hep-lat].

- [98] I. Campos, P. Fritzsche, M. Hansen, M. K. Marinkovic, A. Patella, A. Ramos, and N. Tantalo. "openQCD code: a versatile tool for QCD+QED simulations." In: *Eur. Phys. J. C* 80.3 (2020), p. 195. DOI: 10.1140/epjc/s10052-020-7617-3. arXiv: 1908.11673 [hep-lat].
- [99] D. H. Weingarten and D. N. Petcher. "Monte Carlo Integration for Lattice Gauge Theories with Fermions." In: *Phys. Lett.* 99B (1981), pp. 333–338. DOI: 10.1016/0370-2693(81)90112-X.
- [100] S. Cali, F. Knechtli, and T. Korzec. "How much do charm sea quarks affect the charmonium spectrum?" In: *Eur. Phys. J. C* 79.7 (2019), p. 607. DOI: 10.1140/epjc/s10052-019-7108-6. arXiv: 1905.12971 [hep-lat].
- [101] L. Del Debbio, L. Giusti, M. Lüscher, R. Petronzio, and N. Tantalo. "Stability of lattice QCD simulations and the thermodynamic limit." In: *JHEP* 02 (2006), p. 011. DOI: 10.1088/1126-6708/2006/02/011. arXiv: hep-lat/0512021 [hep-lat].
- [102] D. Mohler and S. Schaefer. *Remarks on strange-quark simulations with Wilson fermions*. Mar. 2020. arXiv: 2003.13359 [hep-lat].
- [103] P. de Forcrand and T. Takaishi. "Fast fermion Monte Carlo." In: *Nucl. Phys. Proc. Suppl.* 53 (1997), pp. 968–970. DOI: 10.1016/S0920-5632(96)00829-8. arXiv: hep-lat/9608093 [hep-lat].
- [104] R. Frezzotti and K. Jansen. "A Polynomial hybrid Monte Carlo algorithm." In: *Phys. Lett.* B402 (1997), pp. 328–334. DOI: 10.1016/S0370-2693(97)00475-9. arXiv: hep-lat/9702016 [hep-lat].
- [105] T. Takaishi and P. de Forcrand. "Odd flavor hybrid Monte Carlo algorithm for lattice QCD." In: *Int. J. Mod. Phys. C* 13 (2002), pp. 343–366. DOI: 10.1142/S0129183102003152. arXiv: hep-lat/0108012 [hep-lat].
- [106] A. D. Kennedy, I. Horvath, and S. Sint. "A New exact method for dynamical fermion computations with nonlocal actions." In: *Nucl. Phys. Proc. Suppl.* 73 (1999), pp. 834–836. DOI: 10.1016/S0920-5632(99)85217-7. arXiv: hep-lat/9809092 [hep-lat].
- [107] M. A. Clark and A. D. Kennedy. "Accelerating dynamical fermion computations using the rational hybrid Monte Carlo (RHMC) algorithm with multiple pseudofermion fields." In: *Phys. Rev. Lett.* 98 (2007), p. 051601. DOI: 10.1103/PhysRevLett.98.051601. arXiv: hep-lat/0608015 [hep-lat].
- [108] M. Lüscher. *Charm and strange quark in openQCD simulations*. <http://luscher.web.cern.ch/luscher/openQCD/>. 2012.
- [109] M. Lüscher and F. Palombi. "Fluctuations and reweighting of the quark determinant on large lattices." In: *PoS LATTICE2008* (2008), p. 049. DOI: 10.22323/1.066.0049. arXiv: 0810.0946 [hep-lat].

- [110] J. C. Sexton and D. H. Weingarten. "Hamiltonian evolution for the hybrid Monte Carlo algorithm." In: *Nucl. Phys.* B380 (1992), pp. 665–677. DOI: 10.1016/0550-3213(92)90263-B.
- [111] M. J. Peardon and J. Sexton. "Multiple molecular dynamics time scales in hybrid Monte Carlo fermion simulations." In: *Nucl. Phys. Proc. Suppl.* 119 (2003), pp. 985–987. DOI: 10.1016/S0920-5632(03)01738-9. arXiv: hep-lat/0209037 [hep-lat].
- [112] M. Hasenbusch. "Speeding up the hybrid Monte Carlo algorithm for dynamical fermions." In: *Phys. Lett.* B519 (2001), pp. 177–182. DOI: 10.1016/S0370-2693(01)01102-9. arXiv: hep-lat/0107019 [hep-lat].
- [113] M. Hasenbusch and K. Jansen. "Speeding up lattice QCD simulations with clover improved Wilson fermions." In: *Nucl. Phys.* B659 (2003), pp. 299–320. DOI: 10.1016/S0550-3213(03)00227-X. arXiv: hep-lat/0211042 [hep-lat].
- [114] C. Urbach, K. Jansen, A. Shindler, and U. Wenger. "HMC algorithm with multiple time scale integration and mass preconditioning." In: *Comput. Phys. Commun.* 174 (2006), pp. 87–98. DOI: 10.1016/j.cpc.2005.08.006. arXiv: hep-lat/0506011 [hep-lat].
- [115] M. A. Clark, B. Joó, A. D. Kennedy, and P. J. Silva. "Improving dynamical lattice QCD simulations through integrator tuning using Poisson brackets and a force-gradient integrator." In: *Phys. Rev.* D84 (2011), p. 071502. DOI: 10.1103/PhysRevD.84.071502. arXiv: 1108.1828 [hep-lat].
- [116] A. D. Kennedy, P. J. Silva, and M. A. Clark. "Shadow Hamiltonians, Poisson Brackets, and Gauge Theories." In: *Phys. Rev.* D87.3 (2013), p. 034511. DOI: 10.1103/PhysRevD.87.034511. arXiv: 1210.6600 [hep-lat].
- [117] A. Bussone, M. Della Morte, V. Drach, M. Hansen, A. Hietanen, J. Rantaharju, and C. Pica. "A simple method to optimize HMC performance." In: *PoS LATTICE2016* (2016), p. 260. DOI: 10.22323/1.256.0260. arXiv: 1610.02860 [hep-lat].
- [118] A. Bussone, M. Della Morte, V. Drach, and C. Pica. "Tuning the hybrid Monte Carlo algorithm using molecular dynamics forces' variances." In: *Comput. Phys. Commun.* 234 (2019), pp. 179–187. DOI: 10.1016/j.cpc.2018.07.012. arXiv: 1801.06412 [hep-lat].
- [119] M. Bruno, P. Korcyl, T. Korzec, S. Lottini, and S. Schaefer. "On the extraction of spectral quantities with open boundary conditions." In: *PoS LATTICE2014* (2014), p. 089. DOI: 10.22323/1.214.0089. arXiv: 1411.5207 [hep-lat].
- [120] M. Lüscher. *Implementation of the lattice Dirac operator*. <http://luscher.web.cern.ch/luscher/openQCD/>. 2016.

