Determination of c_{sw} in $N_f = 3 + 1$ Lattice QCD with massive Wilson fermions

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)
im Fach Physik
Spezialisierung: Theoretische Physik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
Herrn Dipl.-Phys. Felix Stollenwerk

Präsidentin der Humboldt-Universität zu Berlin
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät
Prof. Dr. Elmar Kulke

Gutachter:
1. Prof. Dr. Ulrich Wolff
2. Prof. Dr. Rainer Sommer
3. Prof. Dr. Kari Rummukainen

eingereicht am: 14.12.2016
Tag der mündlichen Prüfung: 07.02.2017

Anything that we scientists can do to weaken the hold of religion should be done, and may in fact be our greatest contribution to civilization.

—Steven Weinberg

Abstract

In order to obtain sensible results from Lattice QCD that may be compared with experiment, extrapolation to the continuum is crucial. This procedure, however, constitutes a significant source of systematic and statistical errors, which one aims to minimize. The well-established Symanzik improvement program systematically reduces the order of cutoff effects, allowing for better control of the aforementioned errors, as well as larger and thus more affordable lattice spacings. Applied to the Wilson fermion action, it entails the addition of the Sheikholeslami–Wohlert term with the $O(a)$ improvement coefficient c_{sw}. In this work, a strategy is developed for the non-perturbative determination of c_{sw} in the theory with $N_f = 3 + 1$ massive sea quarks. It is embedded in a general, mass-dependent renormalization and improvement scheme, for which we lay the foundations. The improvement condition, formulated by means of the PCAC relation in the Schrödinger functional, is imposed along a line of constant physics that is designed to be close to the physical mass of the charm quark. The aim of this rather elaborate approach is to avoid large, mass-dependent $O(a^2)$ effects in future large volume simulations with four dynamical quark species. The numerical results are worked out using the tree-level improved Lüscher–Weisz gauge action. Since the gradient flow coupling is employed in the definition of the line of constant physics, its interdependence with the topological charge in regard to critical slowing down and topology freezing is investigated in a supplemental study.

Keywords:
Lattice QCD, $O(a)$ improvement, non-perturbative methods, charm quark

Zusammenfassung

Um aussagekräftige, mit dem Experiment vergleichbare Resultate aus Berechnungen der Gitter-QCD zu erhalten, ist die Extrapolation zum Kontinuum unabdingbar. Diese Prozedur stellt jedoch eine maßgebliche Quelle systematischer und statistischer Fehler dar, die es zu minimieren gilt. Das bewährte Symanzik-Verbesserungsprogramm führt zu einer systematischen Reduzierung der Ordnung von Cutoff-Effekten, die eine bessere Kontrolle über die genannten Fehler sowie größere und damit erschwinglichere Gitterabstände ermöglicht. Auf die Wilson-Fermionenwirkung bezogen bedarf es nur des Hinzufügens des Sheikholeslami-Wohlert-Terms mit dem $O(a)$-Verbesserungskoeffizienten c_{sw}. In der vorliegenden Arbeit wird eine Strategie zur nicht-perturbativen Bestimmung dieses Koeffizienten in der Theorie mit $N_f = 3 + 1$ massiven Seequarks entwickelt. Diese ist in ein allgemeines, massenabhängiges Renormierungs- und Verbesserungsschema eingebettet, dessen Grundlagen dargelegt werden. Die Auferlegung der Verbesserungsbedingung, bei der die PCAC-Relation im Schrödinger-Funktional Verwendung findet, geschieht entlang einer Linie konstanter Physik, welche dem Charm-Quark näherungsweise seine physikalische Masse zuordnet. Dieser vergleichsweise aufwendige Ansatz hat zum Ziel, große, massenabhängige $O(a^2)$-Effekte in zukünftigen Simulationen im großen Volumen mit vier dynamischen Quarkspezies zu vermeiden. Die numerischen Resultate dieser Arbeit werden unter Verwendung der tree-level-verbesserten Lüscher-Weisz-Eichwirkung gewonnen. Da die sogenannte Gradient-Flow-Kopplung bei der Definition der Linie konstanter Physik Verwendung findet, wird in einer zusätzlichen Untersuchung die Wechselbeziehung dieser Kopplung mit der Topologischen Ladung beleuchtet, insbesondere im Bezug auf die unter den Namen Critical Slowing Down und Topology Freezing bekannten Phänomene.

Schlagwörter:
Gitter-QCD, $O(a)$-Verbesserung, nicht-perturbative Methoden, Charm-Quark

Contents

1 Introduction

Elementary particle physics deals with the most fundamental building blocks of matter and the way they interact. It explores the very basic laws of nature and tackles far-reaching questions about the development of the universe. Our current knowledge is comprised in the *Standard Model of particle physics (SM)*, which consists of a set of quantum field theories that describe the electromagnetic, weak and strong interaction between elementary particles. Their basic properties are compiled in tab. 1.1, and the particles involved in the SM are illustrated in fig. 1.1. Fundamental constituents of matter in the narrow sense are fermions of spin 1/2, the quarks and leptons, for both of which there exist 6 different types, called flavors. While quarks are subject to all interactions, leptons participate only in the electromagnetic and weak interaction. For both quarks and leptons, one distinguishes 3 generations of specific flavor pairs that predominantly appear together in (flavor-changing) processes of the weak interaction. In addition, the SM contains gauge bosons of spin 1, which mediate the forces between the aforementioned fermions. For the strong, weak and electromagnetic interaction, these are gluons g, the W and Z bosons, and the photon γ, respectively.

One of the fundamental principles which guides the interactions between the elementary particles is *gauge invariance*. The postulation of invariance of the action under local gauge transformations of the form[1]

$$SU(3)_c \times SU(2)_L \times U(1)_Y \tag{1.1}$$

not only requires the existence of the gauge fields, but also presents a determining factor that governs the interactions we observe in nature. An immediate consequence of gauge invariance is that the gauge fields are massless. The photon and the gluon are found to be massless in nature indeed. In contrast, the W and Z bosons have a considerable non-vanishing mass, which is responsible for the relatively short range of the weak force. An explanation for this is given by spontaneous breaking of the electroweak symmetry $SU(2)_L \times U(1)_Y$ in terms of the Higgs mechanism. The involved

---

[1] The indices c, L, Y stand for "color", "left" (referring to the chirality of the particles that carry weak isospin) and the "weak hypercharge", respectively. Compare to tab. 1.1.

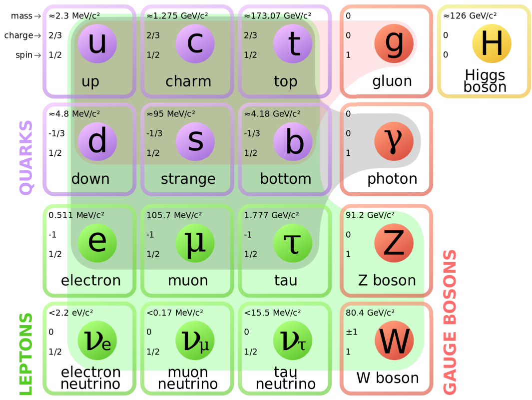


Figure 1.1: Elementary particles of the Standard Model. Quarks and leptons are shown in violet and green, respectively, the three columns representing the different generations. The gauge bosons are displayed in red. As indicated by the shades, gluons interact with quarks, photons with quarks and charged leptons, while the Z and W bosons couple to all quarks and leptons. The Higgs boson is depicted in yellow. The figure is taken from [1].

Higgs field, which also provides mass to the quarks and leptons via Yukawa coupling, completes the SM. The existence of its excitation, the scalar Higgs boson, has been confirmed at the Large Hadron Collider recently [2, 3], around 50 years after it had been predicted theoretically (see e.g. [4, 5]).

The present work is concerned with *quantum chromodynamics (QCD)* and the strong interaction. A striking feature of this interaction is the heavy dependence of its strength on the energy, which manifests itself in a very different phenomenology in the high and low energy realm. While at high energies, the interaction of quarks and gluons becomes weak (*asymptotic freedom*), it becomes so strong at low energies that quarks are not observed as free particles, a phenomenon known as *confinement*. Accordingly, the mathematical approaches differ depending on the energies involved. While at high energies, perturbative methods are applicable as in the case of electromagnetic and weak interactions, the strong coupling at low energies calls for a non-perturbative treatment.

The most successful non-perturbative approach to describe the phenomenology of the strong interaction at low energies is *Lattice QCD*. Its starting point is the path integral formalism, in which a discretized spacetime serves as a regulator. It employs massive computational power as well as advanced numerical methods and algorithms, and allows to compute low energy observables like for instance hadron masses and decay constants. Among many further applications, it also provides a tool to extract

Interaction	Strong interaction	Weak interaction	Electromagnetism
Theory	Quantum chromodynamics	Glashow–Salam–Weinberg theory $SU(2)_L \times U(1)_Y$	
		$\hookrightarrow$ SSM $\rightarrow$	Quantum electrodynamics
Charge	Color (r,g,b)	Weak isospin	Electromagnetic charge
Gauge group	$SU(3)_c$	$SU(2)_L$	$U(1)_Q$
Gauge fields (#)	g (8)	W, Z (3)	γ (1)
Quarks	yes	yes	yes
Leptons	no	yes	yes
Range	≈ 1 fm (confinement)	≈ 0.1 fm (massive gauge bosons)	∞
Relative strength	1	10^{-5}	1/137
CPT	C,P,T	CP, T	C,P,T
Physical phenomena	Binding of atomic nuclei	Radioactive decay of atomic nuclei	Binding of atoms, optics, electricity

Table 1.1: The interactions and theories of the Standard Model of particle physics. Gravity, the 4th known interaction and its corresponding theory, general relativity, is not part of the SM.

the fundamental parameters of QCD, the renormalization group invariant masses and the Λ parameter, which is one of the major objectives of the ALPHA Collaboration (see [6] for an overview).

Simulations in Lattice QCD are necessarily carried out in finite spacetime, and in general at unphysically large quark (or equivalently, meson) masses for reasons of affordability. This can be accounted for by a set of *extrapolations*. First, the infinite volume limit $V \to \infty$ is taken, although in practice finite spacetime effects are sufficiently suppressed if the correlation length of the lowest energy bound state, the pion, fits well into the lattice ($m_\pi L \gtrsim 4$). Second, the pion masses are extrapolated to the physical point, which by reference to the chiral symmetry of the QCD Lagrangian in the massless limit is called chiral extrapolation. Finally, in order to get results comparable with experiment, the continuum limit needs to be taken, i.e. the lattice spacing a is sent to zero, $a \to 0$.

The discretization of the theory leaves room for different formulations on the lattice, and there exist a number of approaches, each with certain advantages and disadvantages. However, all of them are constrained by the requirement to yield the continuum theory if the lattice spacing is removed, such that there is no arbitrariness in the

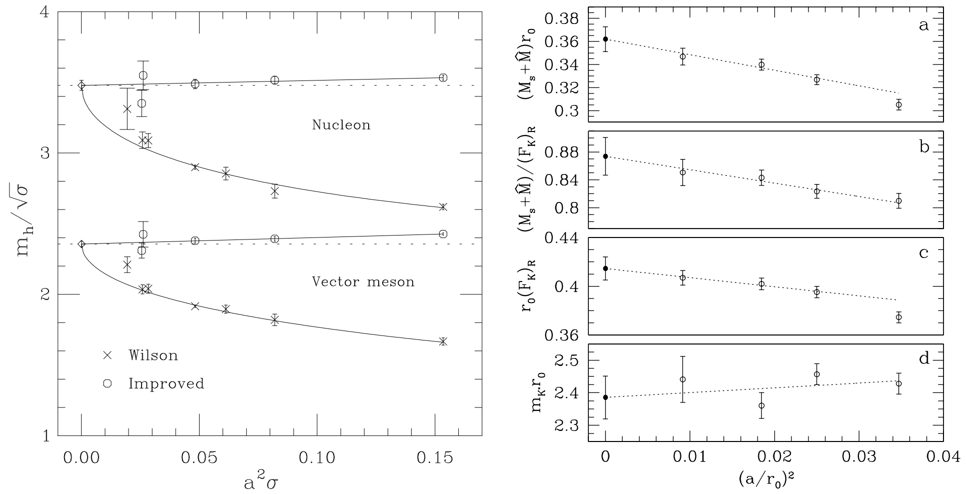


Figure 1.2: *Left:* Comparison of the continuum limit, taken in the (unimproved) Wilson theory and the improved theory [10]. *Right:* Continuum limits of different observables in the improved theory [11]. The figures are taken from the respective references. Physical quantities are given in units of the string tension $\sigma \approx (465\text{ MeV})^2$ [12] and the Sommer scale $r_0 \approx 0.5$ fm [13], i.e. $a \sim 0.1$ fm roughly corresponds to $a^2\sigma \sim 0.05$ and $(a/r_0)^2 \sim 0.04$.

continuum limit. The *Wilson formulation* of Lattice QCD is the most straight forward discretization and a computationally favorable choice, but suffers from cutoff effects of first order in the lattice spacing. In order to be able to continuum extrapolate with reasonable uncertainties, very small lattice spacings are in principle required, which are however computationally expensive and usually unaffordable. Fortunately, the well-established *Symanzik improvement program* [7–9] represents a systematic method to get rid of discretization effects to leading order $O(a)$ in the lattice spacing, rendering the remaining effects smaller at small a and thus enabling a smoother extrapolation to the continuum, with better control of systematic errors. The gain from this is exemplarily depicted in fig. 1.2.

Since the inclusion of dynamical quarks poses another huge computational challenge, simulations in Lattice QCD were conducted in the approximation of no ($N_f = 0$, quenched) or fewer (e.g. $N_f = 2$) dynamical quark flavors in the past. Despite being a rather rough approximation, these computations led to remarkably good results, see e.g. [14]. This is due to the fact that for the low energy observables usually considered in Lattice QCD, the physical effects of dynamical quark flavors decrease in importance as their mass increases. However, for a full description of the strong interaction and high accuracy comparison with experimental data, the dynamical effects also of the heavier flavors (strange, charm) need to be taken into account. Advances in computational power and algorithmic techniques have allowed to gradually approach this goal in recent years.

The present work is the first step towards Lattice QCD simulations with $N_f = 4$ *dynamical quark flavors* within the effort of the ALPHA Collaboration. It covers the improvement of the Wilson fermion action, for which the improvement coefficient c_{sw} [15] as a function of the bare coupling g_0^2 needs to be determined non-perturbatively. This has been done before for various numbers of up to $N_f = 3$ massless, dynamical quark flavors and different actions [16–22]. However, the inclusion of the charm quark poses a new challenge, due to the unprecedented order of magnitude of mass-dependent cutoff effects. This will be accommodated by switching to a mass-dependent renormalization and improvement scheme. Within this framework, c_{sw} will not only be determined under use of massive Wilson fermions, but the strategy is also adapted in the sense that the improvement condition is imposed at a line of constant physics (LCP). The fixation of the volume is realized in terms of the gradient flow coupling. As a supplemental study to the main topic, we investigate it with regard to critical slowing down, which refers to the well-known insufficiency of the Hybrid Monte Carlo algorithm to tunnel between topological sectors as the continuum limit is approached.

The thesis is structured as follows. After a brief summary of the foundations of quantum chromodynamics in sec. 2, its formulation on the lattice is discussed in sec. 3, with an emphasis on the aspects that are important for our purposes. This is followed by a discussion in sec. 4 of Monte Carlo simulations and the associated numerical techniques that find application in this work. After a presentation of the Schrödinger functional in sec. 5, the gradient flow coupling and the topological charge are subject of sec. 6. The interplay of those two quantities, in particular how the gradient flow coupling is affected by critical slowing down, is investigated in the quenched approximation in sec. 7. The subsequent two chapters are then devoted to the Symanzik improvement of Lattice QCD with four dynamical quark flavors. First, in sec. 8, the framework of a massive scheme is discussed. Finally, in sec. 9, the determination of c_{sw} within this scheme is presented. We close with a summary in sec. 10.

2 Quantum chromodynamics

2.1 Color, $SU(3)_c$ and the QCD Lagrangian

Quantum chromodynamics is the theory of strongly interacting particles, i.e. quarks and gluons. At its foundation is the notion of the color degree of freedom, which is supported by experimental evidence of several kinds. First, the existence of color reconciles the appearance of certain baryons with the Pauli principle. The Δ^{++}, for instance, possesses both a symmetric spatial and spin wave function. The additional color degree of freedom ensures that the overall wave function is antisymmetric as required for baryons, through the antisymmetric color wavefunction

$$|qqq\rangle = \frac{1}{\sqrt{6}}\,\epsilon_{abc}|q^a q^b q^c\rangle\ , \tag{2.1}$$

where the color indices $a, b, c = 1, 2, 3$ are summed over. Correspondingly, color-symmetric meson wavefunctions are given by

$$|q\overline{q}\rangle = \frac{1}{\sqrt{3}}\,\sum_c |q^c \overline{q}^c\rangle\ . \tag{2.2}$$

The bound states' wavefunctions (2.1, 2.2) are invariant under global $SU(3)_c$ transformations, i.e. they are color-neutral. In analogy to chromatics, one labels the 3 new degrees of freedom red, green and blue, and the color-neutral hadrons are called white. In principle, antisymmetric baryon color wave functions could also be realized with a different number N_c of colors. However, the color factor $N_c = 3$ enters experimental observables like for instance the decay width $\Gamma(\pi^0 \to 2\gamma)$, the branching fraction $B(\tau^- \to e^-\overline{\nu}_e\nu_\tau)$ and the cross section ratio $R = \frac{\sigma(e^+ + e^- \to \text{hadrons})}{\sigma(e^+ + e^- \to \mu^+ + \mu^-)}$. Moreover, the renormalizability of the Glashow–Salam–Weinberg theory can only be achieved if the contributions of all leptons and quarks to the Adler–Bell–Jackiw (ABJ) anomaly [23, 24] in the electroweak sector cancel each other. This requires an equal number of leptons and quarks (lepton-quark symmetry) in connection with $N_c = 3$ [25].

As massive fermions, quarks are described mathematically by Dirac spinors $\psi, \overline{\psi}$. In

Generation	Flavor	Mass	Q_{em}
I	up (u)	$2.3^{+0.7}_{-0.5}$ MeV	$+2/3$
	down (d)	$4.8^{+0.5}_{-0.3}$ MeV	$-1/3$
II	charm (c)	1.275 ± 0.025 GeV	$+2/3$
	strange (s)	95 ± 5 MeV	$-1/3$
III	top (t)	$173.21 \pm 0.51 \pm 0.71$ GeV	$+2/3$
	bottom (b)	4.18 ± 0.03 GeV	$-1/3$

Table 2.1: Electromagnetic charge and renormalized masses [26] of the quarks in the Standard Model. While the top quark mass corresponds to the so-called direct measurement, the other masses $\overline{m}_i$ are given in the $\overline{\text{MS}}$ scheme at a renormalization scale $\mu = 2$ GeV (i=u,d,s) and $\mu = \overline{m}_i$ (i=c,b).

accordance with $N_c = 3$, each of those carries a color index $c \in \{1,2,3\}$, as well as a Dirac index $\alpha \in \{1,2,3,4\}$ and a flavor index $f \in \{1,\ldots,N_f\}$, where N_f is the number of quark flavors in the theory. Furthermore, the quark fields are functions of space and time, indicated in the continuum by the spacetime coordinates $x = (x_0, x_1, x_2, x_3)$:

$$\psi(x) \equiv \psi^{(f)}(x)^c_\alpha \; . \tag{2.3}$$

The color and flavor indices are often summed over, in which case they will usually be suppressed in favor of the matrix-vector notation shown on the left hand side of (2.3). The Dirac equation, which describes the free propagation of the quark fields, emerges as the Euler-Lagrange equation from the *free quark Lagrangian*

$$\mathcal{L}_F^{\text{free}}[\psi,\overline{\psi}] = \overline{\psi}(x)\left(i\gamma^\mu \partial_\mu - \mathcal{M}\right)\psi(x) \; , \tag{2.4}$$

where γ_μ are the gamma matrices and

$$\mathcal{M} = \text{diag}(m_{0,1},\ldots,m_{0,N_f}) = \text{diag}(m_{0,u}, m_{0,d},\ldots) \tag{2.5}$$

is the matrix of bare quark masses $m_{0,i}$ in flavor space. The corresponding renormalized masses of the 6 different quark flavors in the SM are collected in tab. 2.1

The fundamental principle which rules the construction of interaction terms is the postulation of invariance under local $SU(3)_c$ gauge transformations $\Omega(x)$ in color space (cf. app. A.1):

$$\psi \;\rightarrow\; \psi'(x) = \Omega(x)\; \psi(x) \tag{2.6}$$

$$\overline{\psi} \;\rightarrow\; \overline{\psi}'(x) = \overline{\psi}(x)\; \Omega(x)^\dagger \;. \tag{2.7}$$

In order for the QCD Lagrangian to be invariant under (2.6, 2.7), $\mathcal{L}_F' = \mathcal{L}_F$, the ordinary derivative ∂_μ in (2.4) needs to be replaced by a *covariant derivative* D_μ with the transformation property

$$D_\mu(x) \;\rightarrow\; D_\mu'(x) = \Omega(x) D_\mu(x) \Omega(x)^\dagger \;, \tag{2.8}$$

which together with (2.6) implies

$$D_\mu\psi(x) \;\rightarrow\; D_\mu'\psi'(x) = \Omega(x)\; D_\mu\psi(x) \;, \tag{2.9}$$

i.e. $D_\mu\psi$ transforms the same way as ψ. The required behavior (2.8) can only be achieved by introducing a coupling of the quark fields to gluon (or gauge) fields A_μ, parametrized[1] by g_0:

$$D_\mu(x) = \partial_\mu + i g_0 A_\mu(x) \;. \tag{2.10}$$

The *continuum gluon fields* are represented at each spacetime point x by a traceless, hermitean 3×3 color matrix

$$A_\mu(x) \equiv A_\mu(x)^{cd} \;. \tag{2.11}$$

As an element of the algebra $\mathfrak{su}(3)$, A_μ may be represented in the basis of $SU(3)$ generators (see app. A.1):

$$A_\mu(x) = A_\mu^a(x) T^a \;. \tag{2.12}$$

The Lorentz index $\mu = 1, 2, 3, 4$ reflects the fact that there is one component per spacetime direction. The covariance of D_μ (cf. (2.8)), is fulfilled by having the gluon fields transform according to the adjoint representation of $SU(3)$,

$$A_\mu'(x) = \Omega(x) A_\mu(x) \Omega(x)^\dagger + \frac{i}{g_0} \left(\partial_\mu \Omega(x)\right) \Omega(x)^\dagger \;. \tag{2.13}$$

Consequently, the interaction of the quark fields with the gluon fields is incorporated

---

[1] Note that the sign in (2.10) is a matter of convention, which propagates to (2.13) and (2.16). The opposite choice is also widely used in the literature.

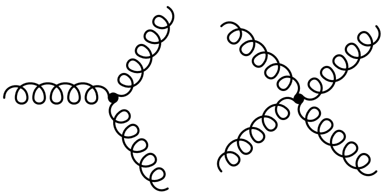


Figure 2.1: Gluon self-interaction in the form of 3- and 4-gluon vertices which emerge from $\mathcal{L}_G$ (2.15). The figure has been created under use of the `feynMF` package [27].

in

$$\mathcal{L}_F[\psi,\overline{\psi},A] = \overline{\psi}(x)\left(i\gamma^\mu D_\mu - \mathcal{M}\right)\psi(x)\ . \tag{2.14}$$

In addition, the kinetic terms of the gluon fields are contained in

$$\mathcal{L}_G[A] = -\frac{1}{2}\,\mathrm{tr}\left[F_{\mu\nu}F^{\mu\nu}\right]\ , \tag{2.15}$$

where tr is the trace in color space, and $F_{\mu\nu}$ is the *gluon field strength tensor*

$$F_{\mu\nu}(x) = -\frac{i}{g_0}[D_\mu(x), D_\nu(x)] \tag{2.16a}$$

$$= \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) + ig_0[A_\mu(x), A_\nu(x)] \tag{2.16b}$$

$$= \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) - g_0 f^{abc} A^a_\mu(x) A^b_\nu(x) T^c\ , \tag{2.16c}$$

where in the last step we used the non-vanishing commutator of the generators, see (A.3) in app. A.1. In particular, (2.16) incorporates the coupling of gluons to themselves, which in connection with (2.15) leads to self-interacting terms that correspond to the Feynman diagrams displayed in fig. 2.1. This is a new feature compared to quantum electrodynamics (QED), reflecting the non-Abelian nature of the gauge group $SU(3)$. It has severe phenomenological implications, which we are going to describe in sec. 2.3. Gauge invariance of $\mathcal{L}_G$ (2.15) is ensured as the field strength tensor transforms covariantly, too:

$$F'_{\mu\nu}(x) = \Omega(x)F_{\mu\nu}(x)\Omega(x)^\dagger\ . \tag{2.17}$$

Finally, in order for the gluon propagator to be well-defined, in the continuum one needs a gauge fixing term

$$\mathcal{L}_{\mathrm{gf}}[A] = -\frac{1}{2\xi}\left(\partial_\mu A^\mu\right)^2\ . \tag{2.18}$$

The *full Lagrangian of QCD*[2] thus reads

$$\mathcal{L}_{QCD}[\psi,\overline{\psi},A] = \mathcal{L}_F[\psi,\overline{\psi},A] + \mathcal{L}_G[A] + \mathcal{L}_{\mathrm{gf}}[A]\ , \tag{2.19}$$

with the individual parts given in (2.14), (2.15) and (2.18), respectively.

Apart from the gauge principle, it is based on Lorentz invariance and renormalizability. In particular, the latter restricts the interaction terms in $\mathcal{L}_{QCD}$ to a mass dimension of ≤ 4. In addition to the local $SU(3)_c$ symmetry, $\mathcal{L}_{QCD}$ is invariant under the global transformations C (charge conjugation), P (parity) and T (time reversal), cf. tab. 1.1. However, in principle, an additional term

$$\mathcal{L}_\Theta = \Theta \frac{g_0^2}{16\pi^2}\ \mathrm{tr}\left[\widetilde{F}_{\mu\nu}F^{\mu\nu}\right] \tag{2.20}$$

with $\widetilde{F}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}F^{\alpha\beta}$ is allowed by the aforementioned principles. It would violate P, T, and (due to the CPT theorem) also the combined transformation CP. Since CP violation would manifest itself in an electric dipole moment of the neutron, which can be ruled out by experimental data up to a very small value, the parameter Θ can be restricted to $|\Theta| \leq 10^{-10}$ [26]. The unanswered question of why the strong interaction does not seem to break CP is called the *strong CP problem*. In any case, as the Θ angle is so small, we may ignore $\mathcal{L}_\Theta$ in the following.

2.2 Chiral symmetry

The QCD Lagrangian exhibits (approximate) global flavor symmetries, which depend on the nature of the quark masses and have important phenomenological consequences. In the *massless* case, the left- and right-hand components of the spinor ψ, defined by

$$\psi_R = P_+\psi, \quad \psi_L = P_-\psi, \quad \overline{\psi}_R = \overline{\psi}P_-, \quad \overline{\psi}_L = \overline{\psi}P_+ \tag{2.21}$$

in terms of the projection operator $P_\pm = \frac{\mathbb{1}\pm\gamma_5}{2}$ and $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$, decouple, and the fermionic Lagrangian $\mathcal{L}_F^0 \equiv \mathcal{L}_F|_{\mathcal{M}=0}$ cf. (2.14) becomes

$$\mathcal{L}_F^0 = \overline{\psi}_L\gamma^\mu D_\mu\psi_L + \overline{\psi}_R\gamma^\mu D_\mu\psi_R\ . \tag{2.22}$$

---

[2] We here ignore the Faddeev-Popov ghosts, which deal with the issue of unphysical, longitudinal degrees of freedom of the gluon fields.

The chiral components can thus be transformed independently, i.e. $\mathcal{L}_F^0$ is invariant under chiral transformations (see app. A.1) of the form

$$U(N_f)_{fL} \times U(N_f)_{fR} = SU(N_f)_{fL} \times SU(N_f)_{fR} \times U(1)_{fL} \times U(1)_{fR} \ . \tag{2.23}$$

Here, the subscripts L and R indicate transformations on the left- and right-handed fields, and the subscript f emphasizes that we now deal with transformations in flavor space in contrast to the color space associated with gauge transformations. The combinations of left- and right-handed transformations in (2.23) may be rewritten in terms of vector and axial transformations as

$$SU(N_f)_{fV} \times {''}SU(N_f)_{fA}^{''} \times U(1)_{fV} \times U(1)_{fA} \ , \tag{2.24}$$

see app. A.1 for details and notation. The underlying reason for the symmetry under axial transformations is that the Dirac operator, $D \equiv \gamma^\mu D_\mu$, anticommutes with γ_5:

$$\{\gamma_5, D\} = 0 \ . \tag{2.25}$$

Details on the gamma matrices can be found in app. A.2.

If one includes *degenerate* masses in the Lagrangian, the axial symmetries are broken explicitly, while the subgroup $SU(N_f)_{fV}$ remains valid for $N_f = N_f^{\mathrm{deg}}$ degenerate quark flavors:

$$SU(N_f^{\mathrm{deg}})_{fV} \times U(1)_{fV} \ . \tag{2.26}$$

Due to $m_u \approx m_d$, the isospin symmetry $SU(2)_{fV}$ is a very good approximate symmetry, and even remnants of $SU(3)_{fV}$ are reasonably manifested in hadron physics. The conserved quantities associated with $SU(2)_{fV}$ and $U(1)_{fV}$ are isospin and baryon number, respectively. The observation of hadron multiplets (cf. fig. 2.2), the lightest of which are approximately degenerate in mass, confirm that $SU(N_f)_{fV}$ is an (approximate) symmetry not just of the Lagrangian, but also the ground state[3] and hence the full theory, i.e it is realized in the Wigner–Weyl mode. For *non-degenerate* masses, $SU(N_f)_{fV}$ breaks down to an N_f-fold $U(1)_{fV}$ symmetry, one for each flavor, such that the corresponding flavor quantum numbers are still conserved.

Since the masses of the up and down, and to lesser extent also the strange quark are quite small, the axial symmetries ${''}SU(N_f)_{fA}^{''}$ and $U(1)_{fA}$ might still be satisfied

---

[3] Specifically, $Q_V^a \, |0\rangle = 0$ for the vector charges Q_V^a with $a = 1, \ldots, N_f^2 - 1$.

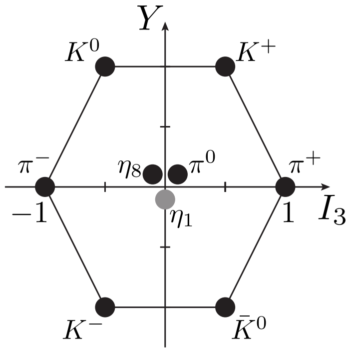


Figure 2.2: $SU(3)_{fV}$ multiplet of approximately mass-degenerate pseudoscalar mesons. The quark-antiquark bound states build an octet (black) and a singlet (gray) according to $3 \otimes \overline{3} = 8 \oplus 1$. They can be classified by the eigenvalues hypercharge Y and isospin I_3.

approximately. However, there is convincing evidence that they—despite being symmetries of the (massless) Lagrangian—are realized differently in nature. First, full $''SU(N_f)''_{fA}$ axial symmetry would imply the occurrence of parity doublets. Instead, the masses of the nucleon N and its negative parity partner $N^\star$ are found to be quite different [26]:

$$m_N \approx 940 \text{ MeV}, \quad m_{N^\star} \approx 1535 \text{ MeV} . \tag{2.27}$$

Second, the three pions, $\pi^\pm$ and π^0, are exceptionally light compared to all other mesons. An explanation for both these observations is that chiral symmetry is spontaneously broken,

$$SU(N_f)_{fL} \times SU(N_f)_{fR} \xrightarrow{SSM} SU(N_f)_{fV} , \tag{2.28}$$

i.e. $''SU(N_f)''_{fA}$ axial symmetry is realized in the Nambu–Goldstone mode and does not leave the ground state invariant[4]. This is reflected by the non-vanishing expectation value of the quark condensate,

$$\langle \overline{\psi}\psi \rangle \neq 0 , \tag{2.29}$$

which transforms like a mass term. In that case, the pions can be identified with the $N_f^2 - 1$ massless Goldstone bosons with pseudoscalar ($J^P = 0^-$) quantum numbers predicted by the Goldstone theorem.

Finally, even in the massless case, the axial current corresponding to $U(1)_{fA}$ is not conserved on the quantum level due to non-vanishing contributions of the ABJ anomaly[5] (cf. sec. 2.1 and sec. 6.2.1) in QCD. An overview on the realization of chiral

---

[4] Specifically, $Q_A^a \, |0\rangle \neq 0$ for the axial charges Q_A^a with $a = 1, \ldots, N_f^2 - 1$.

[5] In the context of QCD, it is usually referred to as *chiral anomaly*.

	$SU(N_f)_{fV}$	$''SU(N_f)''_{fA}$	$U(1)_{fV}$	$U(1)_{fA}$
massless	full theory	only $\mathcal{L}_F$	full theory	only $\mathcal{L}_F$
mass-degenerate	full theory	—	full theory	—
arbitrary masses	$\to U(1)_{fV}$	—	full theory	—
mode	Wigner–Weyl	Nambu–Goldstone	Wigner–Weyl	explicitly broken (chiral anomaly)
conserved quantum number	Isospin, Flavor	—	Baryon number	—
related phenomenology	Hadron multiplets	Goldstone bosons	baryon–antibaryon (a)symmetry	$\pi^0 \to \gamma\gamma$, η' mass

Table 2.2: Overview on chiral symmetry in QCD and its realization in nature.

symmetry in QCD is given in tab. 2.2.

2.3 Asymptotic freedom and confinement

The QCD Lagrangian contains $N_f + 1$ parameters, the bare coupling g_0 and the bare masses m_i with $i \in \{1, \ldots, N_f\}$. In order to get sensible results from any QCD calculation, the theory needs to be renormalized. This necessitates the introduction of a dimensionful renormalization scale μ, and involves the replacement of the bare parameters in favor of renormalized quantities, a renormalized coupling $\overline{g}(\mu)$ and renormalized masses $\overline{m}_i(\mu)$. Irrespective of the renormalization scheme, however, physical observables P do not depend on the unphysical renormalization scale μ. This is expressed in the renormalization group equation (RGE):

$$\left[\mu\, \partial_\mu + \beta(\overline{g})\, \partial_g + \overline{m}_i \tau(\overline{g})\, \partial_{\overline{m}_i}\right] P(\mu, \overline{g}, \overline{m}_i) = 0 \; . \tag{2.30}$$

The β and τ functions are logarithmic derivatives of the renormalized coupling and the renormalized masses, respectively:

$$\beta(\overline{g}) = \mu\, \partial_\mu \overline{g}(\mu) \tag{2.31}$$

$$\overline{m}_i \tau(\overline{g}) = \mu\, \partial_\mu \overline{m}_i(\mu) \; . \tag{2.32}$$

Given their perturbative expansions,

$$\begin{aligned} \beta(\overline{g}) &= -\overline{g}^3\left(b_0 + \overline{g}^2 b_1 + \ldots\right) && (2.33)\\ \text{and}\quad \tau(\overline{g}) &= -\overline{g}^2\left(d_0 + \overline{g}^2 d_1 + \ldots\right)\ , && (2.34) \end{aligned}$$

the following special solutions of the RGE can be found:

$$\Lambda = \mu\ \left(b_0\overline{g}^2\right)^{-\frac{b_1}{2b_0^2}} e^{-\frac{1}{2b_0\overline{g}^2}} \exp\left\{-\int_0^{\overline{g}} dx\left[\frac{1}{\beta(x)} + \frac{1}{b_0 x^3} - \frac{b_1}{b_0^2 x}\right]\right\} \tag{2.35}$$

$$M_i = \overline{m}_i\ \left(2b_0\overline{g}^2\right)^{-\frac{d_0}{2b_0}} \exp\left\{-\int_0^{\overline{g}} dx\left[\frac{\tau(x)}{\beta(x)} - \frac{d_0}{b_0 x}\right]\right\}\ . \tag{2.36}$$

These are called *renormalization group invariants (RGI)*. They are independent of μ, since a variation of μ is compensated by simultaneous variations of $\overline{g}(\mu)$ and $\overline{m}_i(\mu)$. For this reason, they can be considered the fundamental parameters of QCD. While the RGI masses M_i are scheme-independent,

$$M_i' = M_i\ , \tag{2.37}$$

the Λ parameter in two different renormalization schemes,

$$\Lambda' = \Lambda \exp\left\{\frac{c^{(1)}_{\alpha_s}}{2b_0}\right\}\ , \tag{2.38}$$

is related by the one-loop coefficient of

$$\alpha_s'(\mu) = \alpha_s(\mu)\left(1 + c^{(1)}_{\alpha_s}\alpha_s(\mu) + \ldots\right)\ , \tag{2.39}$$

where $\alpha_s(\mu) = \overline{g}^2(\mu)/(4\pi)$. The perturbative expansion (2.33) for β used in (2.35) leads to the following running of the coupling for large energies[6]:

$$\overline{g}^2(\mu) \overset{\mu\to\infty}{=} \frac{1}{b_0 \log\left(\frac{\mu^2}{\Lambda^2}\right)} + \frac{b_1}{b_0^3}\frac{\log\left(\log\left(\frac{\mu^2}{\Lambda^2}\right)\right)}{\left[\log\left(\frac{\mu^2}{\Lambda^2}\right)\right]^2} + \ldots \tag{2.40}$$

---

[6] We here use the terms energy and renormalization scale interchangeably, since for the description of high-energy (scattering) processes, μ is usually set to the typical energy scale of the system in order to ensure a fast convergence of the perturbative expansion.

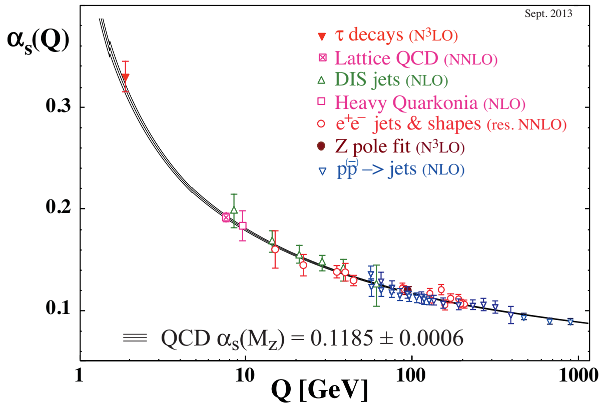


Figure 2.3: Measurements of $\alpha_s(\mu)$ in the $\overline{\text{MS}}$ scheme at different energy scales $\mu = Q$. The curves represent the QCD prediction with the combined world average value of $\alpha_s(M_Z)$ as input. The figure is taken from [26].

Note that this result can also be derived straight forwardly from higher order corrections to the coupling in perturbative QCD, see e.g. [28]. Remarkably, the lowest order coefficients,

$$b_0 = \frac{1}{(4\pi)^2}\left[\frac{11}{3}N_c - \frac{2}{3}N_f\right] \tag{2.41}$$

$$b_1 = \frac{1}{(4\pi)^4}\left[\frac{34}{3}N_c^2 - \left(\frac{13}{3}N_c - \frac{1}{N}\right)N_f\right] \tag{2.42}$$

$$d_0 = \frac{1}{(4\pi)^2}\,\frac{3\left(N_c^2 - 1\right)}{N_c}\,, \tag{2.43}$$

are scheme-independent. In particular, $N_c = 3$ and $N_f = 6$ imply $b_0, b_1 > 0$, and hence $\beta(\bar{g}) < 0$ (cf. (2.33)), or equivalently (cf. (2.40))

$$\bar{g}^2(\mu) \to 0 \quad \text{for} \quad \mu \to \infty\,. \tag{2.44}$$

This is the mathematical expression of the important phenomenon called *asymptotic freedom*, which was first discovered in [29, 30]. It states that strongly interacting particles are free in the limit of very high energies. Note that this behavior is strongly based on the self-interaction of gluons, cf. fig. 2.1, which appear in form of the N_c-dependent terms in b_0, b_1. It can be understood in the way that the amplification of the effective coupling at small distance by vacuum polarizations of quark-antiquark pairs, known already from QED, is overcompensated by polarizations of virtual gluon pairs. Asymptotic freedom is the foundation of the applicability of perturbation theory for the description of high energy QCD processes, studied e.g. at the Large Hadron Collider.

While the running of the coupling $\alpha_s(\mu)$ with the energy is determined in QCD by

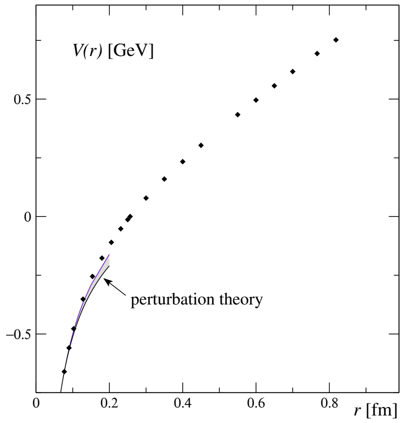


Figure 2.4: Static potential $V(r)$ as a function of the quark-antiquark distance r as computed in [32], converted into physical units by setting the Sommer scale [13] to $r_0 = 0.5$ fm. The figure is taken from [33].

the beta function, its absolute value needs to be fixed by experiment. Measurements of various kinds at different scales can be evolved to and compared at a reference scale, which is usually taken to be the Z boson mass. This serves as a highly non-trivial test of QCD, see fig. 2.3.

At low energies, the renormalization group predicts that the renormalized coupling gets large. This provides an indication of why colored states like quarks or gluons are not observed as free particles. Obviously, at a distance of roughly the diameter of a hadron, a very strong force effectively binds them into color-neutral hadrons, cf. (2.1, 2.2). This phenomenon is known as *confinement* and cannot be quantitatively described with perturbative methods as the large coupling renders these useless in the low energy sector. It requires non-perturbative approaches like Lattice QCD, which we will address in the next section. Although confinement is not yet resolved definitely, computations in Lattice QCD lead to a linearly rising term[7] in the quark-antiquark potential,

$$V(r) = A + \frac{B}{r} + \sigma r \, , \tag{2.45}$$

see fig. 2.4, and therefore give a semi-quantitative understanding of it.

Altogether, the wealth of phenomena and the accuracy with which these can be described impressively confirm QCD as the fundamental theory of the strong interaction, and in the end the principle of gauge invariance that it relies on.

---

[7] This result may also be found analytically in the strong coupling limit, cf. e.g. [31].

3 Quantum chromodynamics on the lattice

Lattice QCD is a straight forward non-perturbative approach to QCD based on the *path integral formalism*, which goes back to the work of Kenneth G. Wilson [34]. In the continuum, the correlation function of an arbitrary combination O of operators is formally given by

$$\langle O \rangle = \frac{1}{Z} \int \mathcal{D}[\psi, \overline{\psi}, A] \; e^{iS_{QCD}[\psi,\overline{\psi},A]} \; O[\psi, \overline{\psi}, A] \; . \tag{3.1}$$

Here, $S_{QCD} = \int d^4x \; \mathcal{L}_{QCD}$ is the action associated with the QCD Lagrangian (2.19), O on the right hand side is a functional of the fields which corresponds to the operator O on the left hand side, the measure $\mathcal{D}[\psi, \overline{\psi}, A] \equiv \mathcal{D}[\psi]\mathcal{D}[\overline{\psi}]\mathcal{D}[A]$ represents an infinite-dimensional integration over all possible field configurations in spacetime, and $Z = \int \mathcal{D}[\psi, \overline{\psi}, A] \; e^{iS_{QCD}[\psi,\overline{\psi},A]}$ is the partition function of QCD.

The functional integral (3.1) as such is neither well-defined nor accessible to numerical evaluation. However, it can be regularized by introduction of a discretized spacetime lattice, where the measure becomes a finite product over the lattice sites. The details of this procedure will be discussed in the upcoming sections 3.1-3.4. Assuming we have a finite functional integral, the Boltzmann factor in (3.1) may still wildly oscillate as the fields are varied. A remedy for this is obtained by performing a *Wick rotation*, i.e. switching from real time t to imaginary time τ,

$$t = -i\tau \; , \tag{3.2}$$

and corresponding rotation of the contour of integration in the path integral. This is equivalent to a transition from Minkowski to Euclidean space,

$$-ds^2 = -dt^2 + dx_1^2 + dx_2^2 + dx_3^2 = d\tau^2 + dx_1^2 + dx_2^2 + dx_3^2 \; , \tag{3.3}$$

where the problem of the oscillating Boltzmann factor in (3.1) is cured as one gets rid

	Minkowski	$\rightarrow$	Euclidean	Ref.
metric	$g^{\mu\nu} = (1,-1,-1,-1)$	$\rightarrow$	$\delta^{\mu\nu} = (1,1,1,1)$	app. A.2
coordinates	$x = (t, \vec{x})$	$\rightarrow$	$x^E = (-i\tau, \vec{x})$	(3.2)
derivative	(∂_0, ∂_j)	$\rightarrow$	$(i\partial_0^E, \partial_j^E)$	(3.6)
gluon field	(A_0, A_j)	$\rightarrow$	(iA_0^E, A_j^E)	(3.7)
cov. derivative	(D_0, D_j)	$\rightarrow$	(iD_0^E, D_j^E)	
γ matrices	(γ_0, γ_j)	$\rightarrow$	$(\gamma_0^E, -i\gamma_j^E)$	(3.8)

Table 3.1: Transition from Minkowski space to Euclidean space.

of the imaginary factor:

$$iS_{QCD} = i \int d^4x \; \mathcal{L}_{QCD}(t) = - \int d^4x^E \; \mathcal{L}_{QCD}^E(\tau) = -S_{QCD}^E \, , \tag{3.4}$$

with $d^4x^E \equiv d^3x d\tau$. The Euclidean QCD Lagrangian defined by (3.4),

$$\mathcal{L}_{QCD}^E = -\mathcal{L}_{QCD} \, , \tag{3.5}$$

can be easily derived from its Minkowski counterpart (cf. (2.14, 2.15, 2.19)). Under consideration that the temporal derivatives translate to Euclidean space according to

$$\partial_0 = \frac{\partial}{\partial t} = i\frac{\partial}{\partial \tau} = i\partial_0^E \, , \tag{3.6}$$

it is useful to apply two reparametrizations in order to obtain $\mathcal{L}_{QCD}^E$ in a compact form that is reminiscent of its Minkowski counterpart. First, in order to adapt the gluon fields to the above derivatives, one introduces Euclidean gluon fields defined by

$$A_0 = iA_0^E \qquad \text{and} \qquad A_j = A_j^E \, , \qquad j = 1,2,3 \, . \tag{3.7}$$

Second, with regard to the fermionic part of the action, Euclidean gamma matrices γ_μ^E are introduced, see app. A.2, which are related to the Minkowski gamma matrices by

$$\gamma_0 = \gamma_0^E \qquad \text{and} \qquad \gamma_j = -i\gamma_j^E \, , \quad j = 1,2,3 \, . \tag{3.8}$$

An overview on the transition from Minkowski space to Euclidean space is given in tab. 3.1.

With regard to the lattice discretization of the Euclidean action, it turns out to be convenient in addition to absorb the bare coupling in the gluon fields,

$$g_0 A_\mu^E \to A_\mu^E \,, \tag{3.9}$$

such that the covariant derivative (cf. (2.10)) and the field strength tensor (cf. (2.16)) read

$$D_\mu^E = \partial_\mu^E + iA_\mu^E \qquad \text{and} \qquad F_{\mu\nu}^E = -i[D_\mu^E(x), D_\nu^E(x)] \,. \tag{3.10}$$

With these ingredients, (3.5)-(3.10), the fermionic part of the *Euclidean QCD action* becomes

$$S_F^E[\psi, \overline{\psi}, A] = \int d^4x^E \; \overline{\psi}(x) \left(\gamma_\mu^E D_\mu^E + \mathcal{M} \right) \psi(x) \,, \tag{3.11}$$

whereas the gauge part can be written as

$$S_G^E[A] = \frac{1}{2g_0^2} \int d^4x^E \; \mathrm{tr} \left[F_{\mu\nu}^E F_{\mu\nu}^E \right] \,. \tag{3.12}$$

Note that from here on, the superscript "*E*" will be dropped everywhere as we will exclusively work in Euclidean space.

In this work, the theory with four dynamical quark flavors is usually considered, where the bare mass matrix (2.5) acquires the form

$$\mathcal{M} = \mathrm{diag}(m_{0,u}, m_{0,d}, m_{0,s}, m_{0,c}) \,. \tag{3.13}$$

We will differentiate between *degenerate* ($m_{0,u} = m_{0,d} = m_{0,s} = m_{0,c}$) and *non-degenerate* ($m_{0,u} \neq m_{0,d} \neq m_{0,s} \neq m_{0,c}$) masses. *While for the former case the notation* $N_f = 4$ *is well-established, we introduce the notation* $N_f = 4^*$ *for the general latter case here, to which all of the following will correspond to, unless noted otherwise.*

3.1 Naive fermion action

Lattice QCD is formulated on discretized Euclidean spacetime, which in our case will be a hypercubic lattice

$$X = \{ x_\mu = a n_\mu \mid \mu = 0, \ldots, 3 \; ; \; n_\mu \in \mathbb{Z} \} \,, \tag{3.14}$$

where a is the lattice spacing. The discretization of the fermion spinors $\psi, \overline{\psi}$ is straight forward, they are simply placed on the lattice sites. In order to write down the free fermion action, cf. (3.11) with $D_\mu \to \partial_\mu$, one simply needs to substitute the spacetime integral by the sum over the lattice sites,

$$\int d^4x \;\to\; a^4 \sum_{x\in X} \, , \tag{3.15}$$

and the derivative by a discretized symmetric expression (cf. app. A.3)

$$\partial_\mu \psi(x) \;\to\; \tilde{\partial}_\mu \psi(x) = \frac{1}{2a} \left[\psi(x+a\hat{\mu}) - \psi(x-a\hat{\mu}) \right] \, , \tag{3.16}$$

where $\hat{\mu}$ denotes a unit vector in direction μ. The free fermion action on the lattice thus becomes

$$\begin{aligned} S_F^{\text{free}}[\psi,\overline{\psi}] &= a^4 \sum_x \overline{\psi}(x) \left(\sum_\mu \gamma_\mu \tilde{\partial}_\mu + \mathcal{M} \right) \psi(x) \\ &= a^4 \sum_x \overline{\psi}(x) \left(\sum_\mu \gamma_\mu \frac{\psi(x+a\hat{\mu}) - \psi(x-a\hat{\mu})}{2a} + \mathcal{M}\psi(x) \right) \, . \end{aligned} \tag{3.17}$$

Note that here, and from now on, we use the abbreviations $\sum_x \equiv \sum_{x\in X}$ and $\sum_\mu \equiv \sum_{\mu=0}^3$.

The gauge fields are introduced like in the continuum, based on the principle that the fermion action shall be invariant under local gauge transformations $\Omega(x) \in SU(3)$, where we now have $x \in X$. The bilinear $\overline{\psi}(x)\psi(x+a\hat{\mu})$ as it appears in (3.17), neglecting γ_μ for the moment as it is trivial in color space, transforms like (cf. (2.6, 2.7))

$$\overline{\psi}(x)\psi(x+a\hat{\mu}) \;\to\; \overline{\psi}(x)\,\Omega(x)^\dagger\,\Omega(x+a\hat{\mu})\,\psi(x+a\hat{\mu}) \, . \tag{3.18}$$

If the fermion fields couple to a directed field $U_\mu(x)$, which transforms as

$$U_\mu(x) \to \Omega(x) U_\mu(x) \Omega^\dagger(x+a\hat{\mu}) \, , \tag{3.19}$$

the expression

$$\overline{\psi}(x) U_\mu(x) \psi(x+a\hat{\mu}) \tag{3.20}$$

is gauge invariant. The same holds for the combination $\overline{\psi}(x) U_\mu(x-a\hat{\mu})^\dagger \psi(x-a\hat{\mu})$.