- [121] R. C. Brower, T. Ivanenko, A. R. Levi, and K. N. Orginos. "Chronological inversion method for the Dirac matrix in hybrid Monte Carlo." In: *Nucl. Phys.* B484 (1997), pp. 353–374. DOI: 10.1016/S0550-3213(96)00579-2. arXiv: hep-lat/9509012 [hep-lat].
- [122] M. Hestenes and E. Stiefel. "Methods of conjugate gradients for solving linear systems." In: *Journal of Research of the National Bureau of Standards* 49.6 (1952), p. 409. ISSN: 0091-0635. DOI: 10.6028/jres.049.044.
- [123] Y. Saad. *Iterative Methods for Sparse Linear Systems*. 2nd. USA: Society for Industrial and Applied Mathematics, 2003. ISBN: 0898715342.
- [124] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. 3rd ed. USA: Cambridge University Press, 2007. ISBN: 0521880688.
- [125] U. Glassner, S. Güsken, T. Lippert, G. Ritzenhofer, K. Schilling, and A. Frommer. "How to compute Green's functions for entire mass trajectories within Krylov solvers." In: *Int. J. Mod. Phys. C7* (1996), p. 635. DOI: 10.1142/S0129183196000533. arXiv: hep-lat/9605008 [hep-lat].
- [126] B. Jegerlehner. *Krylov space solvers for shifted linear systems*. 1996. arXiv: hep-lat/9612014 [hep-lat].
- [127] M. Lüscher. *Multi-shift conjugate gradient algorithm*. <http://luscher.web.cern.ch/luscher/openQCD/>. 2012.
- [128] M. Lüscher. "Lattice QCD and the Schwarz alternating procedure." In: *JHEP* 05 (2003), p. 052. DOI: 10.1088/1126-6708/2003/05/052. arXiv: hep-lat/0304007 [hep-lat].
- [129] M. Lüscher. "Solution of the Dirac equation in lattice QCD using a domain decomposition method." In: *Comput. Phys. Commun.* 156 (2004), pp. 209–220. DOI: 10.1016/S0010-4655(03)00486-7. arXiv: hep-lat/0310048 [hep-lat].
- [130] M. Lüscher. "Schwarz-preconditioned HMC algorithm for two-flavour lattice QCD." In: *Comput. Phys. Commun.* 165 (2005), pp. 199–220. DOI: 10.1016/j.cpc.2004.10.004. arXiv: hep-lat/0409106 [hep-lat].
- [131] T. Banks and A. Casher. "Chiral Symmetry Breaking in Confining Theories." In: *Nucl. Phys.* B169 (1980), pp. 103–125. DOI: 10.1016/0550-3213(80)90255-2.
- [132] A. Jüttner and M. Della Morte. "Heavy quark propagators with improved precision using domain decomposition." In: *PoS LAT2005* (2006), p. 204. DOI: 10.22323/1.020.0204. arXiv: hep-lat/0508023 [hep-lat].

- [133] G. M. de Divitiis, R. Petronzio, and N. Tantalo. "Distance preconditioning for lattice Dirac operators." In: *Phys. Lett.* B692 (2010), pp. 157–160. DOI: 10.1016/j.physletb.2010.07.031. arXiv: 1006.4028 [hep-lat].
- [134] S. Collins, K. Eckert, J. Heitger, S. Hofmann, and W. Soeldner. "Charmed pseudoscalar decay constants on three-flavour CLS ensembles with open boundaries." In: *PoS LATTICE2016* (2017), p. 368. DOI: 10.22323/1.256.0368. arXiv: 1701.05502 [hep-lat].
- [135] U. Wolff. "Monte Carlo errors with less errors." In: *Comput. Phys. Commun.* 156 (2004). [Erratum: *Comput. Phys. Commun.* 176, 383 (2007)], pp. 143–153. DOI: 10.1016/S0010-4655(03)00467-3. arXiv: hep-lat/0306017.
- [136] A. Sokal. "Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms." In: *Functional Integration: Basics and Applications*. Ed. by C. DeWitt-Morette, P. Cartier, and A. Folacci. Boston, MA: Springer US, 1997, pp. 131–192. ISBN: 978-1-4899-0319-8. DOI: 10.1007/978-1-4899-0319-8-6.
- [137] N. Madras and A. D. Sokal. "The Pivot algorithm: a highly efficient Monte Carlo method for selfavoiding walk." In: *J. Statist. Phys.* 50 (1988), pp. 109–186. DOI: 10.1007/BF01022990.
- [138] A. Ramos. "Automatic differentiation for error analysis of Monte Carlo data." In: *Comput. Phys. Commun.* 238 (2019), pp. 19–35. DOI: 10.1016/j.cpc.2018.12.020. arXiv: 1809.01289 [hep-lat].
- [139] W. Janke. "Statistical Analysis of Simulations: Data Correlations and Error Estimation." In: *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*. Ed. by J Groten-dorst, D Marx, and A Muramatsu. Vol. 10. NIC Series. John von Neumann Institute for Computing, Jülich. 2002, pp. 423–445. URL: <http://www.fz-juelich.de/nic-series/volume10>.
- [140] P. Young. *Everything you wanted to know about Data Analysis and Fitting but were afraid to ask*. 2012. arXiv: 1210.3781 [physics.data-an].
- [141] G. Bali. *Computation of hadronic two-point functions in Lattice QCD*. 2012. URL: <https://indico.cern.ch/event/745783/contributions/3082095/attachments/1690338/2720181/corrected.pdf>.
- [142] M. Bruno, T. Korzec, and S. Schaefer. "Setting the scale for the CLS 2 + 1 flavor ensembles." In: *Phys. Rev. D*95.7 (2017), p. 074504. DOI: 10.1103/PhysRevD.95.074504. arXiv: 1608.08900 [hep-lat].