Hence, introducing the convenient notation

$$U_{-\mu}(x) \equiv U_\mu(x - a\hat{\mu})^\dagger , \tag{3.21}$$

the gauge invariant fermion action reads

$$S_F[\psi, \overline{\psi}, U] = a^4 \sum_x \overline{\psi}(x) \left(\sum_\mu \gamma_\mu \frac{U_\mu(x)\psi(x + a\hat{\mu}) - U_{-\mu}(x)\psi(x - a\hat{\mu})}{2a} + \mathcal{M}\psi(x) \right) . \tag{3.22}$$

In terms of the Dirac operator

$$D = \sum_\mu \{\gamma_\mu \tilde{\nabla}_\mu\} = \frac{1}{2} \sum_\mu \left\{ \gamma_\mu \left(\nabla^*_\mu + \nabla_\mu \right) \right\} , \tag{3.23}$$

which employs a symmetric combination $\tilde{\nabla}_\mu$ of the covariant forward and backward derivatives (see also app. A.3),

$$\nabla_\mu \psi = \frac{1}{a} \left[U_\mu(x)\ \psi(x + a\hat{\mu}) - \psi(x) \right] \tag{3.24a}$$

$$\nabla^*_\mu \psi = \frac{1}{a} \left[\psi(x) - U_{-\mu}(x)\ \psi(x - a\hat{\mu}) \right] , \tag{3.24b}$$

it can be written in a very compact way:

$$S_F[\psi, \overline{\psi}, U] = a^4 \sum_x \overline{\psi}(x) \left(D + \mathcal{M} \right) \psi(x) . \tag{3.25}$$

This so-called *naive fermion action* approximates the continuum fermion action (3.11) with leading order discretization errors of only $O(a^2)$, due to the use of the symmetric (covariant) derivative. It is the basis for more sophisticated forms of lattice fermion actions like the Wilson fermion action, which we will address in sec. 3.3.

The nature of the lattice gauge fields U_μ differs from the continuum case. From (3.19), one immediately infers that $U_\mu(x)$ must be an element of the gauge group $SU(3)$, unlike the continuum counterparts $A_\mu(x)$, which belong to the algebra $\mathfrak{su}(3)$. The two are related by

$$U_\mu(x) = \exp\left(iaA_\mu(x) \right) , \tag{3.26}$$

where $aA_\mu(x)$ can be considered a lattice approximation of the continuum integral of

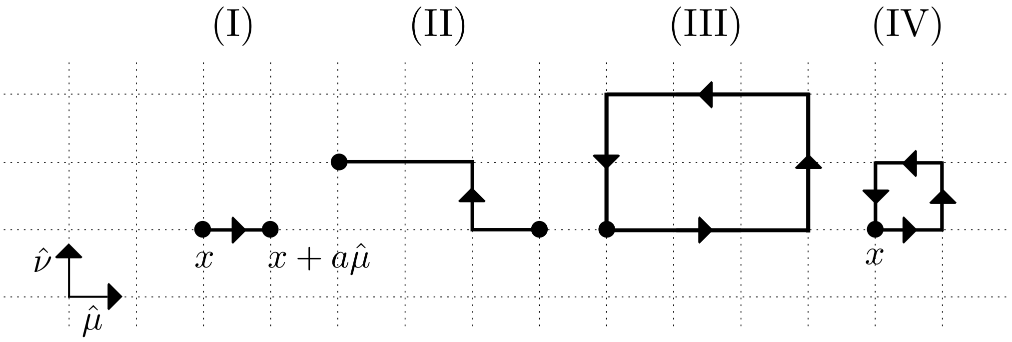


Figure 3.1: Illustration of different lattice gauge field quantities. The dashed lines represent the $\mu\nu$ hyperplane of a lattice with lattice spacing a. The arrows show the orientation of the link variables, and the dots stand for x_{start} or x_{end}. *(I)* The *link variable* $U_\mu(x)$, (3.26). *(II)* Product $P[U]$ of link variables along a general path $\mathcal{P}$, (3.27). *(III)* Product $P_{\text{cl}}[U]$ of link variables along a closed loop $\mathcal{P}_{\text{cl}}$, (3.29). The illustrated rectangular shape in two dimensions is called *Wilson loop*. *(IV) Plaquette variable* $U_{\mu\nu}(x)$ as the shortest, non-trivial, closed loop of link variables on the lattice, (3.30).

$A_\mu(x)$ along the link from x and $x + \mu$. Accordingly, $U_\mu(x)$ is called link variable[1], see fig. 3.1. Evidently, (3.26) employed in the naive fermion action (3.22) yields the continuum fermion action (3.11) in the limit $a \to 0$, as stated above.

3.2 Wilson gauge action

We now turn to the discussion of the *lattice gauge action*. Any such action has to be build up from gauge invariant combinations of link variables. We note that the product

$$P[U] = \prod_{(x,\mu)\in\mathcal{P}} U_\mu(x) \tag{3.27}$$

of link variables along any path $\mathcal{P}$ from the lattice site x_{start} to the lattice site x_{end} transforms as

$$P[U] \to \Omega(x_{\text{start}})\; P[U]\; \Omega^\dagger(x_{\text{end}})\;. \tag{3.28}$$

This is due to the fact that if each link is transformed according to (3.19), the transformation matrix $\Omega^\dagger$ at the end point of one link cancels with the transformation matrix Ω at the starting point of the next link on the path (cf. fig. 3.1). For the special case

---

[1] or gauge transporter

of closed loops $\mathcal{P}_{\text{cl}}$, where $x_{\text{start}} = x_{\text{end}}$, one may hence construct a gauge invariant product of link variables by taking the overall trace tr in color space:

$$\operatorname{tr}\Big[P_{\text{cl}}[U]\Big] \;\to\; \operatorname{tr}\Big[P_{\text{cl}}[U]\Big] \,. \tag{3.29}$$

The shortest, non-trivial, closed loop is the *plaquette*, also depicted in fig. 3.1. The corresponding plaquette variable is defined as[2]

$$U_{\mu\nu}(x) = U_\mu(x)\; U_\nu(x + a\hat{\mu})\; U_\mu(x + a\hat{\nu})^\dagger\; U_\nu(x)^\dagger \,. \tag{3.30}$$

In order to construct a lattice gauge action from $\operatorname{tr}[U_{\mu\nu}(x)]$ that turns into the continuum gauge action (3.12) in the naive continuum limit $a \to 0$, i.e. whose leading term in the lattice spacing is

$$\frac{a^4}{2g_0^2} \sum_x \sum_{\mu,\nu} \operatorname{tr}\left[F_{\mu\nu} F_{\mu\nu}\right] \,, \tag{3.31}$$

we need to express $U_{\mu\nu}(x)$ in terms of the continuum gauge fields A_μ. This can be done by using (3.26) for each of the four link variables, combination of the exponential functions by iterative application of the Baker–Campbell–Hausdorff formula,

$$\exp(A)\exp(B) = \exp(A + B + \frac{1}{2}[A, B] + \ldots) \,, \tag{3.32}$$

and Taylor expansion of the continuum gauge fields A_μ around the spacetime point x. One finds

$$U_{\mu\nu}(x) = \exp\left(ia^2 F_{\mu\nu}(x) + O(a^3)\right) = \mathbb{1} + ia^2 F_{\mu\nu}(x) + O(a^3) \,. \tag{3.33}$$

Note that $\operatorname{tr}[U_{\mu\nu}(x)]$ already contains the $O(a^4)$ term of (3.31) we are seeking for. The task now is to alter this expression in such a way that the contributions up to and including $O(a^3)$ vanish. While $\operatorname{tr}[U_{\mu\nu}(x)] \to \operatorname{tr}[\mathbb{1} - U_{\mu\nu}(x)]$ makes the constant disappear, the other terms vanish if the real part of the trace is taken. Indeed, from

$$\operatorname{tr}[U_{\mu\nu}(x)]^\star = \operatorname{tr}[U_{\mu\nu}^\dagger(x)] = \operatorname{tr}[U_{\nu\mu}(x)] \tag{3.34}$$

---

[2] Note that e.g. $U_\nu(x)^\dagger = U_{-\nu}(x + a\hat{\nu})$, see (3.21).

follows

$$\begin{aligned}
\mathrm{Re}\,\mathrm{tr}[\mathbb{1} - U_{\mu\nu}(x)] &= \frac{1}{2}\left(\mathrm{tr}[\mathbb{1} - U_{\mu\nu}(x)] + \mathrm{tr}[\mathbb{1} - U_{\mu\nu}(x)]^{\star}\right) \\
&\overset{(3.34)}{=} \frac{1}{2}\left(\mathrm{tr}[\mathbb{1} - U_{\mu\nu}(x)] + \mathrm{tr}[\mathbb{1} - U_{\nu\mu}(x)]\right) \\
&= \frac{ia^2}{2}\mathrm{tr}[F_{\mu\nu}(x) + F_{\nu\mu}(x)] + O(a^3)\ , \qquad (3.35)
\end{aligned}$$

where the leading $O(a^2)$ term vanishes due to the antisymmetry of the field strength tensor. The $O(a^3)$ terms cancel each other similarly, whereas the $O(a^4)$ term remains unaltered by the restriction to the real part. Hence, the so-called *Wilson gauge action*

$$S_G^{\mathrm{W}}[U] = \frac{2}{g_0^2}\sum_x \sum_{\mu<\nu} \mathrm{Re}\,\mathrm{tr}[\mathbb{1} - U_{\mu\nu}(x)]\ . \qquad (3.36)$$

fulfills the requirement to lead to (3.31) in the limit $a \to 0$. It is customary to use

$$2\sum_{\mu<\nu} \mathrm{Re}\,\mathrm{tr}\big[U_{\mu\nu}(x)\big] \overset{(3.34)}{=} \sum_{\mu,\nu} \mathrm{tr}\big[U_{\mu\nu}(x)\big] \qquad (3.37)$$

to replace the sum $\sum_{\mu<\nu}$ over all *unoriented plaquettes* in (3.36) by the sum $\sum_{\mu,\nu}$ over all *oriented plaquettes*,

$$S_G^{\mathrm{W}}[U] = \frac{1}{g_0^2}\sum_x \sum_{\mu,\nu} \mathrm{tr}[\mathbb{1} - U_{\mu\nu}(x)]\ . \qquad (3.38)$$

Furthermore, if the set of all *oriented plaquettes* is denoted by $\mathcal{S}_0$, the Wilson gauge action (3.38) may be written under use of $\sum_x \sum_{\mu,\nu} U_{\mu\nu}(x) \equiv \sum_{\mathcal{C}\in\mathcal{S}_0} U(\mathcal{C})$ as

$$S_G^{\mathrm{W}}[U] = \frac{1}{g_0^2}\sum_{\mathcal{C}\in\mathcal{S}_0} \mathrm{tr}[\mathbb{1} - U(\mathcal{C})]\ . \qquad (3.39)$$

In the literature, both formulations are widely used, in terms of *unoriented plaquettes*, (3.36), or *oriented plaquettes*, (3.38) and (3.39).

From further terms in the expansion of the exponential in (3.33), we conclude that the discretization error of (3.36, 3.38, 3.39) is—as for the naive fermion action (3.25)—of the order $O(a^2)$, i.e. additional terms come with a factor a^6. In sec. 3.8, we will encounter another lattice gauge action, whose discretization error is further reduced.

3.3 Wilson fermion action

We now further pursue the discussion of fermions on the lattice, starting from the naive fermion action obtained in sec. 3.1,

$$S_F[\psi,\overline{\psi},U] = a^4 \sum_x \overline{\psi}(x)\,(D+\mathcal{M})\,\psi(x)\;, \tag{3.25}$$

$$\text{with}^3 \qquad D = \frac{1}{2}\sum_\mu \left\{\gamma_\mu\left(\nabla_\mu^* + \nabla_\mu\right)\right\}\;. \tag{3.23}$$

The action as such suffers from the so-called *fermion doubling problem*. To explain this, it is sufficient to consider the free naive fermion action of one quark flavor with mass m_0, i.e. we set $\tilde{\nabla}_\mu \to \tilde{\partial}_\mu$ and $\mathcal{M} \to m_0$. We consider the free fermion propagator in momentum space, which can be found by taking the inverse of the Fourier transformed Dirac operator. It reads

$$\tilde{D}(p)^{-1} = \left(m_0 + ia^{-1}\sum_\mu \gamma_\mu \sin\left(ap_\mu\right)\right)^{-1} = \frac{m_0 - ia^{-1}\sum_\mu \gamma_\mu \sin\left(ap_\mu\right)}{m_0^2 + a^{-2}\sum_\mu \sin^2\left(ap_\mu\right)}\;, \tag{3.40}$$

where the momenta ap_μ may assume $2\pi/L$ equidistant values in the range $-\pi \leq ap_\mu \leq \pi$. While for $a \to 0$, this propagator correctly reproduces the physical pole, there are 15 additional singularities. For $m_0 = 0$, these unphysical so-called *doublers* are found when $ap_\mu \in \{0,\pi\}\ \forall\mu$.

A possible solution to get rid of the doublers is to add an extra Wilson term to the momentum space Dirac operator, cf. (3.40):

$$\tilde{D}_W(p) = m_0 + ia^{-1}\sum_\mu \gamma_\mu \sin\left(ap_\mu\right) + a^{-1}\sum_\mu \left(1-\cos(ap_\mu)\right)\;. \tag{3.41}$$

This term effectively constitutes an additional mass term,

$$m_0 \to m_0 + \frac{2l}{a}\;, \tag{3.42}$$

where l is the number of momentum components with $ap_\mu = \pi$. Hence, while the physical pole remains the same, the doublers decouple from the theory for $a \to 0$ as they become infinitely heavy. The massless *Wilson–Dirac operator*, which incorporates

---

[3] Note that throughout this work, if equations are being reused, they are labelled according to their first appearance.

the Wilson term, can be written as (cf. (3.23))

$$D_W = \frac{1}{2}\sum_{\mu}\left\{\gamma_\mu\left(\nabla^*_\mu + \nabla_\mu\right) - a\nabla^*_\mu\nabla_\mu\right\}\ , \tag{3.43}$$

where the interaction has been switched on again now. The corresponding *Wilson fermion action* S_F^{W} for $N_f = 4^*$ then assumes a form very similar to (3.25),

$$S_F^{\mathrm{W}}[\psi, \overline{\psi}, U] = a^4 \sum_x \overline{\psi}(x)\left(D_W + \mathcal{M}\right)\psi(x)\ . \tag{3.44}$$

It respects the global symmetries C,P and T. Moreover, the Wilson–Dirac operator is γ_5-hermitean,

$$(\gamma_5 D_W)^\dagger = \gamma_5 D_W\ , \tag{3.45}$$

which will be important for its use in Monte Carlo simulations, to be discussed in sec. 4.4.

However, there are also two prices one has to pay for the removal of the fermion doublers. First, it is obvious from (3.43) that the extra Wilson term induces cutoff effects of the order $O(a)$, which compared to the naive fermion action (3.25) is a change for the worse. The $O(a^2)$ scaling to the continuum can be restored with the help of the Symanzik improvement program, whose basic idea is described in sec. 3.7, and whose application to the case of $N_f = 4^*$ non-degenerate, massive Wilson fermions is the main topic of this thesis, see sec. 8 and 9.

Second, the Wilson–Dirac operator explicitly breaks chiral symmetry on the lattice, because the Wilson term is a $\mathbb{1}$ in Dirac space and hence does *not* anticommute with γ_5, cf. (2.25). This is in accordance with the Nielsen–Ninomiya theorem [35–37], which states that it is impossible to both conserve chiral symmetry on the lattice and remove the doublers. The breaking of chiral symmetry by Wilson fermions has far-reaching consequences. They concern the quark mass renormalization as well as the consideration of dynamical quark effects in the numerical computation of Euclidean correlation functions, as we will see in detail in sec. 3.5 and sec. 4, respectively.

3.4 From Euclidean correlation functions to physical observables

Having discussed the fermion and gauge fields on the lattice, we may now return to the Euclidean correlation functions in the path integral formalism. The lattice equivalent to the continuum expression (3.1) is

$$\langle O \rangle = \frac{1}{Z} \int \mathcal{D}[\psi, \overline{\psi}, U] \; e^{-S[\psi,\overline{\psi},U]} O[\psi, \overline{\psi}, U] \; , \tag{3.46}$$

where

$$Z = \int \mathcal{D}[\psi, \overline{\psi}, U] \; e^{-S[\psi,\overline{\psi},U]} \tag{3.47}$$

is again the partition function. $S[\psi, \overline{\psi}, U]$ summarizes the Wilson gauge (3.36) and Wilson fermion (3.44) action,

$$S[\psi, \overline{\psi}, U] \equiv S^{\mathrm{W}}[\psi, \overline{\psi}, U] = S_G^{\mathrm{W}}[U] + S_F^{\mathrm{W}}[\psi, \overline{\psi}, U] \; , \tag{3.48}$$

although the following holds for any valid discretized action. The measures $\mathcal{D}[\psi, \overline{\psi}, U] = \mathcal{D}[\psi, \overline{\psi}]\mathcal{D}[U]$ are now *finite* products of the fermion and link variables:

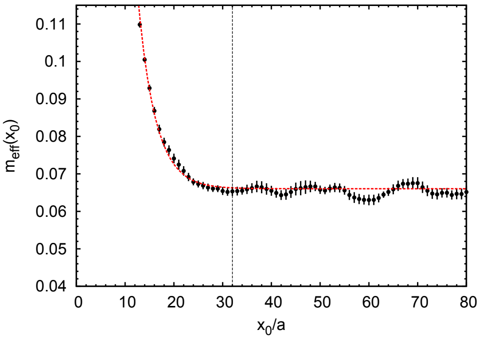
$$\mathcal{D}[\psi, \overline{\psi}] = \prod_x \prod_{f,\alpha,c} d\psi^{(f)}(x)_c^\alpha \; d\overline{\psi}^{(f)}(x)_c^\alpha \tag{3.49}$$

$$\mathcal{D}[U] = \prod_x \prod_\mu dU_\mu(x) \; . \tag{3.50}$$

The gauge invariant $(dU_\mu(x) = d(\Omega(x) \; U_\mu(x) \; \Omega(x+\hat{\mu})))$ and normalized $(\int dU_\mu \; 1 = 1)$ so-called Haar measure $dU_\mu(x)$ integrates the link variables over the whole compact group manifold $SU(3)$. One usually works with gauge invariant observables $O[\psi, \overline{\psi}, U]$, so that also $\langle O \rangle$ is gauge invariant.

Correlation functions like (3.46) can be numerically evaluated within the framework of Monte Carlo simulations, which will be explained in detail in sec. 4. Under use of the canonical operator formalism, the results may be used to extract physical observables like the energy spectrum of the theory. As hadron masses are conceptually the simplest quantities to obtain from lattice computations, we concentrate on those for the moment. If O_H and $\overline{O}_H$ are interpolators which correspond to the Hilbert space operators $\hat{O}_H$ and $\hat{O}_H^\dagger$ that are able to annihilate or create some hadron, the spectral

Figure 3.2: Extraction of the pion mass $E_\pi = m_\pi$ from an $N_f = 2+1$ simulation on a 128×64^3 lattice carried out in the course of the CLS effort [38]. The excited state contributions at small x_0/a are taken account here by the fit shown in red. m_{eff} is given in lattice units (a set to unity). The figure is taken from [38].



decomposition of the propagator of these interpolators reads[4]

$$C(x_0) = \left\langle O_H(x_0)\, \overline{O}_H(0) \right\rangle = \sum_k \langle 0|\hat{O}_H|k\rangle \langle k|\hat{O}_H^\dagger|0\rangle\; e^{-x_0 E_k} \tag{3.51a}$$

$$\sim e^{-x_0 E_H} \left(1 + O(e^{-x_0 \Delta E})\right) , \tag{3.51b}$$

where E_H is the energy of the considered hadron, and ΔE denotes the energy gap to the first excited state with the same quantum numbers. From the behavior for large Euclidean time $x_0 = n_0 \cdot a$, where $n_0 \in [0, T/a]$ labels the time slice number, the energy of the ground state in units of the lattice spacing, aE_H, can be derived. For this purpose, one considers an effective mass[5] $am_{\text{eff}}(x_0) = \log\left(\frac{C(x_0)}{C(x_0+1)}\right)$, which reaches a plateau for sufficiently large x_0, when the excited state contributions have abated and the ground state dominates. An example is shown in fig. 3.2. The value of aE_H may be obtained from a fit to am_{eff}. It depends on the bare parameters of the theory,

$$aE_H \equiv aE_H(g_0, am_{0,u}, am_{0,s}, am_{0,c}) , \tag{3.52}$$

which are at this moment unspecified. In order to arrive at physical predictions E_H^{cont} for E_H in the continuum, the main steps to be taken are the following:

1. *Scale setting:* The lattice spacing a is determined. This allows to relate dimensionless numbers (like aE_H) to quantities in physical units (like E_H).
2. *Renormalization:* The bare parameters are eliminated in favor of physical observables under use of experimental input.

---

[4] We ignore technical aspects like momentum projection and smearing here.

[5] Note that the propagation in x_0 and $T - x_0$ is often identical up a possible minus sign, in which case m_{eff} is usually defined in terms of typical cosh or sinh functions.

3. *Continuum extrapolation:* The limit $a \to 0$ is taken, while the lattice volume is kept large enough for finite volume effects to be small.

We are going to discuss the individual steps in detail in the following. Although one principally follows the above order, they are intrinsically tied to each other and can often not be considered truly isolated in practice.

The basic method to *set the scale* at a given bare coupling g_0 is to use a phenomenological scale as input, say the proton mass m_p^{exp}, and identify with it the ground state energy $E_p = m_p$ obtained on the lattice by use of a proton interpolator in (3.51):

$$a = \frac{am_p}{m_p^{\text{exp}}} \; . \tag{3.53}$$

The proton mass here serves as a simple example. It is actually more common to use other phenomenological scales with more favorable properties (good statistical precision, small systematic uncertainties, weak quark mass dependence), like e.g. the Ω mass or pseudoscalar decay constants f_π, f_K. Also, the use of the Sommer scale r_0 or its variation r_1, defined in terms of the static quark-antiquark potential (cf. (2.45)), have been prominently employed. More recently, various gradient flow observables (to be discussed in sec. 6.1) have emerged, which stand out due to the combination of excellent statistical properties and the absence of excited state contributions with the associated systematic errors. An overview on the most commonly used scales is given in [39]. For simplicity, nevertheless, we will proceed assuming the scale was set with the proton mass.

For the (basic) *renormalization* of the theory, it requires additional phenomenological scales which together are sensitive[6] to *all* bare masses in the theory. Usually one employs pseudoscalar meson masses, for instance those of the pion, the K- and the D-meson. Depending on the bare coupling g_0, the bare masses $m_i = m_i(g_0)$ are adjusted such that the theory reproduces physical values for those observables. With the scale set by the proton mass, this can be achieved by tuning to physical mass ratios:

$$\frac{am_h}{am_p} = \frac{m_h^{\text{exp}}}{m_p^{\text{exp}}} \; , \quad h = \pi, K, D \; . \tag{3.54}$$

Once the bare masses $m_i(g_0)$ are fixed by this, the energy spectrum $(aE_H)(g_0)$ can be

---

[6] While all observables are intrinsically dependent on the sea quark masses (in case those are unquenched), by sensitive we here mean that the respective quark masses enter the observables via *valence quarks*. Moreover, the observables should be independent in the sense that they depend differently on the quark masses.

computed for any hadron $H \neq p, \pi, K, D$ at a given g_0. Subsequently, with the lattice spacing $a(g_0)$ known from scale setting, physical predictions for E_H can be made in the continuum limit,

$$\frac{E_H^{\text{cont}}}{m_p^{\text{exp}}} = \lim_{a \to 0} \frac{aE_H}{am_p}(g_0) \; , \tag{3.55}$$

providing a non-perturbative solution of the relativistic bound state problem. Since small quark masses are associated with increased numerical effort, as we will see in detail in sec. 4.7 and the end of the current section, one often simulates at unphysically large up and down (or equivalently pion) masses and extrapolates to the physical ones. This procedure, called *chiral extrapolation* as the physical point lies at (almost) vanishing masses, may be conducted simultaneously with the continuum extrapolation.

The *continuum limit* is reached for $g_0 \to 0$. This can be derived from the insight that a physical observable[7] $P = P(g_0, a)$ should be independent of a in the limit $a \to 0$. Consequently, a change in a needs to be compensated by a simultaneous change in $g_0(a)$, which may again (cf. (2.30)) be expressed in a renormalization group equation:

$$\left[a \, \partial_a - \beta(g_0) \, \partial_{g_0}\right] P(g_0, a) = 0 \; . \tag{3.56}$$

The beta function is defined here with a relative minus sign in comparison to (2.31),

$$\beta(g_0) = -a \, \partial_a g_0(a) \; . \tag{3.57}$$

The perturbative expansion of $\beta(g_0)$ around $g_0 = 0$ is completely equivalent to (2.33), and the special solution Λ can again be found, by formally replacing $\mu \to 1/a$ and $\overline{g}^2 \to g_0^2$ in (2.35):

$$\Lambda = \frac{1}{a} \left(b_0 g_0^2\right)^{-\frac{b_1}{2b_0^2}} e^{-\frac{1}{2b_0 g_0^2}} \left(1 + O(g_0^2)\right) \; . \tag{3.58}$$

It relates the lattice spacing a to the bare coupling g_0^2, and leads to the *asymptotic scaling*

$$g_0^2 \sim 1/\log\left(a\Lambda\right) \qquad \text{for} \qquad a \to 0 \; . \tag{3.59}$$

Note that the continuum limit $g_0 \to 0$ coincides with a critical point of the system, at

---

[7] We consider pure gauge theory here for brevity. The discussion may be extended to QCD, in which case the bare masses are dependent on the lattice spacing as well, $m_{0,i}(a)$ with $i \in \{u, d, s, c\}$.

which the correlation length $\xi = 1/(am_\pi)$ diverges. The pion governs the correlation length in its capacity as the lightest hadron in the theory. As such, it also rules the finite volume effects, which for an observable P are determined by the ratio of the lattice extent[8] L/a to the correlation length ξ:

$$P|_{V=\infty} - P|_{V=L^4} = O(e^{-m_\pi L}) \;, \tag{3.60}$$

where $(L/a)\,/\xi = m_\pi L$ was used. In principle, one has to take the *infinite volume limit* for observables as well. However, due to the exponential decrease of finite volume effects, (3.60), they can be neglected if $m_\pi L$ is sufficiently large. As a rule of thumb, this is the case for $m_\pi L \gtrsim 4$. Note that the smaller the pion mass, the larger the lattice volume has to be in order to fulfill this requirement. This provides another reason why simulations at low pion masses are expensive.

3.5 Renormalization of the bare parameters

The hadron masses E_H are the simplest quantities one may compute on the lattice. Other physical observables are derived from the matrix elements in (3.51a), or in general from Euclidean correlators with composite operators of various kinds. These require an additional renormalization on top of the previously encountered hadron spectrum renormalization. We postpone the discussion of this matter to the next section.

Meanwhile, we focus on the renormalization of the bare coupling and the bare masses. The necessity to consider this arises for various reasons. It allows, for instance, to make predictions for the strong coupling $\overline{g}^2(\mu)$ (cf. fig. 2.3) and the renormalized quark masses $\overline{m}_i(\mu)$, or equivalently the RGIs Λ [40–44] and M_i [45], see (2.35) and (2.36), respectively. This is done by computation of the renormalized quantity under use of low energy experimental input (scale setting, cf. sec. 3.4), and evolution to high energies, where it may be converted to a commonly used renormalization scheme in the continuum, like $\overline{\text{MS}}$. Moreover, the notion of a renormalized coupling or mass may be employed to keep the physics on the lattice fixed, e.g. in terms of a line of constant physics. We denote renormalized quantities on the lattice with a subscript R (in contrast to the overbar used in the continuum), and—until further notice—work exclusively in a *mass-independent renormalization scheme*, where the renormalization constants depend only on the bare coupling and the renormalization scale $a\mu$, $Z \equiv Z(g_0^2, a\mu)$.

---

[8] We assume a hypercubic lattice with $V = T \times L^3 = L^4$ here.

In principle, the renormalization of the bare parameters reads

$$g_R^2 = Z_g(g_0^2, a\mu)\, g_0^2 \tag{3.61}$$

$$m_\mathrm{R}^i = Z_m(g_0^2, a\mu)\, m_{0,i}\ , \tag{3.62}$$

where the definition of the renormalization constants Z and the scale μ depend on the scheme. In sec. 5, we will encounter the Schrödinger functional, in terms of which a renormalized coupling g_R^2 can be defined. The above principle mass renormalization prescription has to be altered in the Wilson formulation of Lattice QCD, due to the breaking of chiral symmetry, cf. sec. 3.3. As explained in detail in app. D and app. E, the bare mass matrix $\mathcal{M}$ (3.13) may be decomposed in a flavor singlet mass and flavor non-singlet mass components, which renormalize differently, with renormalization constants $Z_{m^0}(g_0^2, a\mu)$ and $Z_m(g_0^2, a\mu)$, respectively. Note, however, that their dependence on the renormalization scale μ is the same, such that the ratio

$$r_m(g_0^2) = \frac{Z_{m^0}(g_0^2, a\mu)}{Z_m(g_0^2, a\mu)} \tag{3.63}$$

is μ-independent. Furthermore, the singlet quark mass is subject to an additive renormalization, which manifests itself in the appearance of the so-called *critical mass* $m_\mathrm{crit} \equiv m_\mathrm{crit}(g_0^2)$. Emphasizing its dependence on g_0^2, one also speaks of the *critical line* in the space spanned by the bare mass and the bare coupling. In terms of the matrix

$$M = \mathrm{diag}(m_{q,u}, m_{q,d}, m_{q,s}, m_{q,c}) \tag{3.64}$$

of *bare subtracted quark masses*

$$m_{q,i} = m_{0,i} - m_\mathrm{crit}\ , \tag{3.65}$$

the quark mass renormalization in the generic case of $N_f = 4^*$ Wilson fermions can be written as[9]

$$m_\mathrm{R}^i = Z_m(g_0^2, a\mu)\left[m_{q,i} + \left(r_m(g_0^2) - 1\right) \mathrm{Tr}\left[M\right] / N_f\right]\ . \tag{3.66}$$

The derivation of this formula can be found in app. E.1. As a consequence, the renor-

---

[9] Note that in contrast to the trace tr in color space, see e.g. (3.29), the trace in flavor space is denoted by Tr.

malized quark mass m_{R}^{i} vanishes if *all* the bare subtracted quark masses become zero:

$$m_{q,j} = 0 \quad \forall\, j \in \{1, \ldots, N_f\} \quad \Rightarrow \quad m_{\mathrm{R}}^{i} = 0\,. \tag{3.67}$$

For the special case of $N_f = 4$ degenerate masses[10], (3.66) reduces to

$$m_{\mathrm{R}} = Z_{m^0}(g_0^2, a\mu)\; m_q\,. \tag{3.68}$$

We close this section with the introduction of the so-called *hopping parameters*, which are closely related to the bare (subtracted) quark masses. By explicitly writing out all the terms in the Wilson–Dirac operator (3.43), the Wilson fermion action (3.44) may be cast into the form

$$S_F^{\mathrm{W}}[\psi, \overline{\psi}, U] = a^4 \sum_{i=u,d,s,c} \sum_{x,y} \overline{\psi}^{(i)}(x)\; C_i \left(\mathbb{1} - \kappa_i H\right)\; \psi^{(i)}(y)\,, \tag{3.69a}$$

with

$$\begin{aligned} C_i &= m_{0,i} + \frac{4}{a} && \text{(3.69b)}\\ \kappa_i &= \frac{1}{2(am_{0,i} + 4)} && \text{(3.69c)}\\ H(x,y) &= \sum_{\mu=\pm 1}^{\pm 4} \left(\mathbb{1} - \gamma_\mu\right) U_\mu(x) \delta_{x+\hat{\mu},y}\,, && \text{(3.69d)} \end{aligned}$$

where in (3.69d) the convenient notation $\gamma_{-\mu} = -\gamma_\mu$ is used. The irrelevant constants C_i can be absorbed in a redefinition of the quark fields, $\psi^{(i)} \to \sqrt{C_i}\psi^{(i)}$ and $\overline{\psi}^{(i)} \to \overline{\psi}^{(i)}\sqrt{C_i}$. The term $H(x,y)$ is called *hopping matrix* as it incorporates all nearest neighbor terms. Accordingly, κ_i are the *hopping parameters*, which are often used instead of the bare masses $m_{0,i}$ in practice. The quark propagator D_{W}^{-1}, for instance, may be expanded in a series of powers of κ_i, the *hopping expansion*, which is useful especially for large masses. The critical mass m_{crit} defines the *critical hopping parameter* κ_{crit}, and the bare subtracted quark masses (3.65) may be written as

$$m_{q,i} = \frac{1}{2a}\left[\frac{1}{\kappa_i} - \frac{1}{\kappa_{\mathrm{crit}}}\right]\,. \tag{3.70}$$

---

[10] The flavor index may be dropped in that case.

3.6 Renormalization of bilinears and the PCAC mass

We now proceed our discussion with the renormalization of composite operators in the general case of $N_f = 4^*$ in a mass-independent scheme, as announced in the previous section. We concentrate on quark bilinears, namely the vector current V_μ, the scalar density S, the axial current A_μ, and the pseudoscalar density[11] P. These are generally of interest e.g. in flavor physics for the determination of hadronic matrix elements and decay constants. Of particular importance for our purposes will be the axial and pseudoscalar currents, and the PCAC relation that relates those, as will be discussed below. The currents are defined in terms of the fermion fields as

$$V_\mu^c(x) = \overline{\psi}(x)\,\gamma_\mu \quad T^c\,\psi(x) \tag{3.71}$$

$$S^c(x) = \overline{\psi}(x) \qquad T^c\,\psi(x) \tag{3.72}$$

$$A_\mu^c(x) = \overline{\psi}(x)\,\gamma_\mu\gamma_5 T^c\,\psi(x) \tag{3.73}$$

$$P^c(x) = \overline{\psi}(x) \quad \gamma_5 T^c\,\psi(x)\,. \tag{3.74}$$

Here, they are represented in the *generator basis* (cf. app. B), T^c extended by $T^0 = \mathbb{1}/N_f$. Correspondingly, the components are labelled by $c \in \{0, \ldots, N_f^2 - 1\}$. We will also frequently employ the bilinears in the *flavor basis* (cf. app. B), where the above currents $\phi = V_\mu, S, A_\mu, P$ with their respective elements $\Gamma = \gamma_\mu, \mathbb{1}, \gamma_\mu\gamma_5, \gamma_5$ of the Clifford Algebra read

$$\phi^{ij} \quad = \overline{\psi}^i(x) \quad \Gamma \quad \psi^j(x)\,, \tag{3.75}$$

with $i, j \in \{1, \ldots, N_f\}$.

In the framework of chiral Ward identities, which are subject of app. C, it can be argued that the pseudoscalar density and the axial current, of which we consider *off-diagonal* ($j \neq i$) components exclusively, renormalize according to

$$(A_R)_\mu^{ij} = Z_A(g_0^2) \qquad A_\mu^{ij} \tag{3.76}$$

$$(P_R)^{ij} = Z_P(g_0^2, a\mu)\,P^{ij}\,, \tag{3.77}$$

and the scale-independent renormalization of the axial current vanishes in the continuum limit, $Z_A(g_0^2) \to 1$ for $g_0^2 \to 0$. Moreover, the chiral Ward identities themselves state that the vector and axial current, V_μ and A_μ, are (partially) conserved. Of partic-

---

[11] We will also use the expression "pseudoscalar current".

ular importance for us will be the *PCAC relation*:

$$\partial_\mu A_\mu^{ij}(x) = 2m_{0,ij}\, P^{ij}(x)\ . \tag{3.78}$$

where $m_{0,ij} = \frac{1}{2}\left(m_{0,i} + m_{0,j}\right)$ is the average bare mass of the quark flavors i and j. Equation (3.78) is an operator relation, which is exact in the continuum. It may well be written in terms of an expectation value with an operator O,

$$\tilde{\partial}_\mu \left\langle A_\mu^{ij}(x)\ O \right\rangle = 2m_{0,ij} \left\langle P^{ij}(x)\ O \right\rangle\ . \tag{3.79}$$

where O is arbitrary apart from the restriction that its support shall not contain x. On the lattice (and after renormalization), the relation is violated due to the breaking of chiral symmetry in the Wilson formulation (cf. app. C):

$$\tilde{\partial}_\mu \left\langle (A_\mathrm{R})_\mu^{ij}(x)\ O_\mathrm{R} \right\rangle = (m_\mathrm{R}^i + m_\mathrm{R}^j) \left\langle P_\mathrm{R}^{ij}(x)\ O_\mathrm{R} \right\rangle + O(a)\ . \tag{C.13b}$$

Isolation of the masses on one side of the equation, and application of the respective renormalization prescriptions for the currents, (3.76) and (3.77), leads to

$$m_\mathrm{R}^{ij} = \frac{Z_A(g_0^2)}{Z_P(g_0^2, a\mu)} \frac{\tilde{\partial}_\mu \left\langle A_\mu^{ij}(x)\ O \right\rangle}{2\left\langle P^{ij}(x)\ O \right\rangle} + O(a)\ , \tag{3.80}$$

where $m_\mathrm{R}^{ij} = \frac{1}{2}\left(m_\mathrm{R}^i + m_\mathrm{R}^j\right)$ is the obvious renormalized counterpart of the average bare mass $m_{0,ij}$. Note that the renormalization of the operator O drops out in the ratio above. Equation (3.80) suggests the definition of the so-called *bare PCAC mass*:

$$m_\mathrm{PCAC}^{ij}(x) = \frac{\tilde{\partial}_\mu \left\langle A_\mu^{ij}(x)\ O \right\rangle}{2\left\langle P^{ij}(x)\ O \right\rangle}\ . \tag{3.81}$$

After introduction of an improved version $m_\mathrm{PCAC,I}^{ij}$ with reduced cutoff effects in the upcoming sec. 3.7, we will concretize the notion of the PCAC mass in the framework of the Schrödinger functional in sec. 5.3. The dependence of m_PCAC^{ij} on the spacetime coordinate x, the operator O, and any other parameter that enters the above correlation functions, is a pure lattice artifact. In connection with (3.80), this will play a central role in the determination of the improvement coefficient c_sw, to be discussed in sec. 9.

3.7 Symanzik improvement program in $N_f = 4$

As explained in sec. 3.4, the continuum extrapolation of results obtained at finite lattice spacing is crucial in order to make predictions suitable for comparison with experiment. Cutoff effects depend on the employed action as well as the observable under consideration and may differ substantially. In general, their leading order is $O(a)$, i.e. the continuum is approached linearly for sufficiently small lattice spacings. For a connected renormalized n-point correlation function, given in terms of a local gauge-invariant composite operator $\phi(x)$ as[12]

$$G_n(x_1, \, \dots \, , x_n) = \left(Z_\phi\right)^n \langle \phi(x_1) \, \dots \, \phi(x_n) \rangle \, , \tag{3.82}$$

this may formally be expressed by

$$G_n(x_1, \, \dots \, , x_n)\big|_{a \neq 0} = G_n(x_1, \, \dots \, , x_n)\big|_{a=0} + O(a) \, . \tag{3.83}$$

The size of the cutoff effects has impact on the statistical error of the continuum extrapolated value. Even more importantly, an uncertainty in their functional dependence on the lattice spacing introduces a systematic error which makes the extrapolation more difficult and in the worst case may impair its reliability. Hence, it is of utter importance to have the cutoff effects under control. Of considerable help in that sense is to reduce the overall cutoff effects to the order $O(a^2)$. The absence of linear order cutoff effects renders the extrapolation smoother and reduces its ambiguousness. This holds in particular as cutoff effects which come in powers of the lattice spacing larger than two, vanish fast for small (but affordable) lattice spacings. In turn, this reduces the need for expensive simulations at very small lattice spacings.

The discretization effects in (3.83) stem from two distinct sources. On the one hand, from the discretized actions in the path integral, on the other hand, from the discretized representations of the operators. Only if these two ingredients exhibit $O(a^2)$ cutoff effects *individually*, the ones of the correlation function in (3.83) are of this order, too. The *Symanzik improvement program* [7–9] is a well-established systematic approach to achieve this.

---

[12] We assume that operator mixing under renormalization is absent.

Symanzik effective theory and on-shell improvement

We discuss the basic idea behind the Symanzik improvement program in the following, restricting ourselves to the case of $N_f = 4$ *degenerate* quark flavors, for which the bare mass matrix $\mathcal{M}$ (cf. (3.13)) and the related bare subtracted mass matrix M (cf. (3.64)) are written as

$$\mathcal{M} = m_0 \cdot \mathbb{1}_4 \qquad \text{and} \qquad M = m_q \cdot \mathbb{1}_4 \,. \tag{3.84}$$

The extension to $N_f = 4^*$ *non-degenerate* quark flavors is subject of sec. 8. The lattice framework is defined by the Wilson gauge $S_G^{\mathrm{W}}[U]$ (3.36) and Wilson fermion action $S_F^{\mathrm{W}}[\psi, \overline{\psi}, U]$ (3.44), which are combined in $S[\psi, \overline{\psi}, U]$, as in (3.48). Symanzik improvement is based on the observation that—sufficiently close to the continuum—the lattice theory may be described in terms of the local so-called *Symanzik effective theory (SET)*, with the action

$$S_{\mathrm{eff}} = S_0 + aS_1 + a^2 S_2 + O(a^3) \,. \tag{3.85}$$

Here, S_0 is the well-known continuum action, while the correction terms are regarded as operator insertions in the continuum theory. They are of the form

$$S_d = \int d^4x \; \mathcal{L}_d(x) \,, \tag{3.86}$$

where $\mathcal{L}_d$ are linear combinations of local, gauge-invariant composite fields of dimension $4 + d$ that respect also the global symmetries of the theory. We will specify $\mathcal{L}_1$ below. Similarly, the renormalized lattice fields $Z_\phi \phi(x)$ are described in the SET by effective operators ϕ_d of dimension $\dim(\phi_d) = \dim(\phi_0) + d$:

$$\phi_{\mathrm{eff}} = \phi_0 + a\phi_1 + a^2\phi_2 + O(a^3) \,. \tag{3.87}$$

The correlation function (3.82) in the SET assumes the form

$$\begin{aligned} G_{n,\mathrm{eff}}(x_1, \, \dots \, , x_n) = {} & \langle \phi_0(x_1) \, \dots \, \phi_0(x_n) \rangle_{\mathrm{cont}} \\ & + a \int d^4y \, \langle \phi_0(x_1) \, \dots \, \phi_0(x_n) \; \mathcal{L}_1(y) \rangle_{\mathrm{cont}} \\ & + a \sum_{k=1}^{n} \langle \phi_0(x_1) \, \dots \, \phi_1(x_k) \, \dots \, \phi_0(x_n) \rangle_{\mathrm{cont}} + O(a^2) \,, \end{aligned} \tag{3.88}$$

where the subscript "cont" indicates that the continuum action S_0 is employed in the evaluation of the expectation values.

We restrict our discussion to *on-shell quantities* that can be derived from those correlation functions G_n (3.82) which only include fields ϕ at a physical distance to one another, $x_1 \neq x_2 \neq \ldots \neq x_n$. In that case, the contact terms that arise from the integration over y in (3.88) may be absorbed in the definition of the field ϕ_1. Consequently, the term $\mathcal{L}_1(y)$ can be considered independent, and the equations of motions may be used to substantially simplify it. To see this concretely, we write down the Symanzik effective Lagrangian to first order, which reads [46]

$$\mathcal{L}_{\text{Sym}} = \mathcal{L}_0 + a\mathcal{L}_1 = \mathcal{L}^0_{\text{Sym}} + \mathcal{L}^m_{\text{Sym}} \,, \tag{3.89a}$$

with[13]

$$\begin{aligned} \mathcal{L}^0_{\text{Sym}} &= \frac{1}{4} F_{\mu\nu} F_{\mu\nu} + \overline{\psi} \not{D} \psi \\ &+ a\rho_1 \, \overline{\psi} D_\mu D_\mu \psi \\ &+ a\rho_2 \, \overline{\psi} i \sigma_{\mu\nu} F_{\mu\nu} \psi \end{aligned} \tag{3.89b}$$

and

$$\begin{aligned} \mathcal{L}^m_{\text{Sym}} &= m_q \overline{\psi} \psi \\ &+ a\sigma_1 \, m_q^2 \, \overline{\psi} \psi \\ &+ a\sigma_5 \, m_q \, \mathrm{tr}[F_{\mu\nu} F_{\mu\nu}] \\ &+ a\sigma_6 \, m_q \, \overline{\psi} \not{D} \psi \,, \end{aligned} \tag{3.89c}$$

where $\sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu]$. It is split here into a mass-independent $\mathcal{L}^0_{\text{Sym}}$ and a mass-dependent part $\mathcal{L}^m_{\text{Sym}}$ for convenience. Note that purely gluonic terms beyond the continuum gauge action are at least of dimension 6 and do not need to be taken into account to the considered order. Moreover, the prefactors of the fermionic $O(a^0)$ terms in (3.89) are set to one. The normalization of the kinetic term can be achieved by rescaling of the fermion fields. In contrast, the mass m_q as it appears in $\mathcal{L}^m_{\text{Sym}}$, is understood to be related to the one in (3.84) by a multiplicative renormalization. Following standard practice, however, this difference is neglected here. The terms with ρ_1 and σ_6, which involve covariant derivatives, can be eliminated under use of the Dirac equation, i.e. they can be converted to terms which are already present in the effective Lagrangian.

We consider the effective bilinears $\phi = A_\mu, P$ (cf. (3.73, 3.74)) next. The principle which underlies the construction of improvement terms ϕ_1 is the same as for the action, they have to maintain the behavior of the respective currents under the symmetry

---

[13] The somewhat peculiar enumeration of the improvement coefficients here serves as preparation for the extension to non-degenerate masses, which leads to additional mass-dependent terms, as we will see in sec. 8.

transformations of the theory. Up to $O(a)$, the terms which fulfill this criterion for the effective axial and pseudoscalar currents are[14] [46]:

$$(A_{\text{Sym}})^c_\mu = A^c_\mu + a\omega_0\ \overline{\psi}\gamma_5 T^c D_\mu\psi + a\omega_1\ m_q A^c_\mu + a\omega_5\ \overline{\psi}\gamma_5 T^c i\sigma_{\mu\nu} D_\nu\psi \tag{3.90}$$

$$(P_{\text{Sym}})^c = P^c + a\omega_1'\ m_q P^c \tag{3.91}$$

Here, the term with ω_5 can be eliminated.

Improved lattice action

The insights from the SET pave the way to the improvement of the scaling behavior of both the lattice action and the lattice operator representations. In order to let the $O(a)$ terms in the Symanzik expansion vanish, one simply needs to add discretized versions of the Symanzik terms to the respective lattice quantities, with suitable *improvement coefficients* c. Those depend on the lattice spacing, or equivalently the bare coupling, $c \equiv c(g_0^2)$.

For *on-shell improvement* of the Wilson fermion and gauge action, the $O(a)$ counterterms that have to be taken into account can be read off from (3.89):

$$\begin{aligned}\mathcal{L}_{F,I} &= \overline{\psi}D_W\psi + m_0\overline{\psi}\psi \\ &+ ac_{\text{sw}}\ \overline{\psi}\frac{i}{4}\sigma_{\mu\nu}F_{\mu\nu}\psi \\ &+ ac_1\ m_q^2\overline{\psi}\psi \end{aligned} \tag{3.92a}$$

and

$$\begin{aligned}\mathcal{L}_{G,I} &= \mathcal{L}_G[U] \\ &+ ac_5\ m_q\ \text{tr}[F_{\mu\nu}F_{\mu\nu}]\ . \end{aligned} \tag{3.92b}$$

Here, $\mathcal{L}_G[U]$ is the Lagrangian associated with the gauge action, whose scaling behavior as discussed in sec. 3.2 is

$$S_G =: a^4\sum_x \mathcal{L}_G[U] = \frac{a^4}{2g_0^2}\sum_x \text{tr}[F_{\mu\nu}F_{\mu\nu}] + O(a^2)\ . \tag{3.93}$$

Note that the mass-dependent improvement terms in (3.92) could just as well contain the bare mass m_0 instead of the bare subtracted mass m_q, as the difference (m_{crit}) can be absorbed in the improvement coefficients c_1 and c_5. However, m_q is advantageous in view of the conjunction of $O(a)$ improvement with renormalization, to be discussed in the following. The mass-dependent improvement terms are both proportional to terms that already appear in the unimproved Lagrangian and may thus be absorbed

---

[14] See footnote 13.

in the bare parameters. While this is straight forward for the bare mass, cf. (3.92a), in the case of the bare coupling it may readily be seen under use of (3.93):

$$\begin{aligned}
\mathcal{L}_{G,I} &= \left(\frac{1}{2g_0^2} + ac_5\, m_q\right) \quad \mathrm{tr}[F_{\mu\nu}F_{\mu\nu}] + O(a^2) \\
&= \frac{1}{2g_0^2}\left(1 + 2ac_5 g_0^2\, m_q\right)\, \mathrm{tr}[F_{\mu\nu}F_{\mu\nu}] + O(a^2) \\
&= \frac{1}{2g_0^2}\,\frac{1}{1 - 2ac_5 g_0^2\, m_q} \quad \mathrm{tr}[F_{\mu\nu}F_{\mu\nu}] + O(a^2) \\
&= \mathcal{L}_G\big|_{g_0^2 \to g_0^2(1 - 2ac_5 g_0^2\, m_q)} \ .
\end{aligned} \tag{3.94}$$

One introduces the improvement coefficients

$$b_m = \sigma_1 \qquad \text{and} \qquad b_g = -2\sigma_5 \cdot g_0^2\,, \tag{3.95}$$

such that the absorption reads

$$g_0^2 \ \rightarrow\ \widetilde{g}_0^2 = g_0^2\,\left(1 + ab_g(g_0^2)m_q\right) \tag{3.96}$$

$$m_q \ \rightarrow\ \widetilde{m}_q = m_q\,\left(1 + ab_m(g_0^2)m_q\right)\ . \tag{3.97}$$

In this form, the $O(a)$ corrections may be accounted for in the renormalization of the bare parameters (cf. (3.61, 3.68)):

$$g_R^2 = Z_g(\widetilde{g}_0^2, a\mu)\ \ \widetilde{g}_0^2 \tag{3.98}$$

$$m_R = Z_{m^0}(\widetilde{g}_0^2, a\mu)\ \widetilde{m}_q\ . \tag{3.99}$$

We stress that this is again the mass-independent renormalization scheme here. The significance of the above formulas lies in the fact that for fixed renormalized quantities (e.g. if the continuum limit is taken), the modified bare parameters scale mass-independently, whereas the scaling of the ordinary bare parameters is mass-dependent if $O(a)$ improvement is to be achieved. Accordingly, if the masses are varied, the bare coupling needs to be varied simultaneously in order to keep $\widetilde{g}_0^2$ and hence the lattice spacing fixed.