- [143] M. Guagnelli, J. Heitger, R. Sommer, and H. Wittig. "Hadron masses and matrix elements from the QCD Schrödinger functional." In: *Nucl. Phys.* B560 (1999), pp. 465–481. DOI: 10.1016/S0550-3213(99)00466-6. arXiv: hep-lat/9903040 [hep-lat].
- [144] A. Vladikas. "Three Topics in Renormalization and Improvement." In: *Modern perspectives in lattice QCD: Quantum field theory and high performance computing. Proceedings, International School, 93rd Session, Les Houches, France, August 3-28, 2009.* 2011, pp. 161–222. arXiv: 1103.1323 [hep-lat].
- [145] C. Michael. "Adjoint Sources in Lattice Gauge Theory." In: *Nucl. Phys.* B259 (1985), pp. 58–76. DOI: 10.1016/0550-3213(85)90297-4.
- [146] M. Lüscher and U. Wolff. "How to Calculate the Elastic Scattering Matrix in Two-dimensional Quantum Field Theories by Numerical Simulation." In: *Nucl. Phys.* B339 (1990), pp. 222–252. DOI: 10.1016/0550-3213(90)90540-T.
- [147] B. Blossier, M. Della Morte, G. von Hippel, T. Mendes, and R. Sommer. "On the generalized eigenvalue method for energies and matrix elements in lattice field theory." In: *JHEP* 04 (2009), p. 094. DOI: 10.1088/1126-6708/2009/04/094. arXiv: 0902.1265 [hep-lat].
- [148] B. Blossier, M. Della Morte, N. Garron, G. von Hippel, T. Mendes, H. Simma, and R. Sommer. "HQET at order $1/m$: II. Spectroscopy in the quenched approximation." In: *JHEP* 05 (2010), p. 074. DOI: 10.1007/JHEP05(2010)074. arXiv: 1004.2661 [hep-lat].
- [149] S. Lottini and R. Sommer. "Data Analysis in Lattice Field Theory." In: *Lattice Practices 2014.* 2014.
- [150] C. Michael. "Fitting correlated data." In: *Phys. Rev.* D49 (1994), pp. 2616–2619. DOI: 10.1103/PhysRevD.49.2616. arXiv: hep-lat/9310026 [hep-lat].
- [151] D. Seibert. "Undesirable effects of covariance matrix techniques for error analysis." In: *Phys. Rev.* D49 (1994), pp. 6240–6243. DOI: 10.1103/PhysRevD.49.6240. arXiv: hep-lat/9305014 [hep-lat].
- [152] B. Yoon, Y.-C. Jang, C. Jung, and W. Lee. "Covariance fitting of highly correlated data in lattice QCD." In: *J. Korean Phys. Soc.* 63 (2013), pp. 145–162. DOI: 10.3938/jkps.63.145. arXiv: 1101.2248 [hep-lat].
- [153] M. Ammer and S. Dürr. "Details of a staggered fermion data analysis." In: *37th International Symposium on Lattice Field Theory (Lattice 2019) Wuhan, Hubei, China, June 16-22, 2019.* 2019. arXiv: 1910.11046 [hep-lat].

- [154] M. Bruno and R. Sommer. to appear.
- [155] M. Lüscher. "Trivializing maps, the Wilson flow and the HMC algorithm." In: *Commun. Math. Phys.* 293 (2010), pp. 899–919. DOI: 10.1007/s00220-009-0953-7. arXiv: 0907.5491 [hep-lat].
- [156] M. Lüscher. "Properties and uses of the Wilson flow in lattice QCD." In: *JHEP* 08 (2010). [Erratum: *JHEP* 03, 092 (2014)], p. 071. DOI: 10.1007/JHEP08(2010)071. arXiv: 1006.4518 [hep-lat].
- [157] M. Lüscher. "Chiral symmetry and the Yang–Mills gradient flow." In: *JHEP* 04 (2013), p. 123. DOI: 10.1007/JHEP04(2013)123. arXiv: 1302.5246 [hep-lat].
- [158] P. Fritzsche and A. Ramos. "The gradient flow coupling in the Schrödinger Functional." In: *JHEP* 10 (2013), p. 008. DOI: 10.1007/JHEP10(2013)008. arXiv: 1301.4388 [hep-lat].
- [159] M. Lüscher and P. Weisz. "Perturbative analysis of the gradient flow in non-abelian gauge theories." In: *JHEP* 02 (2011), p. 051. DOI: 10.1007/JHEP02(2011)051. arXiv: 1101.0963 [hep-th].
- [160] R. Narayanan and H. Neuberger. "Infinite N phase transitions in continuum Wilson loop operators." In: *JHEP* 03 (2006), p. 064. DOI: 10.1088/1126-6708/2006/03/064. arXiv: hep-th/0601210.
- [161] R. Lohmayer and H. Neuberger. "Continuous smearing of Wilson Loops." In: *PoS LATTICE2011* (2011), p. 249. DOI: 10.22323/1.139.0249. arXiv: 1110.3522 [hep-lat].
- [162] C. Runge. "Ueber die numerische Auflösung von Differentialgleichungen." In: *Mathematische Annalen* 46.2 (1895), pp. 167–178. ISSN: 00255831. DOI: 10.1007/BF01446807.
- [163] W. Kutta. "Beitrag zur näherungsweise Integration totaler Differentialgleichungen." In: *Z. Math. Phys.* 46 (1901), pp. 435–453.
- [164] E. Fehlberg. "Klassische Runge-Kutta-Formeln vierter und niedrigerer Ordnung mit Schrittweiten-Kontrolle und ihre Anwendung auf Wärmeleitungsprobleme." In: *Computing* 6.1-2 (1970), pp. 61–71. ISSN: 0010485X. DOI: 10.1007/BF02241732.
- [165] S. Borsanyi et al. "High-precision scale setting in lattice QCD." In: *JHEP* 09 (2012), p. 010. DOI: 10.1007/JHEP09(2012)010. arXiv: 1203.4469 [hep-lat].
- [166] Z. Fodor, K. Holland, J. Kuti, S. Mondal, D. Nogradi, and C. H. Wong. "The lattice gradient flow at tree-level and its improvement." In: *JHEP* 09 (2014), p. 018. DOI: 10.1007/JHEP09(2014)018. arXiv: 1406.0827 [hep-lat].

- [167] A. Cheng, A. Hasenfratz, Y. Liu, G. Petropoulos, and D. Schaich. "Improving the continuum limit of gradient flow step scaling." In: *JHEP* 05 (2014), p. 137. DOI: 10.1007/JHEP05(2014)137. arXiv: 1404.0984 [hep-lat].
- [168] A. Ramos and S. Sint. "Symanzik improvement of the gradient flow in lattice gauge theories." In: *Eur. Phys. J. C* 76.1 (2016), p. 15. DOI: 10.1140/epjc/s10052-015-3831-9. arXiv: 1508.05552 [hep-lat].
- [169] R. Sommer. "Scale setting in lattice QCD." In: *PoS LATTICE2013* (2014), p. 015. DOI: 10.22323/1.187.0015. arXiv: 1401.3270 [hep-lat].
- [170] Z. Fodor, K. Holland, J. Kuti, D. Nogradi, and C. H. Wong. "The Yang-Mills gradient flow in finite volume." In: *JHEP* 11 (2012), p. 007. DOI: 10.1007/JHEP11(2012)007. arXiv: 1208.1051 [hep-lat].
- [171] P. Fritzsche. private communication.
- [172] P. Fritzsche, A. Ramos, and F. Stollenwerk. "Critical slowing down and the gradient flow coupling in the Schrödinger functional." In: *PoS Lattice2013* (2014), p. 461. DOI: 10.22323/1.187.0461. arXiv: 1311.7304 [hep-lat].
- [173] A. A. Belavin, A. M. Polyakov, A. S. Schwartz, and Yu. S. Tyupkin. "Pseudoparticle Solutions of the Yang-Mills Equations." In: *Phys. Lett.* 59B (1975), pp. 85–87. DOI: 10.1016/0370-2693(75)90163-X.
- [174] M. Atiyah and I. Singer. "The Index of elliptic operators. 5." In: *Annals Math.* 93 (1971), pp. 139–149. DOI: 10.2307/1970757.
- [175] M. Lüscher. "Topology of Lattice Gauge Fields." In: *Commun. Math. Phys.* 85 (1982), p. 39. DOI: 10.1007/BF02029132.
- [176] A. Phillips and D. Stone. "Lattice Gauge Fields, Principal Bundles and the Calculation of Topological Charge." In: *Commun. Math. Phys.* 103 (1986), pp. 599–636. DOI: 10.1007/BF01211167.
- [177] M. Lüscher. "Topological effects in QCD and the problem of short distance singularities." In: *Phys. Lett. B* 593 (2004), pp. 296–301. DOI: 10.1016/j.physletb.2004.04.076. arXiv: hep-th/0404034.
- [178] M. Bruno, S. Schaefer, and R. Sommer. "Topological susceptibility and the sampling of field space in $N_f = 2$ lattice QCD simulations." In: *JHEP* 08 (2014), p. 150. DOI: 10.1007/JHEP08(2014)150. arXiv: 1406.5363 [hep-lat].
- [179] M. Lüscher. "Step scaling and the Yang-Mills gradient flow." In: *JHEP* 06 (2014), p. 105. DOI: 10.1007/JHEP06(2014)105. arXiv: 1404.5930 [hep-lat].