Under the provision of (3.96)-(3.99), the $O(a)$ improved Wilson action reads

$$S_{\mathrm{I}}^{\mathrm{W}}[\psi,\overline{\psi},U] = S_G^{\mathrm{W}}[U] + S_F^{\mathrm{W}}[\psi,\overline{\psi},U] + \delta S_{\mathrm{sw}}[\psi,\overline{\psi},U]\,, \tag{3.100}$$

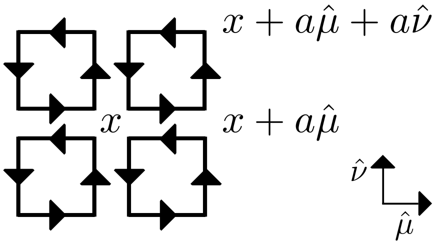


Figure 3.3: Illustration of the term $Q_{\mu\nu}(x)$ as given in (3.102b), which enters the clover discretization $F^{(\mathrm{clov})}_{\mu\nu}$ of the gluon field tensor (3.102a).

where the only remaining improvement term in the action is the *Sheikholeslami–Wohlert* term,

$$\delta S_{\mathrm{sw}}[\psi, \overline{\psi}, U] = a^5 c_{\mathrm{sw}}(g_0^2) \sum_x \overline{\psi}(x) \frac{i}{4} \sigma_{\mu\nu} F_{\mu\nu}(x) \psi(x) \ . \tag{3.101}$$

The coefficient $c_{\mathrm{sw}}(g_0^2)$ is called the *Sheikholeslami–Wohlert coefficient*. There is some arbitrariness in the discretization of the field strength tensor in (3.101), which does not affect the considered order of improvement though. A particular, frequently employed lattice representation is

$$F^{(\mathrm{clov})}_{\mu\nu}(x) = -\frac{1}{8a^2} \left(Q_{\mu\nu}(x) - Q_{\nu\mu}(x) \right) \tag{3.102a}$$

with[15]

$$Q_{\mu\nu}(x) = U_{\mu,\nu}(x) + U_{\nu,-\mu}(x) + U_{-\mu,-\nu}(x) + U_{-\nu,\mu}(x) \ . \tag{3.102b}$$

Due to the shape of $Q_{\mu\nu}(x)$ provided by the involved plaquettes, illustrated in fig. 3.3, it is referred to as *clover term*.

Improved bilinears and PCAC mass

For the improvement of the axial and pseudoscalar current, there is only one term that needs to be considered [46]:

$$(A_{\mathrm{I}})^c_\mu = A^c_\mu + a c_A(g_0^2)\, \tilde{\partial}_\mu P^c \tag{3.103}$$

$$(P_{\mathrm{I}})^c \ = P^c \tag{3.104}$$

---

[15] Note that in the literature, the definition of $Q_{\mu\nu}(x)$ is often such that its indices are interchanged. Correspondingly, in that case, a relative minus sign appears in (3.102a), and the orientation of the plaquettes in fig. 3.3 is reversed.

This is because the terms which correspond to ω_1, ω_1' in (3.90) and (3.91) may again be incorporated by means of the renormalization prescription (cf. (3.76, 3.77)):

$$\begin{aligned} (A_{\mathrm{RI}})^c_\mu &= Z_A(\tilde{g}_0^2) \quad \left[1 + ab_A(g_0^2)\, m_q\right] \; (A_{\mathrm{I}})^c_\mu \qquad (3.105) \\ (P_{\mathrm{RI}})^c &= Z_P(\tilde{g}_0^2, a\mu) \; \left[1 + ab_P(g_0^2)\, m_q\right] \; P^c \; . \qquad (3.106) \end{aligned}$$

It is now straightforward to translate the renormalization in terms of the PCAC mass in the unimproved $N_f = 4^*$ case, (3.80), to the improved $N_f = 4$ theory. It reads

$$m_{\mathrm{R}} = \frac{Z_A(\tilde{g}_0^2)}{Z_P(\tilde{g}_0^2, a\mu)} \, \frac{1 + ab_A(g_0^2)\, m_q}{1 + ab_P(g_0^2)\, m_q} \, m_{\mathrm{PCAC,I}}(x) + O(a^2) \; , \qquad (3.107)$$

in terms of the *improved PCAC mass* (cf. (3.81)),

$$m_{\mathrm{PCAC,I}}(x) = \frac{\tilde{\partial}_\mu \left\langle (A_{\mathrm{I}})_\mu (x)\, O \right\rangle}{2\, \langle P(x)\, O \rangle} \; . \qquad (3.108)$$

Note that the currents, although written without flavor indices due to the degeneracy, are still understood to be *non-singlet* currents.

All the renormalization and improvement coefficients in principle depend on $\tilde{g}_0^2$. Concerning the encountered improvement coefficients c_{sw}, c_A and b_g, b_m, b_A, b_P, however, the difference made by the use of the original and the modified bare coupling is of the order $O(a^2)$. We keep g_0^2 as their argument to emphasize that all expressions are set up to improve the theory up to $O(a)$ only.

There exist non-perturbative as well as perturbative determinations for them, which in general depend on the number N_f of flavors in the theory and the gauge action which is employed. We will return to this issue in sec. 3.9, when we have another, improved gauge action at our disposal. This so-called *tree-level improved Lüscher–Weisz gauge action* is subject of the upcoming section.

3.8 Tree-level improved Lüscher–Weisz gauge action

The scaling violations of the Wilson gauge action are of order $O(a^2)$. According to the Symanzik improvement program, one may add operators of dimension 6 to the gauge action in order to reduce the order of cutoff effects as the continuum limit is approached. These operators are given by the products $U(\mathcal{C})$ of 6 link variables along closed loops $\mathcal{C} \in S_i$ $(i = 1, 2, 3)$, where $\mathcal{S}_i$ is the set of all rectangles $(i = 1)$,

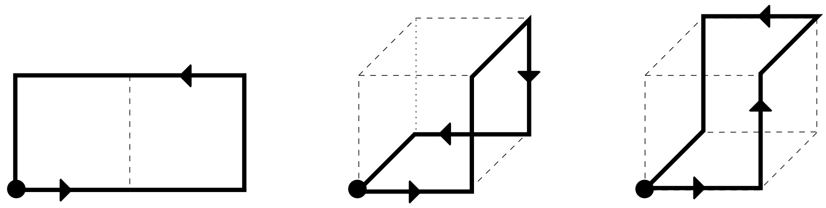


Figure 3.4: Elementary closed loops from which improved gauge actions are built up in addition to the plaquette depicted in fig. 3.1. *From left to right:* rectangle ($i = 1$), chair ($i = 2$), parallelogram ($i = 3$).

bent rectangles or chairs ($i = 2$) and parallelograms ($i = 3$) on the lattice, see fig. 3.4. Together with the set of all oriented plaquettes, $\mathcal{S}_0$, an improved lattice gauge action may conveniently be written as

$$S_G^{\mathrm{W,I}}[U] = \frac{1}{g_0^2} \sum_{i=0}^{3} c_i(g_0^2) \sum_{\mathcal{C} \in \mathcal{S}_i} \mathrm{tr}[\mathbb{1} - U(\mathcal{C})] \, . \tag{3.109}$$

Note that the unimproved Wilson gauge action S_G^{W} (cf. (3.39)) is recovered from (3.109) for $c_0 = 1$ and $c_1 = c_2 = c_3 = 0$. The improvement coefficient are subject to the normalization condition

$$c_0(g_0^2) + 8\, c_1(g_0^2) + 16\, c_2(g_0^2) + 8\, c_3(g_0^2) = 1 \, , \tag{3.110}$$

which is a convention that can always be enforced. In [47], it was argued that on-shell improvement (cf. sec. 3.7) provides two more constraints on the improvement conditions. To fix all the improvement coefficients, $c_2(g_0^2) = 0$ is a particularly suitable choice, for which the remaining coefficients were calculated at tree-level order[16] in perturbation theory. They read [47]

$$c_0 \equiv \; c_0(g_0^2 = 0) \;\; = 5/3 \tag{3.111a}$$

$$c_1 \equiv \; c_1(g_0^2 = 0) \;\; = -1/12 \tag{3.111b}$$

$$c_2 \equiv \; c_2(g_0^2 = 0) \;\; = 0 \tag{3.111c}$$

$$c_3 \equiv \; c_3(g_0^2 = 0) \;\; = 0 \, . \tag{3.111d}$$

---

[16] For a computation of the next order in perturbation theory, see [48].

The coefficients (3.111a, 3.111b) define the *tree-level improved Symanzik* or *Lüscher–Weisz gauge action*, which can be written as

$$S_G^{\mathrm{LW}}[U] = \frac{1}{g_0^2} \sum_{i=0,1} c_i \sum_{\mathcal{C} \in \mathcal{S}_i} \mathrm{tr}[\mathbb{1} - U(\mathcal{C})] \; . \tag{3.112}$$

Its use leads to scaling violations of the order $O(a^2 g_0^2)$. We emphasize that $\mathcal{S}_i$ is understood to contain all *oriented loops*, and (3.112) may also be found in the literature as a sum over all *unoriented loops* with an additional occurrence of $2\,\mathrm{Re}$, see the discussion at the end of sec. 3.2.

3.9 Results for improvement coefficients

3.9.1 c_{sw}

The improvement coefficient c_{sw} with its *perturbative* expansion

$$c_{\mathrm{sw}}(g_0^2) = c_{\mathrm{sw}}^{(0)} + c_{\mathrm{sw}}^{(1)} g_0^2 + O(g_0^4) \tag{3.113a}$$

depends on the employed gauge action. The tree-level coefficient

$$c_{\mathrm{sw}}^{(0)} = 1 \tag{3.113b}$$

is universal, however, and has been computed along with the one-loop coefficient $c_{\mathrm{sw}}^{(1)}$ for the Wilson gauge action S_G^{W} in [49, 50]. The numerical value of the counterpart for the tree-level improved Lüscher–Weisz gauge action S_G^{LW} is known from [51]. They read

$$c_{\mathrm{sw}}^{(1)} = \begin{cases} 0.2659(1) & \text{for } S_G = S_G^{\mathrm{W}} \\ 0.196(6) & \text{for } S_G = S_G^{\mathrm{LW}} \; . \end{cases} \tag{3.113c}$$

Non-perturbative results for c_{sw} were obtained by the ALPHA Collaboration with various numbers of massless, dynamic quark flavors, and under use of both the Wilson gauge action and the tree-level improved Lüscher–Weisz gauge action. Details on the strategies employed in these works as well as a compilation of their findings can be found in app. F.1. For $N_f = 2$ in connection with S_G^{W}, the result reads [17]

$$c_{\mathrm{sw}}(g_0^2) = \frac{1 - 0.454g_0^2 - 0.175g_0^4 + 0.012g_0^6 + 0.045g_0^8}{1 - 0.720g_0^2} \qquad \text{for } \beta \geq 5.2 \ . \tag{3.114}$$

3.9.2 c_A

The *perturbative* expansion is similar to (3.113a) for the improvement coefficient c_A, the difference being that it vanishes at tree-level [52],

$$c_A(g_0^2) = c_A^{(1)} g_0^2 + O(g_0^4) \ . \tag{3.115a}$$

The leading one-loop coefficient $c_A^{(1)}$ was obtained for the Wilson gauge action first in [50] and rederived with improved accuracy in [53]. The case of the tree-level improved Lüscher–Weisz gauge action has again been covered in [51]. The results are

$$c_A^{(1)} = \begin{cases} -0.005680(2) \times C_F & \text{for } S_G = S_G^{\mathrm{W}} \\ -0.004525(25) \times C_F & \text{for } S_G = S_G^{\mathrm{LW}} \ , \end{cases} \tag{3.115b}$$

with $C_F = 4/3$ (cf. (A.12)).

The *non-perturbative* determinations of c_A, again performed for different numbers of massless, dynamic quark flavors and both the gauge actions, S_G^{W} and S_G^{LW}, are covered in app. F.2. The result for $N_f = 2$ in connection with S_G^{W} reads [54]

$$c_A(g_0^2) = -0.00756g_0^2 \times \frac{1 - 0.4485g_0^2}{1 - 0.8098g_0^2} \qquad \text{for } 0.98 \leq g_0^2 \leq 1.16 \ . \tag{3.116}$$

3.9.3 b_g, b_m, b_A, b_P

For the improvement coefficients b_g, b_m, b_A, b_P and their respective *perturbative* expansions,

$$b_X(g_0^2) = b_X^{(0)} + b_X^{(1)} g_0^2 + O(g_0^4) \ , \tag{3.117}$$

the parameters are known up to one-loop order, see tab. 3.2. Note that the numerical one-loop expression of the combination $b_A - b_P$ as it appears in (3.107) is quite small, namely [53]

$$b_A - b_P = -0.00093(8) \cdot g_0^2 + O(g_0^4) \ . \tag{3.118}$$

X	$b_X^{(0)}$	Ref.	$b_X^{(1)}$	Ref.
g	0	[46]	$0.012000(2) \cdot N_f$	[46]
m	$-\frac{1}{2}$	[50]	$-0.07217(2) \cdot C_F$	[53]
A	1	[50, 52]	$0.11414(4) \cdot C_F$	[53]
P	1	[50, 52]	$0.11484(2) \cdot C_F$	[53]

Table 3.2: Coefficients of the perturbative expansions (3.117) of $b_X \in \{b_g, b_m, b_A, b_P\}$.

It was computed also *non-perturbatively*, along with b_m and the renormalization constant Z (cf. (E.17)), in $N_f = 0$ [55–57] as well as $N_f = 2$ [58].

Regarding b_g, there exist no non-perturbative results for any gauge action[17]. Note, however, that the fact that it is associated with sea quark effects is reflected in both, the fact that it vanishes at tree-level order (in accordance with (3.95)), and the occurrence of the explicit factor N_f in $b_g^{(1)}$, see tab. 3.2.

---

[17] This is a significant restriction in large volume simulations with massive sea quarks, where one needs to adopt a strategy to keep $\widetilde{g}_0^2$ and thus the lattice spacing fixed (cf. sec. 3.7) without the knowledge of b_g, see e.g. [38, 59].

4 Monte Carlo simulations and numerical techniques

In this section, we cover the numerical simulation of Euclidean correlation functions, which we recall here for convenience:

$$\langle O \rangle = \frac{1}{Z} \int \mathcal{D}[\psi, \overline{\psi}, U] \; e^{-S[\psi,\overline{\psi},U]} O[\psi, \overline{\psi}, U] \; , \tag{3.46}$$

with

$$Z = \int \mathcal{D}[\psi, \overline{\psi}, U] \; e^{-S[\psi,\overline{\psi},U]} \; . \tag{3.47}$$

The different statistics of the fermion and boson fields demand substantially different methods for the numerical integration of the respective degrees of freedom. Hence, it is useful to split off the fermionic part of the integration:

$$\langle O \rangle = \frac{1}{Z} \int \mathcal{D}[U] \; e^{-S_G[U]} \; Z_F[U] \; \langle O \rangle_F [U] \; , \tag{4.1}$$

where

$$\langle O \rangle_F [U] = \frac{1}{Z_F[U]} \int \mathcal{D}[\psi, \overline{\psi}] \; e^{-S_F[\psi,\overline{\psi},U]} O[\psi, \overline{\psi}, U] \tag{4.2}$$

contains the integration over the fermionic degrees of freedom and

$$Z_F[U] = \int \mathcal{D}[\psi, \overline{\psi}] \; e^{-S_F[\psi,\overline{\psi},U]} \tag{4.3}$$

is the fermionic partition function. Note that S_F will from now on always represent the Wilson fermion action (3.44), whereas S_G may be either the Wilson gauge (3.36) or tree-level improved Lüscher–Weisz gauge action (3.112).

We will explore the techniques to evaluate the different parts of the functional integral (4.1) step by step. In sec. 4.1, we consider the implementation of the fermionic degrees of freedom in the path integral. We will learn that the fermionic expectation value (4.2) can be explicitly computed, whereas the partition function (4.3) is treated numerically together with the gauge field integration in the framework of Monte Carlo simulations. These will be explained in detail in the subsequent sections. First, in sec. 4.2 and sec. 4.3, under complete neglect of the fermionic contributions included in Z_F. The contributions of the different flavors and the techniques they require will then be incorporated gradually in sec. 4.4-4.6. After a brief survey on important numerical techniques in sec. 4.7-4.9, we will finally summarize the estimation of statistical errors using the so-called gamma method in sec. 4.10.

4.1 Fermions in the path integral

We first consider the fermionic path integral

$$\langle O\rangle_F[U] = \frac{1}{Z_F[U]} \int \mathcal{D}[\psi,\overline{\psi}]\; e^{-S_F[\psi,\overline{\psi},U]} O[\psi,\overline{\psi},U] \; . \tag{4.2}$$

The operator O may involve products of fermion fields ψ and $\overline{\psi}$, which need to obey Fermi statistics, i.e. an interchange of any two components of ψ or $\overline{\psi}$ is required to produce a minus sign. To implement this, one needs anti-commuting Grassmann variables η_i and $\bar{\eta}_i$, one for each component of ψ and $\overline{\psi}$, respectively:

$$\eta_i\eta_j = -\eta_j\eta_i \; , \bar{\eta}_i\bar{\eta}_j = -\bar{\eta}_j\bar{\eta}_i \; , \bar{\eta}_i\eta_j = -\eta_j\bar{\eta}_i \; , \tag{4.4}$$

with

$$i, j \in \{1, \ldots, N_{\mathrm{comp}} = 12N_f|X|\} \tag{4.5}$$

given in terms of the number $|X|$ of lattice sites, cf. (3.14). For a given gauge configuration U, the fermionic expectation value (4.2) may be computed in closed form with the help of Wick's theorem for Grassmann variables:

$$\begin{aligned} \left\langle \eta_{i_1}\overline{\eta}_{j_1}\ldots\eta_{i_n}\overline{\eta}_{j_n} \right\rangle_F &= \frac{1}{Z_F} \int \prod_{k=1}^{N_{\mathrm{comp}}} d\eta_k d\overline{\eta}_k \;\; \eta_{i_1}\overline{\eta}_{j_1}\ldots\eta_{i_n}\overline{\eta}_{j_n} \; \exp\left(-\sum_{l,m=1}^{N_{\mathrm{comp}}} \overline{\eta}_l D_{lm}\eta_m \right) \\ &= \sum_P \mathrm{sign}(P) \; (D^{-1})_{i_1 j_{P_1}} \ldots (D^{-1})_{i_n j_{P_n}} \; . \end{aligned} \tag{4.6}$$

Here, the sum runs over all possible permutations $P(1,\dots,n) = (p_1,\dots,p_n)$ of the numbers $1,\dots,n$. The *massive* Wilson–Dirac operator (cf. (3.13, 3.43))

$$D = D_W + \mathcal{M} \tag{4.7}$$

with its $N_{\text{comp}} \times N_{\text{comp}}$ components is equipped with multi-indices, which stand short for a set of four indices labelling spacetime coordinates as well as Dirac, color and flavor components, e.g. $i_1 = (x, \alpha, a, f)$. A proof of the above formula, which employs the generating functional for fermions, can be found e.g. in [31].

As an example of its application, we consider the hadron correlator $\langle O_H(t)\, \overline{O}_H(0)\rangle$ already encountered in (3.51). For the specific case of the correlator of the operators

$$O_{\pi^+}(x) \qquad\qquad = \overline{d}(x)\gamma_5 u(x) \tag{4.8a}$$

$$\overline{O}_{\pi^+}(y) = O_{\pi^-}(y) = \overline{u}(x)\gamma_5 d(x), \tag{4.8b}$$

between two spacetime points $x, y \in X$, one finds

$$\begin{aligned}
\left\langle O_{\pi^+}(x)\; \overline{O}_{\pi^+}(y)\right\rangle_F &= \left\langle \overline{d}(x)\gamma_5 u(x)\; \overline{u}(y)\gamma_5 d(y)\right\rangle_F \\
&= \quad (\gamma_5)_{\alpha_1\beta_1}(\gamma_5)_{\alpha_2\beta_2}\; \left\langle \overline{d}(x)^{\alpha_1}_{c_1} u(x)^{\beta_1}_{c_1}\; \overline{u}(y)^{\alpha_2}_{c_2} d(y)^{\beta_2}_{c_2}\right\rangle_F \\
&= -(\gamma_5)_{\alpha_1\beta_1}(\gamma_5)_{\alpha_2\beta_2}\; \left\langle u(x)^{\beta_1}_{c_1}\overline{u}(y)^{\alpha_2}_{c_2}\right\rangle_u\; \left\langle d(y)^{\beta_2}_{c_2}\overline{d}(x)^{\alpha_1}_{c_1}\right\rangle_d \\
&= -\mathrm{tr}\left[\gamma_5\; D_u^{-1}(x|y)\; \gamma_5\; D_d^{-1}(y|x)\right] \; .
\end{aligned} \tag{4.9}$$

In the third line the order of the Grassmann variables was changed and the fermion expectation value was factorized in flavor contributions. Subsequently, the Wick theorem (4.6) was applied. The notation of the quark propagators indicates that the flavor and spacetime indices are fixed, whereas the other indices are summed over by the trace tr that is to be understood in both color and Dirac space here. The pion propagator in the form (4.9) may be computed numerically for a given gauge configuration U. Note that the correlator with mere time dependence, (3.51), which in the case at hand is suitable to extract the pion mass, is found by momentum projection of (4.9).

The fermionic partition function Z_F (4.3) in terms of the Grassmann variables reads

$$Z_F = \int \prod_{k=1}^{N_{\text{comp}}} d\eta_k d\overline{\eta}_k \; \exp\left(- \sum_{l,m=1}^{N_{\text{comp}}} \overline{\eta}_l D_{lm} \eta_m \right) \; .$$

The Matthews-Salam formula [60, 61] allows to integrate out the fermionic degrees of freedom in the partition function and write it as a *fermion determinant*:

$$Z_F = \det[D] \; . \tag{4.10}$$

It is convenient for later use to split the partition function in contributions of different flavors:

$$Z_F = \prod_f \det[D_f] \; , \tag{4.11}$$

with D_f now being matrices of the size $12|X| \times 12|X|$. The path integral (4.1) under use of (4.11) reads

$$\langle O \rangle = \frac{1}{Z} \int \mathcal{D}[U] \; e^{-S_G[U]} \; \prod_f \det[D_f] \; \langle O \rangle_F [U] \; , \tag{4.12a}$$

where the partition function with the fermions integrated out assumes the form

$$Z = \int \mathcal{D}[U] \; e^{-S_G[U]} \; \prod_f \det[D_f] \; . \tag{4.12b}$$

The splitting of the fermionic part into a fermion determinant and the fermionic expectation value can physically be interpreted as follows. The fermion determinant $\prod_f \det[D_f]$ represents dynamical *sea quark* effects, i.e. it describes the creation and annihilation of virtual quark-antiquark pairs. This viewpoint is suggested by investigation of the hopping expansion (see sec. 3.3) for the determinant, see e.g. [31]. The fermionic expectation value $\langle O \rangle_F [U]$ describes the propagation of the state defined by O under the interaction with the external gauge field U, which is itself affected by the aforementioned dynamical quark effects. In the case of our example (4.9), it is the *valence quarks* of a pion(-like) state that interact with U.

The question arises of how to treat the fermion determinant mathematically. A direct determination of $\det[D_f]$, which formally has $(12|X|)!$ terms, is absolutely out of reach regarding the computational costs that are of the order[1] $O((12|X|)^3)$. Historically, the first Lattice QCD computations were for this reason performed in the *quenched approximation*, where the sea quark effects are simply ignored:

$$Z_F = \prod_f \det[D_f] \to 1 \; . \tag{4.13}$$

---

[1] This refers to the use of Gaussian elimination.

They may be included nowadays by assignment of the fermion determinant to the Boltzmann weight factor of the gauge field, $e^{-S_G[U]}$, in the framework of Monte Carlo simulations. However, since this is both computationally expensive and technically involved, the consideration of the different flavors has been done stepwise. Depending on the dynamical quark content of the theory, one speaks of $N_f = 0$ (quenched) and $N_f = 2$ (including mass-degenerate up and down quarks), for instance. In the case where strange and charm quark flavors with different masses add to the theory, it is labelled by $N_f = 2+1$ or $N_f = 2+1+1$.

We follow the historical route and begin the discussion of Monte Carlo simulations ignoring dynamical quark effects. As we gradually move on to the full theory with four quark flavors, we will take account of more sophisticated numerical techniques.

4.2 Markov chain Monte Carlo simulations

In the $N_f = 0$ theory without dynamical fermions, the full path integral (4.12) reduces to

$$\langle O \rangle_{N_f=0} = \frac{1}{Z_G} \int \mathcal{D}[U]\; e^{-S_G[U]}\, \langle O \rangle_F\,[U] \tag{4.14}$$

with

$$Z_G = \int \mathcal{D}[U]\; e^{-S_G[U]}\ . \tag{4.15}$$

Quarks may only appear as valence (or spectator) quarks in $\langle O \rangle_F$. In pure gauge theory, the fermion expectation value simplifies to $\langle O \rangle_F\,[U] = O[U]$. From counting the link variables in

$$\mathcal{D}[U] = \prod_x \prod_\mu dU_\mu(x)\ , \tag{3.50}$$

we conclude that the above integral is of dimension[2] 4 times the number of lattice sites, i.e. $d = 4 \cdot (L/a)^4$ on a hypercubic lattice of size L/a in each of the four directions. In practical situations one virtually always faces $L/a \geq 8$, which means $d \geq 2^{14} = 16384$. Those integrals cannot to be solved analytically, and they are also way too high-dimensional for standard numerical integration methods (e.g. Simpson) to work.

A technique to compute them numerically is provided by *Monte Carlo simulations*.

---

[2] We ignore the different components of the links here for simplicity.

These are based on the principle of *importance sampling*. One creates an ensemble

$$\{U^{(n)} \mid n = 1, \ldots, N_{\text{cf}}\} \tag{4.16}$$

of N_{cf} gauge field configurations $U^{(n)}$, each of them being a specific choice of $U_\mu(x)$ at all (x, μ), which are sampled from the space of all gauge configurations according to the probability distribution given by the Boltzmann factor:

$$P(U) = \frac{1}{Z_G}\, e^{-S_G[U]} \,. \tag{4.17}$$

In the limit of an infinitely large ensemble, the path integral becomes the arithmetic mean of the observable evaluated on the N_{cf} sample configurations[3]:

$$\begin{aligned} \langle O \rangle_{N_f=0} &= \frac{1}{Z_G} \int \mathcal{D}[U]\, e^{-S_G[U]}\ \langle O \rangle_F\, [U^{(n)}] \\ &= \lim_{N_{\text{cf}} \to \infty} \frac{1}{N_{\text{cf}}} \sum_{n=1}^{N_{\text{cf}}} \langle O \rangle_F\, [U^{(n)}] \,. \end{aligned} \tag{4.18}$$

In practice, this exact expression is approximated by the estimator

$$\overline{O} = \frac{1}{N_{\text{cf}}} \sum_{n=1}^{N_{\text{cf}}} \langle O \rangle_F\, [U^{(n)}] \approx \langle O \rangle_{N_f=0} \ , \tag{4.19}$$

whose statistical error is proportional to[4] $1/\sqrt{N_{\text{cf}}}$, as we will see in sec. 4.10.

The gauge configurations cannot be importance sampled directly as the many link variables are coupled via the action. Instead, one uses a *Markov chain*. Starting from one configuration $U_{th}^{(1)}$, one generates a stochastic sequence of gauge configurations,

$$U_{th}^{(1)} \to U_{th}^{(2)} \to \ldots \to U_{th}^{(N_{\text{th}})} \quad \rightsquigarrow \quad U^{(1)} \to \ldots \to U^{(N_{\text{cf}})} \,. \tag{4.20}$$

The transition amplitude $T(U \to U')$, which is equal to the probability of a configuration U to be transformed into a configuration U' in one Monte Carlo step, needs to fulfill a few conditions in order to guarantee that the ensemble represents the aspired probability distribution. Apart from the trivial statements reflecting that $T(U \to U')$

---

[3] This holds if P has a finite first moment.

[4] This holds if P has a finite second moment.

is a normalized probability distribution,

$$\text{I)} \qquad 0 \leq T(U \to U') \leq 1 \tag{4.21}$$

$$\text{II)} \qquad \sum_{U'} T(U \to U') = 1 \, , \tag{4.22}$$

the so-called *balance condition* is crucial:

$$\text{III)} \qquad \sum_{U} T(U \to U') P(U) = \sum_{U} T(U' \to U) P(U') \, . \tag{4.23}$$

The left hand side is the total probability to find U' after a Monte Carlo step, whereas the right hand side is equal to the total probability to hop out of U'. Note that due to (4.22), the right hand side of (4.23) is the same as $P(U')$. Hence, once the probability distribution is obtained, the balance condition ensures that the Markov chain maintains it and the system is in *equilibrium*. However, the start configuration $U_{th}^{(1)}$ itself is often not selected with the very probability $P(U)$, in which case the system is not in equilibrium from the start. Fortunately though, it can be shown (see e.g. [62]) that the balance condition[5] leads to the equilibrium distribution after a finite number of Markov chain steps. This phase is called *thermalization*, and the N_{th} gauge configurations sampled before the equilibrium is reached, marked with a subscript "th" in (4.20), are usually discarded. A fourth crucial condition is *ergodicity*:

$$\text{IV)} \qquad \forall \; U, U' \;\; \exists \; n \in \mathbb{N} : \;\; T^n(U \to U') > 0 \, . \tag{4.24}$$

Here, T^n is the transition probability to get from U to U' in n Monte Carlo steps. Ergodicity makes sure that the whole configuration space is accessible for the Markov chain. Finally, we note that instead of III), many Monte Carlo algorithms implement the stronger so-called *detailed balance condition*

$$\text{IIIb)} \quad T(U \to U') P(U) = T(U' \to U) P(U') \, , \tag{4.25}$$

which obviously implies III).

---

[5] in combination with ergodicity, see below.

4.3 Metropolis algorithm

The simplest algorithm which fulfills the detailed balance condition (4.25) is the *Metropolis algorithm* [63]. It evolves a configuration U to U' in two steps:

- A candidate configuration U' is chosen with a selection probability $T_S(U \to U')$.
- U' is accepted with the acceptance probability

$$T_A(U \to U') = \min\left(1, \frac{T_S(U' \to U)\exp\left(-S[U']\right)}{T_S(U \to U')\exp\left(-S[U]\right)}\right) \ . \tag{4.26}$$

If U' is rejected, one reuses U in the Markov chain, $U' = U$.

It can easily be shown that the total transition amplitude $T = T_S \cdot T_A$ obtained this way indeed satisfies the detailed balance condition:

$$\begin{aligned} T(U \to U')\ P(U) &= T_S(U \to U')\ \min\left(1, \frac{T_S(U' \to U)\exp\left(-S[U']\right)}{T_S(U \to U')\exp\left(-S[U]\right)}\right)\ \exp\left(-S[U]\right) \\ &= \min\left(T_S(U \to U')\ \exp\left(-S[U]\right), T_S(U' \to U)\ \exp\left(-S[U']\right)\right) \\ &= T_S(U' \to U)\ \min\left(\frac{T_S(U \to U')\exp\left(-S[U]\right)}{T_S(U' \to U)\exp\left(-S[U']\right)}, 1\right)\ \exp\left(-S[U]\right) \\ &= T(U' \to U)\ P(U') \ . \end{aligned} \tag{4.27}$$

In pure gauge theory with the Wilson gauge action S_G (3.36), one usually conducts the above steps individually for a single link, $U_\mu(x)$ to $U'_\mu(x)$. Hence, the Metropolis Monte Carlo algorithm is called *local*. After a so-called sweep through the lattice, where one repeats this update procedure for every link, one obtains the configuration U' from U. The proposal for a concerned link, $U_\mu(x) \to U'_\mu(x)$, can be found by

$$U'_\mu(x) = X U_\mu(x) \ , \tag{4.28}$$

where $X \in SU(3)$ is a random element in the vicinity of $\mathbb{1}$. If X and X^{-1} are proposed with equal probability, the selection probability is symmetric:

$$T_S(U_\mu(x) \to U'_\mu(x)) = T_S(U'_\mu(x) \to U_\mu(x)) \ . \tag{4.29}$$

In that case, (4.26) simplifies to

$$T_A(U_\mu \to U'_\mu) = \min\left(1, \exp\left(-\Delta S_G\right)\right) \tag{4.30}$$

with

$$\Delta S_G = \Delta S_G[U'_\mu(x)] - \Delta S_G[U_\mu(x)] \ . \tag{4.31}$$

Note that ΔS_G can be computed locally, since all but one link stay the same and only six of the plaquettes which are summed over in S_G are affected. From (4.30), we see that a change in the action is always accepted if the action decreases or remains the same, whereas a new configuration with an increased action is accepted occasionally. This is in accordance with the principle that the path integral receives most of its contributions from configurations which (nearly) minimize the action, while quantum fluctuations are accounted for as well.

For completeness, we mention that there exist other update procedures which have been applied for gauge fields. Among them are the heat bath algorithm [64], over-relaxation [65, 66], the microcanonical update [67], the Langevin method [68, 69] and the Hybrid Monte Carlo algorithm [70]. The latter, also suitable for simulations with dynamical fermions, is the most widely-used algorithm in Lattice QCD nowadays and will be the one we focus on in this work.

4.4 Pseudofermions

On our way to the path integral in the full theory, (4.12), we now move on from the quenched approximation (4.14) to the $N_f = 2$ theory,

$$\langle O \rangle_{N_f=0} \ \rightarrow \ \langle O \rangle_{N_f=2} \ , \tag{4.32}$$

by including the fermion determinant for the up and down quark in the Monte Carlo integration:

$$e^{-S_G[U]} \ \rightarrow \ e^{-S_G[U]} \prod_{f=u,d} \det[D_f] \ . \tag{4.33}$$

In order for the interpretation of the right hand side of (4.33) as a Boltzmann weight to be possible, the product of fermion determinants obviously needs to be real and non-negative. The former is ensured by the fact that the massive Wilson–Dirac operator $D_f = D_W + m_f$ for any flavor f is γ_5-hermitean, see (3.45):

$$\det[D_f]^\star = \det[D_f^\dagger] = \det[\gamma_5 D_f \gamma_5] = \det[D_f] \tag{4.34}$$

Correspondingly, the eigenvalues of a γ_5-hermitean operator are either real or come in complex conjugate pairs. The non-negativity, however, is not ensured due to the breaking of chiral symmetry by the Wilson discretization[6]. This problem is circumvented in different ways for the different quarks. For the light up and down quarks, one makes use of the fact that these are to a very good approximation mass-degenerate, in which case one finds that their contributions together are strictly non-negative:

$$\prod_{f=u,d} \det[D_f] \geq 0 \, . \tag{4.35}$$

The inclusion of the heavier strange and charm quark works on a different basis and the discussion of this is postponed to sec. 4.6. Meanwhile, we will concentrate on $N_f = 2$ as outlined at the beginning of this section. We note that the product of fermion determinants for the two degenerate flavors, $D_u = D_d$, can be summarized as

$$\prod_{f=u,d} \det[D_f] = \det[D_u] \det[D_u^\dagger] = \det[D_u D_u^\dagger] \, , \tag{4.36}$$

where in the first step the γ_5-hermiticity was used again. The fermion determinant can be numerically implemented avoiding anti-commuting Grassmann variables, with the help of so-called *pseudofermions*. These are bosonic degrees of freedom represented by complex numbers $\phi = \phi_R + i\phi_I$ with $12|X|$ components (cf. (4.5)), which may be used to replace the integral over Grassmann variables by [72]

$$\det[D_u D_u^\dagger] = \pi^{-N_{\text{comp}}} \int \mathcal{D}[\phi] \; e^{-\phi^\dagger (D_u D_u^\dagger)^{-1} \phi} \, , \tag{4.37}$$

where

$$S_{PF,ud}[U, \phi] = \phi^\dagger (D_u D_u^\dagger)^{-1} \phi \tag{4.38}$$

is the *pseudofermion action*. The approach is based on a formal analogy of the Gaussian integrals of fermions and bosons, provided that all the eigenvalues of the (normal and thus diagonalizable) operator in the determinant are positive, in which case the integral on the right hand side of (4.37) converges. This condition is fulfilled for the combination of a Dirac operator and its adjoint, whose appearance, as we stress again, is based on the even number of degenerate quark flavors, see (4.36). Under use of the

---

[6] It can in fact be shown that for Ginsparg-Wilson fermions [71], which respect chiral symmetry on the lattice, the eigenvalues λ of the associated Dirac operator lie on a circle in the complex half-plane with $\mathrm{Re}\,\lambda \geq 0$.

above equations, (4.37) and (4.38), the path integral average in $N_f = 2$ may be written as

$$\langle O \rangle_{N_f=2} = \frac{1}{Z} \int \mathcal{D}[U]\mathcal{D}[\phi]\; e^{-S[U,\phi]}\; \langle O \rangle_F [U] \,, \tag{4.39}$$

where

$$S[U,\phi] = S_G[U] + S_{PF,ud}[U,\phi] \tag{4.40}$$

summarizes both the gauge and the pseudofermion action, the abbreviation $\mathcal{D}[\phi] = \mathcal{D}[\phi_R]\mathcal{D}[\phi_I]$ is used, and

$$Z = \int \mathcal{D}[U]\mathcal{D}[\phi]\; e^{-S[U,\phi]} \tag{4.41}$$

is the gauge and pseudofermion field partition function. The trade-off for the use of pseudofermion fields that one has to take into account is that in contrast to the local gauge action S_G, the pseudofermion action $S_{PF,ud}$ (4.38) is highly non-local as it contains the *inverse* of $D_u D_u^\dagger$. Somewhat simplifying, the computer time to determine the inverse is proportional to the lattice volume $V = (L/a)^4$, which makes a factor V^2 for a whole sweep over the lattice. Hence, in a *local* Monte Carlo algorithm like the Metropolis, the computation of the change of the total action after a local update $U \to U'$, (cf. (4.30, 4.31))

$$\begin{aligned}\exp(-\Delta S) &= \exp\left(-S[U',\phi'] + S[U,\phi]\right) \\ &= \exp\left(-S_G[U'] + S_G[U]\right)\; \exp\left(-S_{PF}[U',\phi'] + S_{PF}[U,\phi]\right) \,,\end{aligned} \tag{4.42}$$

would become very expensive. Updating all the fields globally by the Metropolis algorithm would not lead to a gain as it would compromise the acceptance rate.

Obviously, the pseudofermion fields require new techniques. A very effective Monte Carlo algorithm which updates the configurations globally in a specific way is the aforementioned *Hybrid Monte Carlo* algorithm, which we are going to describe next.

4.5 Hybrid Monte Carlo

The Hybrid Monte Carlo (HMC) algorithm [70] combines the so-called *molecular dynamics* evolution with a Metropolis acceptance/rejection step. The idea is that candidate gauge field configurations U' for a global update are created from U by numerical integration of the classical equations of motion in order to keep the change in the action small, which allows for large steps in configuration space with a reasonable acceptance probability. For convenience, we use anti-hermitean generators $\tilde{T}^a = iT^a$ (see app. A.1) in the following. Link variables may be written in terms of those as

$$U_\mu(x) = \exp\left(u_\mu(x)\right) \quad \in SU(3)\ , \tag{4.43}$$

with

$$u_\mu(x) = \sum_{a=1}^{8} u_\mu^{(a)}(x)\ \tilde{T}^a \in \mathfrak{su}(3)\ , \quad u_\mu^{(a)}(x) \in \mathbb{R}\ . \tag{4.44}$$

One then introduces conjugate momenta

$$\pi_\mu(x) = \sum_{a=1}^{8} \pi_\mu^{(a)}(x)\ \tilde{T}_a \in \mathfrak{su}(3)\ , \quad \pi_\mu^{(a)}(x) \in \mathbb{R}\ , \tag{4.45}$$

which allow to reformulate the expectation value (4.39) according to

$$\begin{aligned} \langle O \rangle_{N_f=2} &= \frac{\int \mathcal{D}[U]\mathcal{D}[\phi]\ \exp\left(-S[U,\phi]\right)\ \langle O \rangle_F[U]}{\int \mathcal{D}[U]\mathcal{D}[\phi]\ \exp\left(-S[U,\phi]\right)} \\ &= \frac{\int \mathcal{D}[U]\mathcal{D}[\phi]\mathcal{D}[\pi]\ \exp\left(-\frac{1}{2}\pi^2 - S[U,\phi]\right) \langle O \rangle_F[U]}{\int \mathcal{D}[U]\mathcal{D}[\phi]\mathcal{D}[\pi]\ \exp\left(-\frac{1}{2}\pi^2 - S[U,\phi]\right)}\ , \end{aligned} \tag{4.46}$$

where

$$\pi^2 = \sum_{x,\mu,a} \pi_\mu^{(a)}(x)\ \pi_\mu^{(a)}(x)\ . \tag{4.47}$$

While the introduction of these auxiliary fields does not change the physical content of the theory, (4.46) may be interpreted as a statistical system with the Hamiltonian

$$H[\pi, U, \phi] = \frac{1}{2}\pi^2 + S[U,\phi]\ . \tag{4.48}$$

Since this is reminiscent of the classical description of a molecular gas, the associated Hamilton equations—expressed here in terms of $U_\mu(x)$ instead of $u_\mu(x)$—

$$\dot{\pi}_\mu(x) = -\tilde{T}^a\ \partial^a_{x,\mu} S[U,\phi] =: -F[U,\phi] \tag{4.49a}$$

$$\dot{U}_\mu(x) = \pi_\mu(x)\ U_\mu(x)\ , \tag{4.49b}$$

are also called *molecular dynamics equations*. The derivatives on the left hand sides of (4.49a) and (4.49b) are understood to be with respect to the molecular dynamics (computer) time t_m in which the configurations are evolved. The *force* $F[U,\phi]$ on the right hand side of (4.49a), formally a derivative with respect to the $SU(3)$ group element $U_\mu(x)$, is defined as follows:

$$\frac{\partial S[U,\phi]}{\partial U_\mu(x)} = \tilde{T}^a \left. \frac{\partial S[e^\omega U,\phi]}{\partial \omega_\mu^a(x)} \right|_{\omega=0} \equiv \tilde{T}^a\, \partial^a_{x,\mu} S[U,\phi] \,. \tag{4.50}$$

Based on (4.49a) and (4.49b), the individual steps to obtain a gauge configuration U' from U in the HMC algorithm, including the contributions of the pseudofermions in $N_f = 2$, are the following:

1. The pseudofermion integral is evaluated by importance sampling, i.e. one selects a complex vector ϕ with the selection probability

   $$P(\phi) \sim \exp\left(-\phi^\dagger \left(D_u D_u^\dagger\right)^{-1} \phi\right) \tag{4.51}$$

   for the given gauge field U. In practice, this is done by randomly choosing a complex vector χ from the Gaussian distribution $P(\chi) \sim \exp\left(-\chi^\dagger\chi\right)$, such that $\phi = D_u\chi$ follows (4.51).

2. One selects a set[7] of $8 \cdot 4 \cdot |X|$ conjugate momenta $\pi_\mu^{(a)}(x)$ from the Gaussian distribution

   $$P(\pi) \sim \exp\left(-\frac{1}{2}\pi^2\right) \,. \tag{4.52}$$

3. Both the link field U and the conjugate field π are evolved in t_m via integration of the molecular dynamics equations, yielding a candidate configuration U' as well as π' after a certain molecular dynamics time t'_m. The standard choice for the integration length which defines this Monte Carlo step is $t'_m = 2$, called 2 *molecular dynamics units (MDU)*. The use of the Hamilton equations in principle restricts the produced configurations to a hypersurface of constant energy in phase space. However, the numerical integration is necessarily performed in discrete steps and hence leads to a non-vanishing value of ΔH through numerical errors.

---

[7] Recall that $|X|$ denotes the number of lattice sites, cf. (4.5).

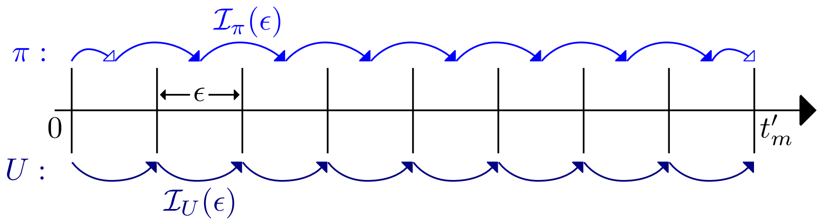


Figure 4.1: Illustration of the leapfrog integration.

4. The candidate configuration (U', π') is accepted or rejected just as in the Metropolis algorithm, with an acceptance probability

$$T_A(\pi, U \to \pi', U') = \min\left(1, \exp\left(-\Delta H\right)\right) \ , \tag{4.53a}$$

where

$$\Delta H = H[\pi', U', \phi] - H[\pi, U, \phi] \ . \tag{4.53b}$$

We stress here that the pseudofermion field ϕ is not part of the molecular dynamics evolution, but affects the acceptance probability via (4.53b).

After this outline, we now have to define the way the molecular dynamics equations are integrated numerically. The basic method to evolve an initial configuration $(\pi, U) = (\pi(t_m), U(t_m))$ to a new configuration $(\pi', U') = (\pi(t_m + t'_m), U(t_m + t'_m))$ is the *leapfrog integration*. The evolution of (4.49a) and (4.49b) is split into small steps of the size $\epsilon = t'_m / N_m$,

$$\mathcal{I}_\pi(\epsilon): \quad \pi(t_m) \quad \longrightarrow \quad \pi(t_m + \epsilon) = \pi(t_m) - \epsilon F[U, \phi]\big|_{t_m} \tag{4.54a}$$

$$\mathcal{I}_U(\epsilon): \quad U(t_m) \quad \longrightarrow \quad U(t_m + \epsilon) = e^{\epsilon \pi(t_m)} U(t_m) \ , \tag{4.54b}$$

and the gauge field configuration U and the field π are integrated alternatingly. One begins with a half step $I_\pi(\epsilon/2)$, then applies $I_U(\epsilon)$ and $I_\pi(\epsilon)$, both $(N_m - 1)$ times, and closes with another full step $I_U(\epsilon)$ and a half step $I_\pi(\epsilon/2)$:

$$\begin{array}{ccccccccccc} \pi(0) & \rightsquigarrow & \pi(\frac{1}{2}\epsilon) & & \longrightarrow & & \pi(\frac{3}{2}\epsilon) & \longrightarrow & \ldots & \rightsquigarrow & \pi(N_m\epsilon) \\ U(0) & & & \longrightarrow & & U(\epsilon) & & \longrightarrow & \ldots \quad \ldots & \longrightarrow & U(N_m\epsilon) \ . \end{array} \tag{4.55}$$

The procedure is illustrated in fig. 4.1. It is equivalent to an N_m-fold application of the elementary leapfrog integrator

$$\mathrm{LF}(\epsilon) = \mathcal{I}_\pi(\frac{1}{2}\epsilon)\ \mathcal{I}_U(\epsilon)\ \mathcal{I}_\pi(\frac{1}{2}\epsilon)\ , \tag{4.56}$$

and the interlacing of (4.54a) and (4.54b) leads to an error of only $O(\epsilon^3)$ in the integration of such a step. It ensures *reversibility* of the Monte Carlo step, as the selection probability $T_S = T_m$ in the molecular dynamics evolution follows

$$T_m(\pi, U \to \pi', U') = T_m(-\pi', U' \to -\pi, U)\ , \tag{4.57}$$

i.e. a configuration evolved from U to U' by integration of the MD equations with conjugate momentum π is evolved back to itself if the opposite conjugate momentum $-\pi$ is used in the subsequent step. This can easily be verified by explicit application of the leapfrog integration steps (4.54a) and (4.54b). Reversibility is crucial for the fulfillment of the detailed balance condition, the proof of which we address in the following. The total transition probability for the step $U \to U'$ is obtained by integrating over all π, π' (we ignore the pseudofermion fields ϕ here):

$$T(U \to U') = \int D[\pi]\mathcal{D}[\pi']\ T_A(\pi, U \to \pi', U')\ T_m(\pi, U \to \pi', U')\ e^{-\frac{1}{2}\pi^2}\ . \tag{4.58}$$

The acceptance probability (4.53a) may be reformulated as

$$\begin{aligned}
T_A(\pi, U \to \pi', U') &= \min\left(1, \frac{e^{-\pi'^2 - S[U']}}{e^{-\pi^2 - S[U]}}\right) \\
&= \frac{e^{-\pi'^2 - S[U']}}{e^{-\pi^2 - S[U]}} \cdot \min\left(1, \frac{e^{-(-\pi)^2 - S[U]}}{e^{-(-\pi')^2 - S[U']}}\right) \\
&= \frac{e^{-\pi'^2 - S[U']}}{e^{-\pi^2 - S[U]}} \cdot T_A(-\pi', U' \to -\pi, U)\ .
\end{aligned} \tag{4.59}$$

Using this together with the reversibility of the molecular dynamics evolution, (4.57), the total transition probability (4.58) multiplied by $e^{-S[U]}$ becomes

$$\begin{aligned}
&e^{-S[U]}T(U \to U') \\
&= \int D[\pi]\mathcal{D}[\pi']\ T_A(-\pi', U' \to -\pi, U)\ T_m(-\pi', U' \to -\pi, U)\ e^{-\frac{1}{2}\pi'^2 - S[U']} \\
&= \int D[\pi]\mathcal{D}[\pi']\ T_A(\pi', U' \to \pi, U)\ T_m(\pi', U' \to \pi, U)\ e^{-\frac{1}{2}\pi'^2 - S[U']} \\
&= e^{-S[U']}T(U' \to U)\ ,
\end{aligned} \tag{4.60}$$

where in the second step the invariance of the integration measure under sign change was used. Hence, the detailed balance condition (4.25) is indeed fulfilled, and the HMC is an exact algorithm.

The acceptance rate P_{acc} of the algorithm is governed by the number N_m of integration steps by which the integration interval t'_m is divided. A fine-grit integration with large N_m leads to large P_{acc}, but is also associated with large computational costs. These effects need to be balanced in order for the algorithm to be most efficient. The optimal N_m depends on the parameters of the simulation, like e.g. the lattice size L/a, the bare parameters and the number N_f of dynamical fermions in the theory. Even more importantly, it would differ substantially for the contributions of the gauge and pseudofermion action to the force $F[U,\phi]$ in (4.54a) if these were considered in isolation. The way individual step scales for the different forces can be achieved will be described in sec. 4.9.

We finally remark that, although the HMC is currently the most widely used algorithm in simulations of dynamical fermions, it may also effectively be applied in the quenched approximation or pure gauge theory. In the upcoming section, we describe an extension of the HMC which allows the inclusion of the strange and charm quark in Monte Carlo simulations.