- [180] M. Albanese et al. "Glueball Masses and String Tension in Lattice QCD." In: *Phys. Lett. B* 192 (1987), pp. 163–169. DOI: 10.1016/0370-2693(87)91160-9.
- [181] C. W. Bernard and T. A. DeGrand. "Perturbation theory for fat link fermion actions." In: *Nucl. Phys. B Proc. Suppl.* 83 (2000), pp. 845–847. DOI: 10.1016/S0920-5632(00)91822-X. arXiv: hep-lat/9909083.
- [182] M. Della Morte, A. Shindler, and R. Sommer. "On lattice actions for static quarks." In: *JHEP* 08 (2005), p. 051. DOI: 10.1088/1126-6708/2005/08/051. arXiv: hep-lat/0506008.
- [183] A. Hasenfratz and F. Knechtli. "Flavor symmetry and the static potential with hypercubic blocking." In: *Phys. Rev. D* 64 (2001), p. 034504. DOI: 10.1103/PhysRevD.64.034504. arXiv: hep-lat/0103029.
- [184] A. Hasenfratz, R. Hoffmann, and F. Knechtli. "The Static potential with hypercubic blocking." In: *Nucl. Phys. B Proc. Suppl.* 106 (2002), pp. 418–420. DOI: 10.1016/S0920-5632(01)01733-9. arXiv: hep-lat/0110168.
- [185] M. Della Morte, S. Dürr, J. Heitger, H. Molke, J. Rolf, A. Shindler, and R. Sommer. "Lattice HQET with exponentially improved statistical precision." In: *Phys. Lett. B* 581 (2004). [Erratum: *Phys.Lett.B* 612, 313–314 (2005)], pp. 93–98. DOI: 10.1016/j.physletb.2005.03.017. arXiv: hep-lat/0307021.
- [186] S. Güsken, U. Löw, K. H. Mütter, R. Sommer, A. Patel, and K. Schilling. "Nonsinglet Axial Vector Couplings of the Baryon Octet in Lattice QCD." In: *Phys. Lett. B* 227 (1989), pp. 266–269. DOI: 10.1016/S0370-2693(89)80034-6.
- [187] S. Basak, I. Sato, S. Wallace, R. Edwards, D. Richards, G. T. Fleming, U. M. Heller, A. C. Lichtl, and C. Morningstar. "Combining quark and link smearing to improve extended baryon operators." In: *PoS LAT2005* (2006). Ed. by C. Michael, p. 076. DOI: 10.22323/1.020.0076. arXiv: hep-lat/0509179.
- [188] M. Papinutto, F. Scardino, and S. Schaefer. "New extended interpolating fields built from three-dimensional fermions." In: *Phys. Rev. D* 98.9 (2018), p. 094506. DOI: 10.1103/PhysRevD.98.094506. arXiv: 1807.08714 [hep-lat].
- [189] E. Fermi. "Versuch einer Theorie der β -Strahlen. I." In: *Zeitschrift für Physik* 88.3-4 (1934), pp. 161–177. ISSN: 14346001. DOI: 10.1007/BF01351864.
- [190] S. Weinberg. "Phenomenological Lagrangians." In: *Physica* A96.1-2 (1979), pp. 327–340. DOI: 10.1016/0378-4371(79)90223-1.

- [191] M. Neubert. "Heavy quark symmetry." In: *Phys. Rept.* 245 (1994), pp. 259–396. DOI: 10.1016/0370-1573(94)90091-4. arXiv: hep-ph/9306320 [hep-ph].
- [192] T. Mannel. "Effective Field Theories for Heavy Quarks: Heavy Quark Effective Theory, Heavy Quark Expansion, Non-Relativistic QCD." In: *Les Houches summer school: EFT in Particle Physics and Cosmology*. <https://indico.in2p3.fr/event/13465/sessions/2550/attachments/12152/14956/MannelLectureNotes.pdf>. 2017.
- [193] M. B. Wise. "Heavy quark physics: Course." In: *Probing the standard model of particle interactions. Proceedings, Summer School in Theoretical Physics, NATO Advanced Study Institute, 68th session, Les Houches, France, July 28-September 5, 1997. Pt. 1, 2.* 1997, pp. 1051–1089. arXiv: hep-ph/9805468 [hep-ph].
- [194] M. Neubert. "Heavy quark effective theory." In: *Subnucl. Ser.* 34 (1997), pp. 98–165. arXiv: hep-ph/9610266 [hep-ph].
- [195] E. Eichten and B. R. Hill. "Static effective field theory: $1/m$ corrections." In: *Phys. Lett. B* 243 (1990), pp. 427–431. DOI: 10.1016/0370-2693(90)91408-4.
- [196] H. Georgi. "An Effective Field Theory for Heavy Quarks at Low-energies." In: *Phys. Lett. B* 240 (1990), pp. 447–450. DOI: 10.1016/0370-2693(90)91128-X.
- [197] N. Isgur and M. B. Wise. "Weak transition form-factors between heavy mesons." In: *Phys. Lett. B* 237 (1990), pp. 527–530. DOI: 10.1016/0370-2693(90)91219-2.
- [198] A. F. Falk and M. Neubert. "Second order power corrections in the heavy quark effective theory. 1. Formalism and meson form-factors." In: *Phys. Rev. D* 47 (1993), pp. 2965–2981. DOI: 10.1103/PhysRevD.47.2965. arXiv: hep-ph/9209268 [hep-ph].
- [199] A. F. Falk, M. Neubert, and M. E. Luke. "The Residual mass term in the heavy quark effective theory." In: *Nucl. Phys. B* 388 (1992), pp. 363–375. DOI: 10.1016/0550-3213(92)90617-K. arXiv: hep-ph/9204229 [hep-ph].
- [200] A. F. Falk and M. Neubert. "Second order power corrections in the heavy quark effective theory. 2. Baryon form-factors." In: *Phys. Rev. D* 47 (1993), pp. 2982–2990. DOI: 10.1103/PhysRevD.47.2982. arXiv: hep-ph/9209269.
- [201] R. Sommer. "Introduction to Non-perturbative Heavy Quark Effective Theory." In: *Les Houches Summer School: Session 93: Modern perspectives in lattice QCD: Quantum field theory and high performance computing*. Aug. 2010, pp. 517–590. arXiv: 1008.0710 [hep-lat].

- [202] R. Sommer. “Non-perturbative Heavy Quark Effective Theory: Introduction and Status.” In: *Nucl. Part. Phys. Proc.* 261-262 (2015). Ed. by J. Blümlein, K. Jansen, M. Krämer, and J. H. Kühn, pp. 338–367. DOI: 10.1016/j.nuclphysbps.2015.03.022. arXiv: 1501.03060 [hep-lat].
- [203] E. Eichten. “Heavy Quarks on the Lattice.” In: *Nucl. Phys. B Proc. Suppl.* 4 (1988). Ed. by A. Billoire, R. Lacaze, A. Morel, O. Napoloy, and J. Zinn-Justin, p. 170. DOI: 10.1016/0920-5632(88)90097-7.
- [204] E. Eichten and B. R. Hill. “An Effective Field Theory for the Calculation of Matrix Elements Involving Heavy Quarks.” In: *Phys. Lett. B* 234 (1990), pp. 511–516. DOI: 10.1016/0370-2693(90)92049-0.
- [205] M. Kurth and R. Sommer. “Renormalization and $O(a)$ improvement of the static axial current.” In: *Nucl. Phys. B* 597 (2001), pp. 488–518. DOI: 10.1016/S0550-3213(00)00750-1. arXiv: hep-lat/0007002.
- [206] J. Heitger and R. Sommer. “Nonperturbative heavy quark effective theory.” In: *JHEP* 02 (2004), p. 022. DOI: 10.1088/1126-6708/2004/02/022. arXiv: hep-lat/0310035.
- [207] B. Blossier, M. Della Morte, N. Garron, and R. Sommer. “HQET at order $1/m$: I. Non-perturbative parameters in the quenched approximation.” In: *JHEP* 06 (2010), p. 002. DOI: 10.1007/JHEP06(2010)002. arXiv: 1001.4783 [hep-lat].
- [208] M. Della Morte, S. Dooling, J. Heitger, D. Hesse, and H. Simma. “Matching of heavy-light flavour currents between HQET at order $1/m$ and QCD: I. Strategy and tree-level study.” In: *JHEP* 05 (2014), p. 060. DOI: 10.1007/JHEP05(2014)060. arXiv: 1312.1566 [hep-lat].
- [209] S. Aoki et al. “FLAG Review 2019: Flavour Lattice Averaging Group (FLAG).” In: *Eur. Phys. J. C* 80.2 (2020), p. 113. DOI: 10.1140/epjc/s10052-019-7354-7. arXiv: 1902.08191 [hep-lat].
- [210] C. McNeile, C. Davies, E. Follana, K. Hornbostel, and G. Lepage. “High-Precision c and b Masses, and QCD Coupling from Current-Current Correlators in Lattice and Continuum QCD.” In: *Phys. Rev. D* 82 (2010), p. 034512. DOI: 10.1103/PhysRevD.82.034512. arXiv: 1004.4285 [hep-lat].
- [211] F. Bernardoni et al. “Decay constants of B-mesons from non-perturbative HQET with two light dynamical quarks.” In: *Phys. Lett. B* 735 (2014), pp. 349–356. DOI: 10.1016/j.physletb.2014.06.051. arXiv: 1404.3590 [hep-lat].

- [212] M. Dalla Brida, P. Fritzsche, T. Korzec, A. Ramos, S. Sint, and R. Sommer. "Determination of the QCD Λ -parameter and the accuracy of perturbation theory at high energies." In: *Phys. Rev. Lett.* 117.18 (2016), p. 182001. DOI: 10.1103/PhysRevLett.117.182001. arXiv: 1604.06193 [hep-ph].
- [213] M. Lüscher, P. Weisz, and U. Wolff. "A Numerical method to compute the running coupling in asymptotically free theories." In: *Nucl. Phys. B* 359 (1991), pp. 221–243. DOI: 10.1016/0550-3213(91)90298-C.
- [214] M. Bruno, M. Dalla Brida, P. Fritzsche, T. Korzec, A. Ramos, S. Schaefer, H. Simma, S. Sint, and R. Sommer. "QCD Coupling from a Nonperturbative Determination of the Three-Flavor Λ Parameter." In: *Phys. Rev. Lett.* 119.10 (2017), p. 102001. DOI: 10.1103/PhysRevLett.119.102001. arXiv: 1706.03821 [hep-lat].
- [215] B. Blossier, M. Della Morte, P. Fritzsche, N. Garron, J. Heitger, H. Simma, R. Sommer, and N. Tantalo. "Parameters of Heavy Quark Effective Theory from $N_f = 2$ lattice QCD." In: *JHEP* 09 (2012), p. 132. DOI: 10.1007/JHEP09(2012)132. arXiv: 1203.6516 [hep-lat].
- [216] Fritzsche, Patrick and Korzec, Tomasz. *Simulating the QCD Schrödinger Functional with three massless quark flavors*. in preparation.
- [217] Sommer, Rainer and others. *Internal notes of the ALPHA collaboration*.
- [218] S. Capitani, M. Lüscher, R. Sommer, and H. Wittig. "Non-perturbative quark mass renormalization in quenched lattice QCD." In: *Nucl. Phys. B* 544 (1999). [Erratum: *Nucl.Phys.B* 582, 762–762 (2000)], pp. 669–698. DOI: 10.1016/S0550-3213(98)00857-8. arXiv: hep-lat/9810063.
- [219] M. Dalla Brida, R. Höllwieser, F. Knechtli, T. Korzec, A. Ramos, and R. Sommer. "Non-perturbative renormalization by decoupling." In: (Dec. 2019). DOI: 10.1016/j.physletb.2020.135571. arXiv: 1912.06001 [hep-lat].
- [220] R. Höllwieser, F. Knechtli, and T. Korzec. "Scale setting for $N_f = 3 + 1$ QCD." In: *Eur. Phys. J. C* 80.4 (2020), p. 349. DOI: 10.1140/epjc/s10052-020-7889-7. arXiv: 2002.02866 [hep-lat].
- [221] M. Lüscher. *Parameters of the openQCD main programs*. <http://luscher.web.cern.ch/luscher/openQCD/>. 2016.
- [222] A. Hasenfratz and O. Witzel. *Dislocations under gradient flow and their effect on the renormalized coupling*. Apr. 2020. arXiv: 2004.00758 [hep-lat].
- [223] G. de Divitiis and R. Petronzio. "Nonperturbative renormalization constants on the lattice from flavor nonsinglet Ward identities." In: *Phys. Lett. B* 419 (1998), pp. 311–316. DOI: 10.1016/S0370-2693(97)01444-5. arXiv: hep-lat/9710071.