4.6 Rational Hybrid Monte Carlo

In sec. 4.4, we have argued that the combination $D_u D_u^\dagger$ of the operator D_u and its adjoint allows to implement the fermion determinant for the up and down quarks by pseudofermions ϕ. For the strange and the charm quark there is no (nearly) mass-degenerate other quark flavor which would allow for such an approach. However, the relatively large masses for those quarks lead to a substantial mass gap in the eigenvalue spectrum, such that the weight of configurations in the functional integral that lead to a negative determinant is negligible. Consequently, one may replace D_f ($f = s, c$ in the following) by the hermitean, non-negative operator $|Q_f|$, defined by

$$D_f \;\to\; |Q_f| = \left(Q_f Q_f^\dagger\right)^{1/2} \quad \text{with} \quad Q_f = \gamma_5 D_f \,, \tag{4.61}$$

without changing the outcome of the simulations. This allows to replace the respective fermion determinant by a pseudofermion integral again (cf. (4.37)), which under use of (4.61) may be written as

$$\det[D_f] \sim \int \mathcal{D}[\phi]\; e^{-\phi^\dagger |Q_f|^{-1}\phi}\;. \tag{4.62}$$

However, the operator combination $|Q_f|^{-1}$ is not directly applicable as it contains a square root, which is why it gets approximated by a rational function[8] R_f, such that $|Q_f| R_f \approx$ const. Instead of (4.62), one then uses

$$\det[D_f] = W_f \cdot \det[R_f]^{-1}\;, \tag{4.63a}$$

where the pseudofermion integration

$$\det[R_f]^{-1} \sim \int \mathcal{D}[\phi]\; e^{-S_{PF,f}[U,\phi]} \tag{4.63b}$$

with

$$S_{PF,f}[U,\phi_1,\phi_2] = \phi^\dagger\; R_f\; \phi \tag{4.63c}$$

is included in the path integral, whereas the correction term

$$W_f = \det[D_f R_f] \sim \int \mathcal{D}[\phi']\; e^{-\phi'^\dagger \left(|Q_f| R_f\right)^{-1}\phi'} \tag{4.63d}$$

is treated as a reweighting factor. This means that the path integral expectation values are evaluated in the modified theory with the replacement $\det[D_f] \to \det[R_f]^{-1}$, $\langle\cdot\rangle_{\text{mod}}$, from which the original path integral expectation value $\langle\cdot\rangle$ can be regained via

$$\langle O\rangle = \frac{\langle OW\rangle_{\text{mod}}}{\langle W\rangle_{\text{mod}}}\;. \tag{4.64}$$

Here, $W = \prod_f W_f$ is the total reweighting factor for all quark flavors f that are implemented in the simulation as described above. This approach is called the *Rational Hybrid Monte Carlo (RHMC)* [74, 75]. The rational functions $R_f \approx (Q_f Q_f^\dagger)^{-1/2}$ can be built from the Zolotarev rational function of degree n,

$$R_{n,\epsilon}(y) = A_\epsilon\, \frac{(y+a_{1,\epsilon})(y+a_{3,\epsilon})\ldots(y+a_{2n-1,\epsilon})}{(y+a_{2,\epsilon})(y+a_{4,\epsilon})\ldots(y+a_{2n,\epsilon})}\;, \tag{4.65}$$

---

[8] Other approximations, e.g. by polynomial functions [73] are possible, too.

which provides an optimal approximation of the function $y^{-1/2}$ in the range $y \in [\epsilon, 1]$ for some $\epsilon > 0$. The corresponding (positive) coefficients $\{A_\epsilon, a_{1,\epsilon}, \dots, a_{2n,\epsilon}\}$ as well as the relative deviation[9]

$$\delta = \max_{\epsilon \leq y \leq 1} \left| 1 - y^{1/2} R_{n,\epsilon}(y) \right| \tag{4.66}$$

that they minimize may be found analytically. If the spectral range of $Q_f Q_f^\dagger$ is given by $[r_{f,a}^2, r_{f,b}^2]$ with $r_{f,a} > 0$, the optimal rational approximation R_f for $(Q_f Q_f^\dagger)^{-1/2}$ reads[10]

$$R_f = r_{f,b}^{-1}\, R_{n,\epsilon}(r_{f,b}^{-2} Q_f Q_f^\dagger)\ , \tag{4.67}$$

with $\epsilon = (r_{f,a}/r_{f,b})^2$. An example which demonstrates the approximation of the exact function by R_f, taken from a real RHMC simulation, can be found in fig. 4.2. The accuracy of the rational approximation manifests itself in the smallness of fluctuations of the reweighting factors W_f (cf. (4.63d)), which may be estimated stochastically [76, 77] by

$$W_{f,N} = \frac{1}{N} \sum_{j=1}^{N} e^{-\eta_{j,e}^\dagger \left(|Q_f| R_f\right)^{-1} \eta_{j,e}} \xrightarrow{N\to\infty} W_f \tag{4.68}$$

under use of N independent, random, and normally distributed pseudofermion fields $\eta_{j,e}$ defined on the even lattice sites[11]. It may be argued [77] that a single random source, $N = 1$, is already sufficient for a proper estimation of W_f if the maximal relative deviation δ obeys

$$12\ N_e\ \delta^2 \lesssim 10^{-4}\ , \tag{4.69}$$

where N_e denotes the number of even lattice sites[12].

With the RHMC for the strange and charm quark at hand, one is finally able to evaluate the full path integral (4.12) for the general case of $N_f = 2 + 1 + 1$ quark

---

[9] We follow standard notation here. Note however, that the relative deviation δ actually depends on the degree n of the Zolotarev rational function, i.e. $\delta \equiv \delta_n$.

[10] $R_f \equiv R_{n,f}$, see footnote 9.

[11] This statement refers to the case where even-odd preconditioning is applied to the fermion determinant, see (4.96) in sec. 4.7.

[12] See footnote 11.

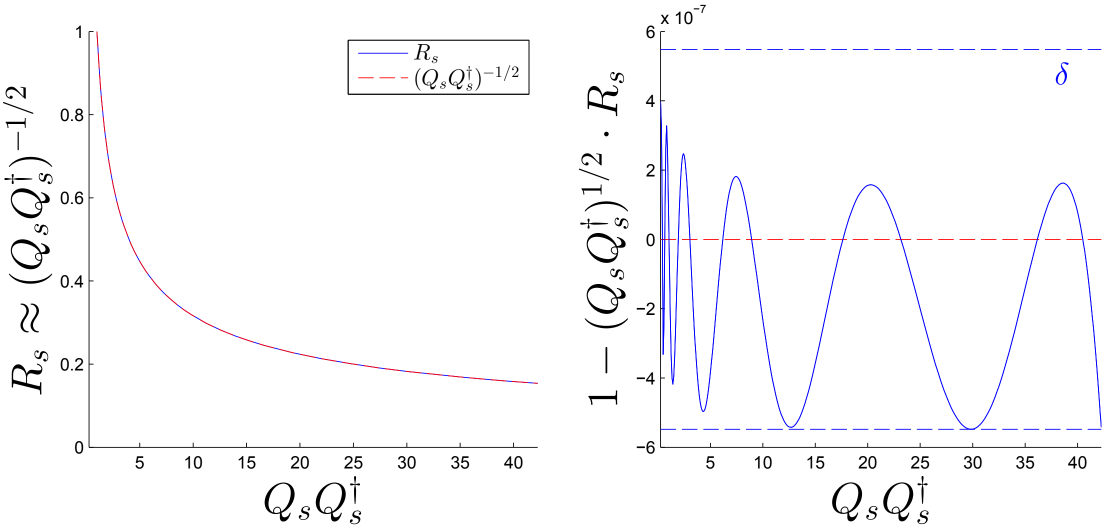


Figure 4.2: *Left*: Zolotarev rational approximation R_s ((4.67), blue) of degree $n = 6$ for $(Q_sQ_s^\dagger)^{-1/2}$ (red) as it appears in the pseudofermion representation of the strange quark determinant, see (4.62). The operator $Q_sQ_s^\dagger$ is formally replaced by a real variable, and the example is taken from an actual simulation on a $T/a = L/a = 12$ lattice with a dynamical strange-like quark (cf. sec. 9). The spectrum of $Q_sQ_s^\dagger$ was found to lie between $r_{s,a}^2 = 0.5^2$ and $r_{s,b}^2 = 6.5^2$, which is also the range of the real variable $Q_sQ_s^\dagger$ displayed above. *Right*: Relative deviation $1-(Q_sQ_s^\dagger)^{1/2}R_s$ of the approximation R_s from the exact expression $(Q_sQ_s^\dagger)^{-1/2}$, as it appears in (4.66). The dashed blue lines represent $\pm\delta$.

flavors with degenerate up and down quark:

$$\langle O\rangle \;\equiv\; \langle O\rangle_{N_f=2+1+1} = \frac{\langle OW\rangle_{N_f=2+1+1,\text{ mod}}}{\langle W\rangle_{\text{mod}}}\;, \tag{4.70}$$

with

$$\langle O\rangle_{N_f=2+1+1,\text{ mod}} = \frac{1}{Z}\int \mathcal{D}[U]\mathcal{D}[\phi]\; e^{-S[U,\phi]}\;\; \langle O\rangle_F\,[U]\;. \tag{4.71}$$

The action is split into the different contributions (cf. (3.36, 3.112, 4.38, 4.63c))

$$S[U,\phi] = S_G[U] + S_{PF,ud}[U,\phi_{ud}] + S_{PF,s}[U,\phi_s] + S_{PF,c}[U,\phi_c] \tag{4.72}$$

with a number of different pseudofermion fields whose measures are all collectively abbreviated by $\mathcal{D}[\phi]$ here, such that the partition function as before reads

$$Z = \int \mathcal{D}[U]\mathcal{D}[\phi]\; e^{-S[U,\phi]}\;. \tag{4.73}$$

Note that in applications where the strange and the charm quark are mass-degenerate (for instance in the massless $N_f = 4$ theory), they can be treated just like the up and down quark as described in sec. 4.4, without the need to use the RHMC.

The remainder of this section is mainly devoted to the question of optimizing the efficiency of the algorithm, sec. 4.7-4.9, and will be concluded with the topic of autocorrelations and error estimation in sec. 4.10.

4.7 Solver and even-odd preconditioning

The Monte Carlo simulations outlined in the previous sections involve the multiple computation of the force $F[U,\phi]$ as it appears in (4.49a) and (4.54a). The pseudofermion parts of the force (cf. (4.72)), in particular, contain derivatives of either the inverse massive Dirac operator D_f^{-1} ($f \in \{u,d\}$, cf. (4.38)), or the rational approximation R_f ($f \in \{s,c\}$, cf. (4.63c)), applied on the pseudofermion field ϕ, respectively. In both cases, Dirac equations need to be solved. While this is obvious for $f \in \{u,d\}$, it may be readily seen for $f \in \{s,c\}$ if the Zolotarev rational function (4.65) in R_f (4.67) is expanded in partial fractions,

$$R_{n,\epsilon}(y) = A\;\left(1 + \frac{r_2}{y+a_2} + \ldots + \frac{r_{2n}}{y+a_{2n}}\right)\;, \tag{4.74}$$

with positive residues r_{2k} due to $A > 0$ and $a_1 > a_2 > \ldots > a_{2n} > 0$ (see e.g. [78]). The application of the different partial fractions with $y = r_{f,b}^{-2} Q_f Q_f^\dagger$ (cf. (4.67)) in the denominators then amounts to finding the solutions ψ_k of

$$\left(Q_f Q_f^\dagger + r_{f,b}^2 a_{2k} \right) \psi_k = r_{f,b}^2 \phi \,, \qquad k \in \{1, \ldots, n\} \,. \tag{4.75}$$

Solving the Dirac equations is the most time-consuming part of the algorithm. Consequently, a lot of effort has been put into the advancement of suitable solvers and the optimization of their efficiency. We will only cover a few basics here and refer to the literature, e.g. [79], for details.

Solver

We consider the general Dirac equation

$$A_f \psi = \eta \,, \tag{4.76}$$

and aim to solve it for ψ. In accordance with the discussion above, A_f may be the massive Dirac operator D_f of flavor f with its $12|X| \times 12|X|$ components (cf. (4.11)), or the combination of the hermitean counterpart $Q_f = \gamma_5 D_f$ (cf. (4.61)) as it appears on the left hand side of (4.75) in the framework of the RHMC. The most effective approach to a large linear system like (4.76) is to solve it iteratively, i.e. by the successive creation of approximate solutions $\psi^{(1)}, \psi^{(2)}, \ldots$, built such that $\psi^{(n)}$ minimizes the residue

$$r^{(n)} = A_f \psi^{(n)} - \eta \tag{4.77}$$

in Krylov space $\mathcal{K}_n$, defined by

$$\mathcal{K}_n = \mathrm{span} \left(\eta, \; A_f \eta, \ldots, \left(A_f\right)^n \eta \right) \,. \tag{4.78}$$

The procedure is stopped once the convergence criterion

$$||r^{(\mathrm{final})}|| < \epsilon \; ||\eta|| \tag{4.79}$$

is fulfilled for some $\psi^{(k)} = \psi^{(\mathrm{final})}$ after the k-th iteration. The convergence rate is governed by the *condition number*, which in the case of $A_f = D_f$ is given by the largest

Abbreviation	Prerequisite for A_f	Minimized norm	recurrence
CG	hermitean, positive definite	$\|\| \cdot \|\|_{A_f}$	short
CR	hermitean	$\|\| \cdot \|\|_2$	short
GCR	none (of the above)	$\|\| \cdot \|\|_2$	long

Table 4.1: Overview on some iterative solvers which may be employed in Lattice QCD. The abbreviations CG, CR and GCR stand for conjugate gradient, conjugate residual and generalized conjugate residual, respectively.

and smallest eigenvalues $\lambda_{\max}$ and $\lambda_{\min}$ of the operator $D_f^\dagger D_f$:

$$\kappa(D_f) = \sqrt{\lambda_{\max}/\lambda_{\min}}\ . \tag{4.80}$$

This becomes an issue for both small lattice spacings and small quark masses, as the condition number for D_f is proportional to the inverse of those:

$$\kappa(D_f) \sim 1/(am_f)\ . \tag{4.81}$$

Together with the choice of ϵ, it also determines a bound for the relative deviation of the found approximate solution from the exact one,

$$\frac{||\psi - \psi^{(\text{final})}||}{||\psi||} < \epsilon\ \kappa(A_f)\ . \tag{4.82}$$

Hence, in total, a large condition number makes more iterations necessary and limits the accuracy of any approximate solution (together with machine precision).

There is a list of solvers that are slight variations of one another, suitable for different cases depending on the properties of the matrix to solve on the left hand side of (4.76). An overview is given in tab. 4.1. The *conjugate gradient* (CG) algorithm employs the function

$$Q(x) = \frac{1}{2}\langle x, x\rangle_{A_f} - \langle x, \eta\rangle\ , \tag{4.83}$$

whose minimum is the solution ψ sought for. Here, $\langle \cdot, \cdot \rangle$ represent the standard scalar product and $\langle x, y\rangle_{A_f} = \langle A_f x, y\rangle$ the one which corresponds to A_f. An iteration starts from the last approximate solution $x^{(i)}$ (or an initial guess $x^{(0)}$). One calculates the corresponding residue $r^{(0)}$, and constructs the search direction vector

$$p^{(i)} = r^{(i)} + \beta_{i-1} p^{(i-1)} \tag{4.84}$$

with

$$\beta_{i-1} = \begin{cases} \frac{\left\langle r^{(i)}, r^{(i)} \right\rangle}{\left\langle r^{(i-1)}, r^{(i-1)} \right\rangle} , & i \geq 1 \\ 0 , & i = 0 . \end{cases} \tag{4.85}$$

The next approximate solution in the iteration is taken to be

$$x^{(i+1)} = x^{(i)} + \alpha_i \, p^{(i)} , \tag{4.86}$$

where

$$\alpha_i = \frac{\left\langle r^{(i)}, r^{(i)} \right\rangle}{\left\langle A_f p^{(i)}, p^{(i)} \right\rangle} \tag{4.87}$$

is determined by the condition that that $Q(x^{(i)} + \alpha_i \, p^{(i)})$ is minimized. The algorithm gets its name from the fact that the residue r is the *gradient* of the function Q. Adding the β_{i-1} term in (4.84) ensures that the p's are *conjugate* to one another, i.e. $p^{(i)} \perp Ap^{(j)}$ $\forall \, j \neq i$, which is a requirement for (4.87) to exist.

The only changes of the *conjugate residual* (CR) with respect to the CG are

$$\beta_{i-1} = \frac{\left\langle \vec{r}^{(i)}, A_f \vec{r}^{(i)} \right\rangle}{\left\langle \vec{r}^{(i-1)}, A_f \vec{r}^{(i-1)} \right\rangle} , \quad i \geq 1 , \tag{4.88a}$$

$$\text{and} \qquad \alpha_i = \frac{\left\langle \vec{r}^{(i)}, A_f \vec{r}^{(i)} \right\rangle}{\left\langle A_f \vec{p}^{(i)}, A_f \vec{p}^{(i)} \right\rangle} . \tag{4.88b}$$

Here, it is the residuals that are conjugate to one another instead of the search direction vectors. While the CG works for a positive definite, hermitean matrix and is optimized for the minimization of the scalar product $\langle . \rangle_{A_f}$, the CR does not require positive definiteness, and is optimized for the scalar product $\langle . \rangle$. The *generalized conjugate residual* (GCR) has no prerequisites on the matrix, but the solutions $\psi^{(n)}$ are determined recursively, which is why in practice the algorithm needs to be restarted with the last approximate solution from time to time. Note that the above methods only require the application of A_f on a vector as well as the determination of scalar products.

They can be combined with *deflation*. The guiding principle of this method is to separate the low and high modes of the operator and treat them differently. The *deflated system* with the high modes can be solved as before, but is equipped with an improved condition number since its minimal eigenvalue is larger compared to the original system, see (4.80). The solution of the *deflation system* is found otherwise, in the case of a hermitean A_f in principle e.g. by explicit calculation of the eigenvectors. We refer to [78] for details.

Finally, we note that equations like (4.75) can be solved for all k simultaneously with so-called *multi-shift solvers* [80], as the right hand sides of the equations are the same. A review of this matter is given in [77].

Even-odd preconditioning

The solving of the Dirac equation (4.76) can be accelerated by the use of preconditioning. In general[13], this involves a matrix M that is cheap to invert and similar to A_f, e.g. in the sense that the small eigenvalues of M and A_f are the same. In that case, the spectral properties (4.80) of $M^{-1}A_f$ are more favorable than those of A_f, which leads to a faster convergence towards the solution of the preconditioned Dirac equation:

$$M^{-1}A_f\,\psi = M^{-1}\,\eta \tag{4.89}$$

A particular type of preconditioning which makes use of the fact that the Dirac operator is a sparse matrix with non-vanishing contributions only for identical or neighboring lattice sites (cf. (3.69)), is the so-called *even-odd preconditioning*. Classifying the lattice sites as even or odd, depending on the sum of their respective coordinates, $\sum_{i=1}^{4} x_i$, the sites are reordered such that the Dirac matrix and the vectors in (4.89) read

$$A_f = \begin{pmatrix} A_f^{ee} & A_f^{eo} \\ A_f^{oe} & A_f^{oo} \end{pmatrix} , \qquad \psi = \begin{pmatrix} \psi^e \\ \psi^o \end{pmatrix} , \qquad \eta = \begin{pmatrix} \eta^e \\ \eta^o \end{pmatrix} . \tag{4.90}$$

While the off-diagonal parts of A_f incorporate the hopping terms, (3.69d), the diagonal parts A_f^{ee} and A_f^{oo} only contain the mass part of the Dirac operator, (3.69b). As such, they are themselves diagonal and trivially invertible. The even-odd preconditioning

---

[13] We consider only left-preconditioning here.

matrix

$$M^{-1} = \begin{pmatrix} \mathbb{1} & -A_f^{eo}(A_f^{oo})^{-1} \\ 0 & 0 \end{pmatrix} \tag{4.91}$$

is hence cheap to build, and the combination $M^{-1}A_f$ on the left hand side of (4.89) assumes the form

$$M^{-1}A_f = \begin{pmatrix} \hat{A}_f & 0 \\ 0 & 0 \end{pmatrix} , \tag{4.92}$$

with

$$\hat{A}_f = A_f^{ee} - A_f^{eo}(A_f^{oo})^{-1}A_f^{oe} \, . \tag{4.93}$$

The preconditioned Dirac equation (4.89) thus becomes an equation only for ψ^e,

$$\hat{A}_f \, \psi^e = \eta^e - A_f^{eo}(A_f^{oo})^{-1} \, \eta^o \, . \tag{4.94}$$

Once the solution ψ^e of this equation is found, it leads to the remaining components ψ^o via the second row of the original Dirac equation (4.76),

$$\psi^o = (A_f^{oo})^{-1} \, \left(\eta^o - A_f^{oe}\psi^e\right) \, . \tag{4.95}$$

Apart from its use in solvers, even-odd preconditioning may also serve to improve the performance of the HMC algorithm. To this end, the determinant of the Dirac operator is decomposed according to

$$\det\left[A_f\right] = \det\left[A_f^{ee}\right] \, \det\left[A_f^{oo}\right] \, \det\left[(A_f^{ee})^{-1}\hat{A}_f\right] \, , \tag{4.96}$$

where the first two determinants are easy to evaluate.

4.8 Hasenbusch frequency splitting

A way to improve the performance of the pseudofermion integration for the up and down quark, (4.37), is to introduce a mass parameter μ^2 and split the corresponding determinant (4.36) according to [81, 82]

$$\det\left[D_u D_u^\dagger\right] = \det\left[D_u D_u^\dagger + \mu^2\right] \cdot \det\left[\frac{D_u D_u^\dagger}{D_u D_u^\dagger + \mu^2}\right] . \tag{4.97}$$

Each of the determinant factors gets implemented by different pseudofermion fields ϕ_1 and ϕ_2 with the action (cf. (4.38)):

$$S_{PF}[U,\phi] = \phi_1^\dagger (D_u D_u^\dagger + \mu^2)^{-1}\phi_1 + \phi_2^\dagger \underbrace{(D_u D_u^\dagger + \mu^2)(D_u D_u^\dagger)^{-1}}_{\mathbb{1}+\mu^2(D_u D_u^\dagger)^{-1}} \phi_2 . \tag{4.98}$$

The factorization separates the contributions of eigenvalues λ of $DD^\dagger$ which are larger than μ^2 from those which are smaller, and improves the condition number (4.80) for both factors with respect to the original pseudofermion representation. As this leads to smaller residues (4.77), the splitting reduces the fluctuations of the pseudofermion force along the molecular dynamics trajectory, and thus allows for coarser integration steps associated with lower computational costs at constant numerical error. The parameter μ^2 may be tuned for efficiency, and depending on the specifics of the simulation, it may also be useful to split the pseudofermion determinant in more than two factors. Moreover, as was pointed out in [83], the contributions from the different factors are quite different in size, which may be made use of in the framework of multiple time scale integration, to be described in sec. 4.9. Since the term $(D_u D_u^\dagger + \mu^2)$ in the ϕ_2 pseudofermion action plays the role of the preconditioning matrix M^{-1} in (4.89), the approach is also called *mass preconditioning*.

Twisted mass reweighting

Finally, we remark that there exists a related method called *twisted mass reweighting* [76]. The Wilson–Dirac operator is not protected against near-zero modes in the case of very light up and down quarks, which may lead to instabilities in the HMC algorithm. This can be averted by introduction of an infrared regulator μ in a similar manner to (4.97), or variations thereof:

$$\det\left[D_u^\dagger D_u\right] \to \det\left[D_u^\dagger D_u + \mu^2\right] \tag{4.99}$$

$$\text{or} \qquad \det\left[D_u^\dagger D_u\right] \to \det\left[(D_u^\dagger D_u + \mu^2)^2 (D_u^\dagger D_u + 2\mu^2)^{-1}\right] . \tag{4.100}$$

While the modified determinant enters the HMC and improves its performance, the term which corrects for this,

$$W_1 = \det\left[D_u^\dagger D_u \cdot (D_u^\dagger D_u + \mu^2)^{-1}\right] \tag{4.101}$$

$$\text{or} \qquad W_2 = \det\left[D_u^\dagger D_u \cdot (D_u^\dagger D_u + \mu^2)^{-2}(D_u^\dagger D_u + 2\mu^2)\right] \, , \tag{4.102}$$

can be estimated stochastically in a manner similar to the case of the RHMC (cf. (4.68)). In contrast to the Hasenbusch frequency splitting, W_1 or W_2 is also treated as a reweighting factor (cf. (4.64)), whose fluctuations are modest in case that μ is chosen appropriately. This holds especially for W_2 as the contributions of modes of the Dirac operator with large eigenvalues are strongly suppressed [76].

4.9 Multiple time scale integration

In sec. 4.5, we have discussed the leapfrog method which integrates the molecular dynamics equation for the conjugate momenta π and the gauge configuration U alternatingly. This method can be refined in two ways in order to make the integration more efficient. The first one is *multiple time scale integration* [84]. The force $F[U,\phi]$ which determines the update of π, see (4.54a), gets contributions from different sources, e.g the gauge action, the pseudofermion action of the up and down quark, the rationally approximated pseudofermion action, and different parts of the fermion determinant in case Hasenbusch frequency splitting (sec. 4.8) is applied. The contributions of these different actions are usually not only of different orders of magnitude, but also differ substantially regarding the effort to compute them (sec. 4.7). Hence, it is generally advantageous to integrate the different forces on individual time scales. To this end, one divides the integration interval t'_m hierarchically, with step sizes

$$\epsilon_2 = t'_m / N_2 \tag{4.103a}$$

$$\epsilon_1 = t'_m / (N_1 N_2) \tag{4.103b}$$

$$\epsilon_0 = t'_m / (N_0 N_1 N_2) \, . \tag{4.103c}$$

As the gauge force is the main contribution, it is integrated with the finest step size ϵ_0. The other forces are smaller, but more expensive to compute, so they are usually distributed on the other scales. A whole Monte Carlo step with the leapfrog integration on multiple time scales is performed by the following sequence of elementary

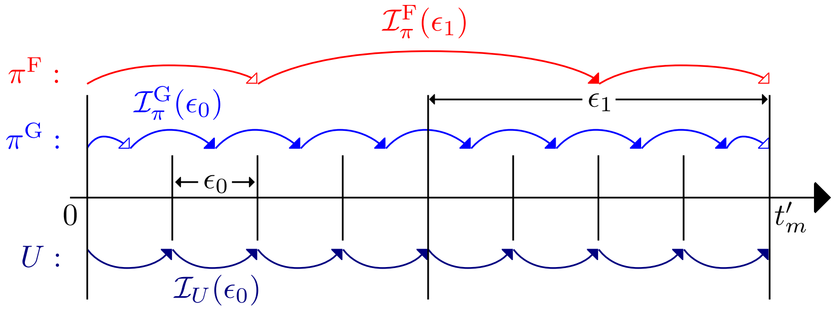


Figure 4.3: Illustration of multiple time scale integration on two levels with $N_1 = 2$, $N_0 = 4$ and the leapfrog as elementary integrator. This corresponds to (4.104) without the occurrence of the coarsest level $(\mathcal{I}_\pi(\frac{1}{2}\epsilon_2), N_2 \to 1)$. As an example, the pseudofermion forces are integrated on the coarse scale (ϵ_1, red) and the gauge force on the fine scale (ϵ_0, blue), indicated by the superscripts F and G. Compare to fig.4.1.

integrators (cf. (4.56)):

$$\left\{\mathcal{I}_\pi(\frac{1}{2}\epsilon_2)\ \left\{\mathcal{I}_\pi(\frac{1}{2}\epsilon_1)\ \left\{\mathcal{I}_\pi(\frac{1}{2}\epsilon_0)\ \mathcal{I}_U(\epsilon_0)\ \mathcal{I}_\pi(\frac{1}{2}\epsilon_0)\right\}^{N_0}\ \mathcal{I}_\pi(\frac{1}{2}\epsilon_1)\right\}^{N_1}\ \mathcal{I}_\pi(\frac{1}{2}\epsilon_2)\right\}^{N_2} . \tag{4.104}$$

This scheme is also depicted in fig. 4.3. The second way to possibly improve the efficiency of the integration is the use of more sophisticated elementary integrators than the leapfrog. One of them is the so-called Omelyan–Mryglod–Folk (OMF) integrator of 2nd order [85]:

$$\mathrm{OMF}_{2,\lambda}(\epsilon) = \mathcal{I}_\pi(\lambda\epsilon)\ \mathcal{I}_U(\frac{1}{2}\epsilon)\ \mathcal{I}_\pi((1-2\lambda)\epsilon)\ \mathcal{I}_U(\frac{1}{2}\epsilon)\ \mathcal{I}_\pi(\lambda\epsilon)\ . \tag{4.105}$$

This is a generalization of the leapfrog, to which it reduces for the special cases of $\lambda = 1/2$ and $\lambda = 1/4$:

$$\mathrm{OMF}_{2,1/2}(\epsilon) = \ \mathrm{LF}(\epsilon) \tag{4.106a}$$

$$\mathrm{OMF}_{2,1/4}(\epsilon) = \{\mathrm{LF}(\epsilon/2)\}^2\ . \tag{4.106b}$$

Note also that, for any $0 < \lambda < 1/2$, the OMF integrator of 2nd order updates the gauge field twice. Furthermore, an OMF integrator of 4th order may as well be applied, which updates the gauge field five times.