- [224] T. Bhattacharya, S. Chandrasekharan, R. Gupta, W.-J. Lee, and S. R. Sharpe. “Nonperturbative renormalization constants using Ward identities.” In: *Phys. Lett. B* 461 (1999), pp. 79–88. DOI: 10.1016/S0370-2693(99)00796-0. arXiv: hep-lat/9904011.
- [225] T. Bhattacharya, R. Gupta, W.-J. Lee, and S. R. Sharpe. “Order α improved renormalization constants.” In: *Phys. Rev. D* 63 (2001), p. 074505. DOI: 10.1103/PhysRevD.63.074505. arXiv: hep-lat/0009038.
- [226] T. Bhattacharya, R. Gupta, W. Lee, and S. R. Sharpe. “Scaling behavior of discretization errors in renormalization and improvement constants.” In: *Phys. Rev. D* 73 (2006), p. 114507. DOI: 10.1103/PhysRevD.73.114507. arXiv: hep-lat/0509160.
- [227] M. Guagnelli, R. Petronzio, J. Rolf, S. Sint, R. Sommer, and U. Wolff. “Nonperturbative results for the coefficients b_m and $b_A - b_P$ in $O(\alpha)$ improved lattice QCD.” In: *Nucl. Phys. B* 595 (2001), pp. 44–62. DOI: 10.1016/S0550-3213(00)00675-1. arXiv: hep-lat/0009021.
- [228] J. Heitger and J. Wennekers. “Effective heavy light meson energies in small volume quenched QCD.” In: *JHEP* 02 (2004), p. 064. DOI: 10.1088/1126-6708/2004/02/064. arXiv: hep-lat/0312016.
- [229] P. Fritzsche, J. Heitger, and N. Tantalo. “Non-perturbative improvement of quark mass renormalization in two-flavour lattice QCD.” In: *JHEP* 08 (2010), p. 074. DOI: 10.1007/JHEP08(2010)074. arXiv: 1004.3978 [hep-lat].
- [230] G. M. de Divitiis, M. Firrotta, J. Heitger, C. C. Köster, and A. Vladikas. “Non-perturbative determination of improvement b-coefficients in $N_f = 3$.” In: *EPJ Web Conf.* 175 (2018). Ed. by M. Della Morte, P. Fritzsche, E. Gámiz Sánchez, and C. Pena Ruano, p. 10008. DOI: 10.1051/epjconf/201817510008. arXiv: 1710.07020 [hep-lat].
- [231] C. Wittmeier. “Non-perturbative Renormalization and Improvement of the Isovector Axial Current in Three-Flavor Lattice QCD with Wilson Fermions.” PhD thesis. University of Münster, 2018.
- [232] Heitger, Jochen and Joswig, Fabian. *The renormalized $O(\alpha)$ improved vector current in three-flavour lattice QCD with Wilson quarks.* in preparation. 2020.
- [233] S. Aoki, K.-i. Nagai, Y. Taniguchi, and A. Ukawa. “Perturbative renormalization factors of bilinear quark operators for improved gluon and quark actions in lattice QCD.” In: *Phys. Rev. D* 58 (1998), p. 074505. DOI: 10.1103/PhysRevD.58.074505. arXiv: hep-lat/9802034.

- [234] Y. Taniguchi and A. Ukawa. "Perturbative calculation of improvement coefficients to $O(g^2 a)$ for bilinear quark operators in lattice QCD." In: *Phys. Rev. D* 58 (1998), p. 114503. DOI: 10.1103/PhysRevD.58.114503. arXiv: hep-lat/9806015.
- [235] G. S. Bali, E. E. Scholz, J. Simeth, and W. Söldner. "Lattice simulations with $N_f = 2 + 1$ improved Wilson fermions at a fixed strange quark mass." In: *Phys. Rev. D* 94.7 (2016), p. 074501. DOI: 10.1103/PhysRevD.94.074501. arXiv: 1606.09039 [hep-lat].
- [236] D. Mohler, S. Schaefer, and J. Simeth. "CLS 2+1 flavor simulations at physical light- and strange-quark masses." In: *EPJ Web Conf.* 175 (2018), p. 02010. DOI: 10.1051/epjconf/201817502010. arXiv: 1712.04884 [hep-lat].
- [237] J. Heitger, F. Joswig, and A. Vladikas. *Ward identity determination of Z_S/Z_P for $N_f = 3$ lattice QCD in a Schrödinger functional setup*. May 2020. arXiv: 2005.01352 [hep-lat].
- [238] M. Della Morte, R. Hoffmann, F. Knechtli, R. Sommer, and U. Wolff. "Non-perturbative renormalization of the axial current with dynamical Wilson fermions." In: *JHEP* 07 (2005), p. 007. DOI: 10.1088/1126-6708/2005/07/007. arXiv: hep-lat/0505026.
- [239] F. Bernardoni et al. "The b-quark mass from non-perturbative $N_f = 2$ Heavy Quark Effective Theory at $O(1/m_h)$." In: *Phys. Lett. B* 730 (2014), pp. 171–177. DOI: 10.1016/j.physletb.2014.01.046. arXiv: 1311.5498 [hep-lat].
- [240] J. Gasser and H. Leutwyler. "Light Quarks at Low Temperatures." In: *Phys. Lett. B* 184 (1987), pp. 83–88. DOI: 10.1016/0370-2693(87)90492-8.
- [241] J. Gasser and H. Leutwyler. "Spontaneously Broken Symmetries: Effective Lagrangians at Finite Volume." In: *Nucl. Phys. B* 307 (1988), pp. 763–778. DOI: 10.1016/0550-3213(88)90107-1.
- [242] M. Lüscher. "Volume Dependence of the Energy Spectrum in Massive Quantum Field Theories. 1. Stable Particle States." In: *Commun. Math. Phys.* 104 (1986), p. 177. DOI: 10.1007/BF01211589.
- [243] P. Hasenfratz and H. Leutwyler. "Goldstone Boson Related Finite Size Effects in Field Theory and Critical Phenomena With $O(N)$ Symmetry." In: *Nucl. Phys. B* 343 (1990), pp. 241–284. DOI: 10.1016/0550-3213(90)90603-B.
- [244] G. Colangelo, S. Dürr, and C. Haefeli. "Finite volume effects for meson masses and decay constants." In: *Nucl. Phys. B* 721 (2005), pp. 136–174. DOI: 10.1016/j.nuclphysb.2005.05.015. arXiv: hep-lat/0503014.

- [245] M. Bruno, J. Finkenrath, F. Knechtli, B. Leder, and R. Sommer. "Effects of Heavy Sea Quarks at Low Energies." In: *Phys. Rev. Lett.* 114.10 (2015), p. 102001. DOI: 10.1103/PhysRevLett.114.102001. arXiv: 1410.8374 [hep-lat].
- [246] S. Cali, F. Knechtli, and T. Korzec. "Comparison between models of QCD with and without dynamical charm quarks." In: *PoS LATTICE2018* (2018), p. 082. DOI: 10.22323/1.334.0082. arXiv: 1811.05285 [hep-lat].
- [247] A. Hasenfratz, R. Hoffmann, and S. Schaefer. "Reweighting towards the chiral limit." In: *Phys. Rev. D* 78 (2008), p. 014515. DOI: 10.1103/PhysRevD.78.014515. arXiv: 0805.2369 [hep-lat].
- [248] W. Bietenholz et al. "Tuning the strange quark mass in lattice simulations." In: *Phys. Lett. B* 690 (2010), pp. 436–441. DOI: 10.1016/j.physletb.2010.05.067. arXiv: 1003.1114 [hep-lat].
- [249] M. Gell-Mann, R. J. Oakes, and B. Renner. "Behavior of current divergences under $SU(3) \times SU(3)$." In: *Phys. Rev.* 175 (1968), pp. 2195–2199. DOI: 10.1103/PhysRev.175.2195.
- [250] O. Bär and M. Golterman. "Chiral perturbation theory for gradient flow observables." In: *Phys. Rev. D* 89.3 (2014). [Erratum: *Phys.Rev.D* 89, 099905 (2014)], p. 034505. DOI: 10.1103/PhysRevD.89.034505. arXiv: 1312.4999 [hep-lat].
- [251] T. Korzec. *Mass Derivatives*. Internal notes. 2018.
- [252] T. Korzec. URL: <https://github.com/to-ko/mesons>.
- [253] M. Antonelli et al. "Flavor Physics in the Quark Sector." In: *Phys. Rept.* 494 (2010), pp. 197–414. DOI: 10.1016/j.physrep.2010.05.003. arXiv: 0907.5386 [hep-ph].
- [254] Y.-B. Yang et al. "Charm and strange quark masses and f_{D_s} from overlap fermions." In: *Phys. Rev. D* 92.3 (2015), p. 034517. DOI: 10.1103/PhysRevD.92.034517. arXiv: 1410.3343 [hep-lat].
- [255] K. Nakayama, B. Fahy, and S. Hashimoto. "Short-distance charmonium correlator on the lattice with Möbius domain-wall fermion and a determination of charm quark mass." In: *Phys. Rev. D* 94.5 (2016), p. 054507. DOI: 10.1103/PhysRevD.94.054507. arXiv: 1606.01002 [hep-lat].
- [256] N. Carrasco et al. "Up, down, strange and charm quark masses with $N_f = 2+1+1$ twisted mass lattice QCD." In: *Nucl. Phys. B* 887 (2014), pp. 19–68. DOI: 10.1016/j.nuclphysb.2014.07.025. arXiv: 1403.4504 [hep-lat].
- [257] B. Chakraborty, C. Davies, B. Galloway, P. Knecht, J. Koponen, G. C. Donald, R. J. Dowdall, G. P. Lepage, and C. McNeile. "High-precision quark masses and QCD coupling from $n_f = 4$ lattice QCD." In: *Phys. Rev. D* 91.5 (2015), p. 054508. DOI: 10.1103/PhysRevD.91.054508. arXiv: 1408.4169 [hep-lat].

- [258] A. Bazavov et al. "Charmed and Light Pseudoscalar Meson Decay Constants from Four-Flavor Lattice QCD with Physical Light Quarks." In: *Phys. Rev. D* 90.7 (2014), p. 074509. DOI: 10.1103/PhysRevD.90.074509. arXiv: 1407.3772 [hep-lat].
- [259] A. Bazavov et al. "Up-, down-, strange-, charm-, and bottom-quark masses from four-flavor lattice QCD." In: *Phys. Rev. D* 98.5 (2018), p. 054517. DOI: 10.1103/PhysRevD.98.054517. arXiv: 1802.04248 [hep-lat].
- [260] S. Capitani, M. Guagnelli, M. Lüscher, S. Sint, R. Sommer, P. Weisz, and H. Wittig. "Nonperturbative quark mass renormalization." In: *Nucl. Phys. B Proc. Suppl.* 63 (1998). Ed. by C. Davies, I. Barbour, K. Bowler, R. Kenway, B. Pendleton, and D. Richards, pp. 153–158. DOI: 10.1016/S0920-5632(97)00707-X. arXiv: hep-lat/9709125.
- [261] J. Rolf and S. Sint. "A precise determination of the charm quark's mass in quenched QCD." In: *JHEP* 12 (2002), p. 007. DOI: 10.1088/1126-6708/2002/12/007. arXiv: hep-ph/0209255.
- [262] J. Heitger and A. Jüttner. "Lattice cutoff effects for F(D(s)) with improved Wilson fermions: A Final lesson from the quenched case." In: *JHEP* 05 (2009). [Erratum: *JHEP* 01, 036 (2011)], p. 101. DOI: 10.1007/JHEP01(2011)036. arXiv: 0812.2200 [hep-lat].
- [263] J. Heitger, G. M. von Hippel, S. Schaefer, and F. Virota. "Charm quark mass and D-meson decay constants from two-flavour lattice QCD." In: *PoS LATTICE2013* (2014), p. 475. DOI: 10.22323/1.187.0475. arXiv: 1312.7693 [hep-lat].
- [264] J. Heitger, F. Joswig, and S. Kuberski. "Towards the determination of the charm quark mass on $N_f = 2 + 1$ CLS ensembles." In: *37th International Symposium on Lattice Field Theory (Lattice 2019) Wuhan, Hubei, China, June 16-22, 2019*. 2019. arXiv: 1909.05328 [hep-lat].
- [265] F. Joswig. in preparation. PhD thesis. University of Münster, 2021.
- [266] J. Heitger, F. Joswig, and S. Kuberski. *The charm quark mass in $N_f = 2 + 1$ lattice QCD with Wilson fermions*. in preparation. 2020.
- [267] P. A. Boyle, L. Del Debbio, A. Jüttner, A. Khamseh, F. Sanfilippo, and J. T. Tsang. "The decay constants f_D and f_{D_s} in the continuum limit of $N_f = 2 + 1$ domain wall lattice QCD." In: *JHEP* 12 (2017), p. 008. DOI: 10.1007/JHEP12(2017)008. arXiv: 1701.02644 [hep-lat].
- [268] G. Bali. *Fixing the charm quark mass*. Internal notes. 2018.
- [269] S. Collins and W. Söldner. private communication.

- [270] A. Gérardin, M. Cè, G. von Hippel, B. Hörz, H. B. Meyer, D. Mohler, K. Ottnad, J. Wilhelm, and H. Wittig. “The leading hadronic contribution to $(g - 2)_\mu$ from lattice QCD with $N_f = 2 + 1$ flavours of $O(a)$ improved Wilson quarks.” In: *Phys. Rev. D* 100.1 (2019), p. 014510. DOI: 10.1103/PhysRevD.100.014510. arXiv: 1904.03120 [hep-lat].
- [271] M. Della Morte, N. Garron, M. Papinutto, and R. Sommer. “Heavy quark effective theory computation of the mass of the bottom quark.” In: *JHEP* 01 (2007), p. 007. DOI: 10.1088/1126-6708/2007/01/007. arXiv: hep-ph/0609294.
- [272] B. Blossier, M. Della Morte, N. Garron, G. von Hippel, T. Mendes, H. Simma, and R. Sommer. “HQET at order $1/m$: III. Decay constants in the quenched approximation.” In: *JHEP* 12 (2010), p. 039. DOI: 10.1007/JHEP12(2010)039. arXiv: 1006.5816 [hep-lat].
- [273] F. Bahr, D. Banerjee, F. Bernardoni, A. Joseph, M. Koren, H. Simma, and R. Sommer. “Continuum limit of the leading-order HQET form factor in $B_s \rightarrow K\ell\nu$ decays.” In: *Phys. Lett. B* 757 (2016), pp. 473–479. DOI: 10.1016/j.physletb.2016.03.088. arXiv: 1601.04277 [hep-lat].
- [274] F. Bahr, D. Banerjee, F. Bernardoni, M. Koren, H. Simma, and R. Sommer. “Extraction of bare form factors for $B_s \rightarrow K\ell\nu$ decays in nonperturbative HQET.” In: *Int. J. Mod. Phys. A* 34.28 (2019), p. 1950166. DOI: 10.1142/S0217751X19501665. arXiv: 1903.05870 [hep-lat].
- [275] F. Bernardoni, J. Bulava, M. Donnellan, and R. Sommer. “Precision lattice QCD computation of the $B^*B\pi$ coupling.” In: *Phys. Lett. B* 740 (2015), pp. 278–284. DOI: 10.1016/j.physletb.2014.11.051. arXiv: 1404.6951 [hep-lat].
- [276] M. Bruno. “The energy scale of the 3-flavour Lambda parameter.” PhD thesis. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät, 2016. DOI: <http://dx.doi.org/10.18452/17516>.
- [277] A. Anastassov et al. “First measurement of $\Gamma(D^{*+})$ and precision measurement of $m(D^{*+}) - m(D_0)$.” In: *Phys. Rev. D* 65 (2002), p. 032003. DOI: 10.1103/PhysRevD.65.032003. arXiv: hep-ex/0108043.
- [278] D. Becirevic and F. Sanfilippo. “Theoretical estimate of the $D^* \rightarrow D\pi$ decay rate.” In: *Phys. Lett. B* 721 (2013), pp. 94–100. DOI: 10.1016/j.physletb.2013.03.004. arXiv: 1210.5410 [hep-lat].

- [279] J. Flynn, P. Fritzsche, T. Kawanai, C. Lehner, B. Samways, C. Sachrajda, R. Van de Water, and O. Witzel. "The $B^*B\pi$ Coupling Using Relativistic Heavy Quarks." In: *Phys. Rev. D* 93.1 (2016), p. 014510. DOI: 10.1103/PhysRevD.93.014510. arXiv: 1506.06413 [hep-lat].
- [280] G. de Divitiis, L. Del Debbio, M. Di Pierro, J. Flynn, C. Michael, and J. Peisa. "Towards a lattice determination of the $B^*B\pi$ coupling." In: *JHEP* 10 (1998), p. 010. arXiv: hep-lat/9807032.
- [281] H. Ohki, H. Matsufuru, and T. Onogi. "Determination of $B^*B\pi$ coupling in unquenched QCD." In: *Phys. Rev. D* 77 (2008), p. 094509. DOI: 10.1103/PhysRevD.77.094509. arXiv: 0802.1563 [hep-lat].
- [282] D. Becirevic, B. Blossier, E. Chang, and B. Haas. " $g_{B^*B\pi}$ -coupling in the static heavy quark limit." In: *Phys. Lett. B* 679 (2009), pp. 231–236. DOI: 10.1016/j.physletb.2009.07.031. arXiv: 0905.3355 [hep-ph].
- [283] W. Detmold, C. Lin, and S. Meinel. "Calculation of the heavy-hadron axial couplings g_1 , g_2 and g_3 using lattice QCD." In: *Phys. Rev. D* 85 (2012), p. 114508. DOI: 10.1103/PhysRevD.85.114508. arXiv: 1203.3378 [hep-lat].
- [284] G. Burdman and J. F. Donoghue. "Union of chiral and heavy quark symmetries." In: *Phys. Lett.* B280 (1992), pp. 287–291. DOI: 10.1016/0370-2693(92)90068-F.
- [285] T.-M. Yan, H.-Y. Cheng, C.-Y. Cheung, G.-L. Lin, Y. C. Lin, and H.-L. Yu. "Heavy quark symmetry and chiral dynamics." In: *Phys. Rev. D* 46 (1992). [Erratum: *Phys. Rev. D* 55, 5851 (1997)], pp. 1148–1164. DOI: 10.1103/PhysRevD.46.1148, 10.1103/PhysRevD.55.5851.
- [286] M. B. Wise. "Chiral perturbation theory for hadrons containing a heavy quark." In: *Phys. Rev. D* 45.7 (1992), p. 2188. DOI: 10.1103/PhysRevD.45.R2188.
- [287] R. Casalbuoni, A. Deandrea, N. Di Bartolomeo, R. Gatto, F. Feruglio, and G. Nardulli. "Phenomenology of heavy meson chiral Lagrangians." In: *Phys. Rept.* 281 (1997), pp. 145–238. DOI: 10.1016/S0370-1573(96)00027-0. arXiv: hep-ph/9605342 [hep-ph].
- [288] B. Grinstein. "An Introduction to Heavy Mesons." In: *6th Mexican School of Particles and Fields*. Aug. 1995, pp. 122–184. arXiv: hep-ph/9508227.
- [289] S. Scherer. "Introduction to chiral perturbation theory." In: *Adv. Nucl. Phys.* 27 (2003). Ed. by J. W. Negele and E. Vogt, p. 277. arXiv: hep-ph/0210398.

- [290] J. Bulava, M. Donnellan, and R. Sommer. "On the computation of hadron-to-hadron transition matrix elements in lattice QCD." In: *JHEP* 01 (2012), p. 140. DOI: 10.1007/JHEP01(2012)140. arXiv: 1108.3774 [hep-lat].
- [291] J. Bulava, M. Donnellan, and R. Sommer. "The $B^*B\pi$ Coupling in the Static Limit." In: *PoS LATTICE2010* (2010). Ed. by G. Rossi, p. 303. DOI: 10.22323/1.105.0303. arXiv: 1011.4393 [hep-lat].
- [292] W. Detmold, C.-J. Lin, and S. Meinel. "Axial couplings in heavy hadron chiral perturbation theory at the next-to-leading order." In: *Phys. Rev. D* 84 (2011), p. 094502. DOI: 10.1103/PhysRevD.84.094502. arXiv: 1108.5594 [hep-lat].
- [293] A. Gérardin. "Heavy-light meson properties from lattice QCD." PhD thesis. Université Paris Sud - Paris XI, 2014. URL: <https://tel.archives-ouvertes.fr/tel-01126879>.