We complete our discussion with the remark that the methods to increase the efficiency of the HMC algorithm discussed here, namely even-odd preconditioning, Hasenbusch frequency splitting, twisted mass reweighting, and multiple time scale integration all go hand in hand and may well be used together. There exist further techniques (e.g. domain decomposition [86], cf. sec. 9.5) which are not covered here as they find no application in this work.

4.10 Autocorrelation and error estimation

So far, we have described in detail the way the estimator (cf. (4.19))

$$\overline{O} = \frac{1}{N_\mathrm{cf}} \sum_{n=1}^{N_\mathrm{cf}} O[U^{(n)}] \tag{4.107}$$

for the path integral average $\langle O \rangle$ (cf. (3.46)) is obtained from an ensemble $\{U^{(1)}, \ldots, U^{(N_\mathrm{cf})}\}$ of configurations importance sampled with the help of Monte Carlo simulations. We now discuss the error analysis for this Monte Carlo estimator.

If the configurations were statistically independent, the standard deviation or error $\Delta\overline{O}$ of the above estimator would simply be given by

$$\Delta\overline{O} = \frac{1}{\sqrt{N_\mathrm{cf}}}\, \sigma_O\ , \tag{4.108}$$

where

$$\sigma_O^2 = \frac{1}{N_\mathrm{cf}-1} \sum_{i=1}^{N_\mathrm{cf}} \left(O_i - \overline{O}\right)^2 \tag{4.109}$$

is the variance of the ensemble. However, while the steps in a Markov chain (cf. (4.20)) are independent in the sense that a produced configuration U' can only be traced back to its direct predecessor and does not know about any previous configurations, the length of a Monte Carlo trajectory in configuration space is limited. In general this leads to a correlation of subsequently produced configurations, which is reflected to different extent in different observables O. Since it is a correlation between configurations of one and the same Markov chain, one speaks of *autocorrelation*. It increases the uncertainty ΔO for the estimate and needs to be taken into account in the error analysis. A quantitative treatment is provided by the *autocorrelation function* for the observable O,

$$\Gamma_O(t_m) = \langle O^{(i)}\, O^{(i+t_m)}\rangle - \langle O^{(i)}\rangle\langle O^{(i+t_m)}\rangle \ , \tag{4.110}$$

where the notation $O^{(i)} = O[U^{(i)}]$ is used. As indicated, the autocorrelation function is translation invariant, i.e. it depends only on the *difference* t_m of the Monte Carlo time of two configurations. Moreover, note that $\Gamma_O(t_m = 0) = \sigma_O^2$. The effect of autocorrelations can thus be disentangled from ordinary statistical fluctuations by the use of the *normalized autocorrelation function* $\Gamma_O^{\mathrm{norm}}(t_m)$, which for large t_m decreases exponentially:

$$\Gamma_O^{\mathrm{norm}}(t_m) := \frac{\Gamma_O(t_m)}{\Gamma_O(0)} \sim \exp\left(-\frac{t_m}{\tau_{O,\mathrm{exp}}}\right) \ , \tag{4.111}$$

where $\tau_{O,\mathrm{exp}}$ is the *exponential autocorrelation time* for O which encodes the capability of the algorithm to efficiently sample the observable O. For a quantitative analysis of autocorrelations, one uses the related *integrated autocorrelation time*

$$\tau_{O,\mathrm{int}} = \frac{1}{2} + \sum_{t_m=1}^{\infty} \Gamma_O^{\mathrm{norm}}(t_m) \approx \int_0^\infty dt_m\, e^{-t_m/\tau_{O,\mathrm{exp}}} \overset{\tau_{O,\mathrm{exp}}\gg 1}{=} \tau_{O,\mathrm{exp}} \ , \tag{4.112}$$

which can be obtained from the normalized autocorrelation function $\Gamma_O^{\mathrm{norm}}(t_m)$. The error estimate which corrects for the autocorrelation is then given in terms of the integrated autocorrelation time by

$$\Delta\overline{O} = \frac{\sqrt{2\,\tau_{O,\mathrm{int}}}}{\sqrt{N_{\mathrm{cf}}}}\,\sigma_O \ . \tag{4.113}$$

Note that this allows for the interpretation that the ensemble effectively contains $N_{\mathrm{indep.}} = N_{\mathrm{cf}}/\,(2\tau_{O,\mathrm{int}})$ independent configurations.

The inclusion of the autocorrelation with the above formulas relies on the assumption that the autocorrelation function can be explicitly and reliably estimated. A method which accomplishes this is the so-called *gamma method* [87, 88]. It uses

$$\overline{\Gamma}_O(t_m) = \frac{1}{N-t_m}\sum_{t'_m=1}^{N-t_m}\left[O_{t_m+t'_m} - \overline{O}\right]\left[O_{t'_m} - \overline{O}\right] \ . \tag{4.114}$$

as an estimate for the autocorrelation function (4.110). The summation in (4.112) is truncated at a window W, which is chosen such that the statistical error of the estimators (4.114) is balanced against the systematic error from the truncation itself. The optimal choice for W can be automatically obtained under the assumption that

$$\tau_{O,\mathrm{exp}} = S_\tau \cdot \tau_{O,\mathrm{int}} \tag{4.115}$$

for some parameter S_τ. It should be given such that the autocorrelation function (4.114) visibly builds a plateau without summing up too much noise, and can be adjusted if necessary. For most systems and observables, $S_\tau \approx 2$ is a reasonable choice.

The gamma method can be extended to *derived observables* F, which are functions of a number of primary observables like O above, now labelled by an index α:

$$F \equiv f(O_\alpha) \; . \tag{4.116}$$

Estimators $\overline{F}$ for them are obtained from those of the primary observables according to $\overline{F} = f(\overline{O}_\alpha)$, and the corresponding errors can be found under use of the partial derivatives $f_\alpha = \frac{\partial f}{\partial O_\alpha}$ as

$$\Gamma_F(t_m) = \sum_{\alpha\beta} f_\alpha f_\beta \Gamma_{\alpha\beta}(t_m) \tag{4.117}$$

Here, $\Gamma_{\alpha\beta}$ is the obvious extension of the autocorrelation function (4.110) to two different primary observables O_α and O_β,

$$\Gamma_{\alpha\beta}(t_m) = \langle O_\alpha^{(i)} \, O_\beta^{(i+t_m)} \rangle - \langle O_\alpha^{(i)} \rangle \langle O_\beta^{(i+t_m)} \rangle \; , \tag{4.118}$$

and the estimator (4.114) is modified accordingly. Moreover, replica can easily be included, allow for consistency checks and may in addition be used to cancel the leading bias in the estimator for F. For details, we refer to [88]. Finally, we note that in [89], it was shown how contributions from slow modes may be incorporated in the autocorrelation function even if they are not directly detectable (cf. sec. 6.2.2). Other approaches for error analyses, which handle the autocorrelation only implicitly however, are binning, statistical bootstrap and the jackknife method.

5 The Schrödinger functional

The Schrödinger functional (SF) provides a useful tool to study QCD in finite volume. It involves a set of boundary conditions which facilitates perturbative calculations as well as MC simulations, and is particularly suitable as a renormalization (cf. sec. 3.5 and sec. 3.6) and improvement scheme (cf. sec. 3.7).

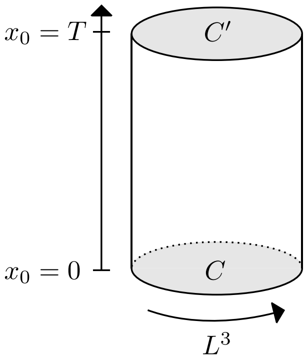


Figure 5.1: The Schrödinger functional and its gauge field boundary conditions.

After a brief summary of its definition and main properties in the continuum (sec. 5.1), we discuss its formulation on the lattice including boundary improvement in sec. 5.2. We then introduce the Schrödinger functional correlation functions (sec. 5.3), which will play an important role in the $O(a)$ improvement of the action, sec. 8 and sec. 9. The presentation mostly follows the original literature [46, 90, 91], as well as the review given in [92].

5.1 The Schrödinger functional in the continuum

Definition and quantum mechanical interpretation

For the Schrödinger functional, spacetime is defined to be a $V = T \times L^3$ cylinder with Dirichlet boundary conditions in the temporal direction and periodic boundary conditions in the spatial directions for both the gauge and fermion fields, see fig. 5.1. We restrict ourselves to the gauge fields A_μ first. At the boundaries in time, they are fixed to

$$A_k(x) = \begin{cases} C_k^\Lambda(\vec{x}) & \text{at } x_0 = 0 \\ C_k'(\vec{x}) & \text{at } x_0 = T \ , \end{cases} \tag{5.1}$$

where C and C' are classical gauge fields, and $\Lambda(\vec{x}) \in SU(3)$ is a time-independent local gauge transformation which transforms the fields according to[1] (cf. (2.13, 3.9))

$$A'_k(\vec{x}) = \Lambda(\vec{x}) A_k(\vec{x}) \Lambda(\vec{x})^\dagger - \left(\partial_\mu \Lambda(\vec{x})\right) \Lambda(\vec{x})^\dagger \, . \tag{5.2}$$

In space, the periodic boundary conditions imply

$$A_k(x + L\hat{k}) = A_k(x) \tag{5.3a}$$

$$\Lambda(\vec{x} + L\hat{k}) = \Lambda(\vec{x}) \tag{5.3b}$$

for $k = 1, 2, 3$, where $\hat{k}$ is a unit vector in the corresponding spatial direction. The partition function

$$Z[C, C'] = \int D[\Lambda] \int D[A] \; \exp\left\{-S_G[A]\right\} \tag{5.4}$$

involves the products $D[A] = \prod_{\vec{x},\mu,a} dA^a_\mu$ and $D[\Lambda] = \prod_{\vec{x}} d\Lambda(\vec{x})$ of Haar measures, and the gauge action $S_G[A]$ as given in (3.12). In the quantum mechanical Schrödinger representation, $Z[C, C']$ can be interpreted (cf. [90]) as the transition amplitude from the state $|C\rangle$ to the state $|C'\rangle$, and the appearance of Λ in (5.1) and (5.4) plays the role of the projector onto the physical (i.e. gauge invariant) subspace of Hilbert space. Accordingly, the partition function is a gauge invariant functional of the boundary gauge fields,

$$Z[C^\Omega, C'^{\Omega'}] = Z[C, C'] \, , \tag{5.5}$$

referred to as the Schrödinger functional.

Induced background field

If the coupling g_0^2 is small, the Schrödinger functional is dominated by fields close to the (in general multiple) minima of the action. Let us assume for the moment that the boundary fields C, C' lead to a *unique* minimum. In that case, the associated *(boundary induced) background field* B_μ is unique as well, up to gauge transformations Ω which preserve the boundary conditions:

$$S_G[A] > S_G[B] \quad \forall A \neq B^\Omega \, , \tag{5.6}$$

---

[1] Note that in accordance with the original literature, we work with anti-hermitean instead of hermitean gauge fields here, see app. A.1.

where B^{Ω} denotes all possible gauge transforms of B. The boundary fields are usually taken to be constant diagonal matrices [90]

$$C_k^{(\prime)} = \frac{i}{L}\,\mathrm{diag}\left(\phi_1^{(\prime)}, \phi_2^{(\prime)}, \phi_3^{(\prime)}\right) \ , \tag{5.7}$$

where the angles $\phi_i^{(\prime)}$ are subject to the conditions

$$\sum_{i=1}^{3} \phi_i = \sum_{i=1}^{3} \phi_i' = 0\ , \tag{5.8}$$

so that the boundary fields are elements of the algebra $\mathfrak{su}(3)$. If they are required to belong to the so-called *fundamental domain*, defined by

$$\phi_1 < \phi_2 < \phi_3 \qquad \text{and} \qquad |\phi_3 - \phi_1| < 2\pi\ , \tag{5.9}$$

the above statement about the uniqueness of the minimal action and the associated background field is fulfilled [90]. The background field induced by (5.7) linearly interpolates between the boundaries,

$$B_0(x) = 0\ , \qquad B_k(x) = \frac{x_0 C_k' + (T - x_0) C_k}{T}\ , \qquad k = 1,2,3\ , \tag{5.10}$$

and represents a constant color-electric field as the only non-vanishing components of the associated field tensor are

$$F_{0k} = \partial_0 B_k = \left(C_k' - C_k\right)/T\ , \qquad k = 1,2,3\ . \tag{5.11}$$

Hence, it is also referred to as *Abelian background field*. The boundary conditions (5.7) preserve spatial translation invariance, a feature which for various reasons is useful in the context of correlation functions formulated by means of the Schrödinger functional, as we will see in detail in sec. 5.3.

The Schrödinger functional as a renormalization scheme

Due to (5.6), the effective action[2]

$$\Gamma[B] = -\log Z[C, C'] \tag{5.12}$$

---

[2] B and C, C' may be used interchangeably as arguments, since there is a one-to-one correspondence between them.

may be perturbatively saddle point expanded for weak couplings g_0^2 around B, according to

$$\Gamma[B] = \frac{1}{g_0^2}\Gamma_0[B] + \Gamma_1[B] + g_0^2\Gamma_1[B] + \ldots \; . \tag{5.13}$$

Under use of this, it was first argued in [90] that the SF is finite after renormalization of the bare parameters[3] (cf. sec. 3.5). This property lays the ground for its use as a *finite volume renormalization scheme*, in which a renormalized coupling may be defined via the variation of the action with the boundary fields. More precisely, with the specific choices [40]

$$\begin{aligned} \phi_1 &= \quad \eta - \frac{\pi}{3} & \qquad \phi_1' &= -\phi_1 - \frac{4\pi}{3} \\ \phi_2 &= -\frac{1}{2}\eta & \qquad \phi_2' &= -\phi_3 + \frac{2\pi}{3} \\ \phi_3 &= -\frac{1}{2}\eta + \frac{\pi}{3} & \qquad \phi_3' &= -\phi_2 + \frac{2\pi}{3} \; , \end{aligned} \tag{5.14}$$

the non-perturbatively defined *Schrödinger functional coupling* reads

$$\bar{g}^2_{\mathrm{SF}}(L) = \left.\frac{\Gamma_0'[B]}{\Gamma'[B]}\right|_{\eta=0} \; , \tag{5.15}$$

where the derivatives are to be understood with respect to η. The numerator of (5.15) is given analytically, and the denominator can be estimated as a path integral expectation value (cf. (5.4, 5.12)),

$$\Gamma'[B] \equiv \frac{\partial \Gamma[B]}{\partial \eta} = \left\langle \frac{\partial S_G}{\partial \eta} \right\rangle \; . \tag{5.16}$$

As indicated in (5.15), the only scale which enters the SF coupling is L. The coupling has prominently been used for the non-perturbative determination of the Λ parameter from hadronic low-energy input within the effort of the ALPHA Collaboration [40–44, 95]. A direct approach on the lattice would have to satisfy

$$a \ll \mu^{-1} \ll m_H^{-1} \ll L \tag{5.17}$$

---

[3] Actually, this has explicitly been shown only in perturbation theory to finite order, see also [93, 94], but it was confirmed by numerous non-perturbative computations.

$$\begin{tikzcd}
L_{\max} = \mathrm{O}(\frac{1}{2}\mathrm{fm}): & \mathrm{HS} \arrow[r] & \mathrm{SF}(\mu = 1/L_{\max}) \arrow[d] \\
 & & \mathrm{SF}(\mu = 2/L_{\max}) \arrow[d] \\
 & & \bullet \\
 & & \bullet \\
 & & \bullet \arrow[d] \\
 & & \mathrm{SF}(\mu = 2^n/L_{\max}) \arrow[d, "\mathrm{PT:}"'] \\
\text{DIS, jet-physics, at } s = M_Z^2 & & \Lambda_{\mathrm{QCD}} \arrow[ll, "\mathrm{PT}"']
\end{tikzcd}$$

Figure 5.2: Illustration of the ALPHA Collaboration strategy to determine the Λ parameter from low-energy phenomenological input under use of the SF coupling (5.15). The figure is taken from [6]. HS, PT and DIS stand for hadronic scheme, perturbation theory and deep inelastic scattering, respectively. The Lambda parameter Λ is denoted by Λ_{QCD}.

in order to avoid large discretization and finite size effects. Direct compliance with these conditions, however, is not feasible regarding the computational costs. This multiscale problem can be overcome by *finite size scaling*, where the renormalization scale is identified with the inverse of the spatial extent of the SF,

$$\mu = 1/L\,. \tag{5.18}$$

The SF coupling is computed at some large $L_{\max}$, which is put in relation to a hadron mass in a hadronic renormalization scheme, as explained in sec. 3.4. Subsequently, $\overline{g}^2_{\mathrm{SF}}(L_{\max})$ is evolved to high energies (i.e. $L_{\min}$) under use of the so-called *step scaling function*, which is a non-perturbatively defined finite step analog of the β function (2.31). The Λ parameter can then be derived from $\overline{g}^2_{\mathrm{SF}}(L_{\min})$, see (2.35). Finally, given the perturbative expansion of the SF coupling, Λ may be converted to the $\overline{\mathrm{MS}}$ scheme via (2.38). The procedure is illustrated in fig. 5.2.

Inclusion of quarks

The SF was extended from the pure Yang–Mills theory to QCD in [91]. The inclusion of quarks in accordance with the quantum mechanical interpretation described above can be obtained by imposing the following Dirichlet boundary conditions in time:

$$P_+\ \psi(x)|_{x_0=0} = \rho(\vec{x}) \qquad\qquad P_-\ \psi(x)|_{x_0=T} = \rho'(\vec{x}) \tag{5.19a}$$

$$\overline{\psi}(x)\ P_-|_{x_0=0} = \overline{\rho}(\vec{x}) \qquad\qquad \overline{\psi}(x)\ P_+|_{x_0=T} = \overline{\rho}'(\vec{x})\,, \tag{5.19b}$$

with $P_\pm = \frac{1}{2}(\mathbb{1} + \gamma_0)$. Only half of the fermion components are specified, as the Dirac equation has a unique solution in that case. In principle, the fermion fields at the boundaries introduce additional counterterms[4] in the action [91]. However, for vanishing boundary fields,

$$\rho = \overline{\rho} = \rho' = \overline{\rho}' = 0\,, \tag{5.20}$$

no further renormalization apart from the one of the bare parameters is required, as before without quarks. The most general boundary conditions in space are given by

$$\psi(x + L\hat{k}) = \exp\{\;\;i\theta_k\}\;\psi(x) \tag{5.21a}$$

$$\overline{\psi}(x + L\hat{k}) = \exp\{-i\theta_k\}\;\overline{\psi}(x)\,. \tag{5.21b}$$

The phases (or fermion angles) θ_k respect the periodicity of bilinears, and are usually chosen identically, $\theta \equiv \theta_1 = \theta_2 = \theta_3$. In that case, the SF reads (cf. (5.4))

$$Z[C, C'] = Z[C, \rho, \overline{\rho}, C', \rho', \overline{\rho}]|_{\rho=\overline{\rho}=\rho'=\overline{\rho}'=0} = \int D[\psi, \overline{\psi}, A]\;\exp\left\{-S[\psi, \overline{\psi}, A]\right\}\,, \tag{5.22}$$

where $S[\psi, \overline{\psi}, A]$ comprises the fermion as well as the gauge action, (2.14) and (2.15), respectively.

It has been pointed out in [91] that the fermionic boundary conditions introduce a finite gap in the spectrum of the Dirac operator. For the special case of $C = C' = 0$ and $\theta = 0$, the smallest eigenvalue of the massless and free squared Dirac operator D^2, which amounts to $D_u^\dagger D_u$ on the lattice (cf. (4.36)), is given by [91]

$$\lambda^2_{\text{min}} = \left(\frac{\pi}{2T}\right)^2\,. \tag{5.23}$$

As explained in sec. 4.7, this largely facilitates the numerical simulation of Wilson fermions (cf. (4.80)), and the associated feasibility to simulate massless fermions allows to use the SF as a *mass-independent renormalization scheme*. Such a scheme has the advantage of mass-independent and thus simpler renormalization group equations (cf. sec. 8), and the fact that all quarks in the theory are mass-degenerate makes their numerical implementation easier. First, only one quark mass has to be tuned, and second, in the case of $N_f = 2, 4$, one may simulate doublets and hence avoid the more expensive use of the RHMC, see sec. 4.6.

---

[4] These appear in QCD in contrast to the pure Yang–Mills theory due to the presence of gauge invariant composite fields of dimension 3.

5.2 The Schrödinger functional on the lattice

Boundary conditions

Given the discussion in the previous section, the formulation of the SF on the lattice is straightforward. The boundary conditions for the fermion fields in time and space stay the same, (5.19) and (5.21). The introduction of the fermion phase angle θ is equivalent to substituting

$$U_k(x) \to \lambda\, U_k(x) \qquad \text{with} \qquad \lambda = \exp\left(i\theta a/L\right) \tag{5.24}$$

in the gauge covariant derivatives ∇_μ and ∇^*_μ (3.24) as they appear in the Wilson–Dirac operator (3.43), see e.g. [46] for details. The boundary conditions for the continuum gauge fields A_μ translate to those of the lattice gauge fields U_μ via

$$U_k(x)|_{x_0=0} = \exp\left\{aC_k\right\} \tag{5.25a}$$

$$U_k(x)|_{x_0=T} = \exp\left\{aC'_k\right\} \ , \tag{5.25b}$$

where $C_k^{(\prime)}$ is the same as in (5.7), and

$$U_\mu(x + L\hat{k}) = U_\mu(x) \ . \tag{5.26}$$

Accordingly, the lattice background field (cf. (5.10)) reads

$$V_\mu(x) = \exp\left\{aB_\mu(x)\right\} \ , \tag{5.27}$$

and the partition function (5.22) assumes the familiar form

$$Z[C, C'] = \int D[\psi, \overline{\psi}, U]\ \exp\left\{-S[\psi, \overline{\psi}, U]\right\} \ . \tag{5.28}$$

Actions and (boundary) improvement in the Schrödinger functional

In order to write down the various actions in the framework of the SF, it is convenient to extend the definition of the fermion and gauge fields to the outside of the SF cylinder, where they are set to 0 and 1, respectively. For the fermion fields, this includes the components at the boundaries which were not explicitly specified in (5.19). In that case, the respective actions may be written again as a sum over all spacetime points, and the Wilson fermion action in fact simply maintains the form given in (3.44). The gauge actions, in contrast, get slightly modified at the boundaries, where the purely spatial plaquettes and—if present—rectangles are weighted with certain factors for convenience, see below. In the formulation of the improved action on the lattice, the clover term (3.101) is restricted to the bulk of the SF:

$$\delta S_{\mathrm{sw}}[\psi,\overline{\psi},U] = a^5 \sum_{x_0=a}^{T-a} \sum_{\vec{x}} \overline{\psi}(x) \frac{i}{4} \sigma_{\mu\nu} F_{\mu\nu} \psi(x) \; . \tag{5.29}$$

For full improvement, however, one also needs to account for the effects of the boundaries in the form of boundary counterterms $\delta S_{G,b}$ and $\delta S_{F,b}$. The full $O(a)$ improved lattice action in the SF then reads

$$S_{\mathrm{I}}^{\mathrm{SF}}[\psi,\overline{\psi},U] = S_G^{\mathrm{W}}[U] + \delta S_{G,b}[U] + S_F[\psi,\overline{\psi},U] + \delta S_{F,b}[\psi,\overline{\psi},U] + \delta S_{\mathrm{sw}}[\psi,\overline{\psi},U] \; . \tag{5.30}$$

To obtain the counterterms, one constructs the most general irrelevant boundary terms in the Symanzik effective action first, similar to the procedure described in sec. 3.7. These are generally of the form [46]

$$S_k = \lim_{\epsilon\to 0} \int d^3\vec{x} \;\; \left\{ \mathcal{B}_d(\vec{x})|_{x_0=\epsilon} + \mathcal{B}'_d(\vec{x})|_{x_0=T-\epsilon} \right\} \;\; , \tag{5.31}$$

where $\mathcal{B}_d^{(\prime)}$ are local composite fields of dimension $3+d$ at the respective boundaries.

The boundary fermion field counterterms that enter $\mathcal{B}_1^{(\prime)}$ can be reduced under use of the equations of motion and symmetry considerations. The corresponding lattice counterterms read [46]

$$\begin{aligned} \delta S_{F,b}[U] = a^4 \sum_{\vec{x}} \Big\{ \; & (\tilde{c}_s - 1) \left[\mathcal{O}_s(\vec{x}) - \mathcal{O}'_s(\vec{x}) \right] \\ & + (\tilde{c}_t - 1) \left[\mathcal{O}_t(\vec{x}) - \mathcal{O}'_t(\vec{x}) \right] \Big\} \;\; , \end{aligned} \tag{5.32}$$

with

$$\mathcal{O}_s(\vec{x}) = \overline{\rho}(\vec{x}) \;\; \gamma_k \widetilde{\nabla}_k \, \rho(\vec{x}) \tag{5.32a}$$

$$\mathcal{O}'_s(\vec{x}) = \overline{\rho}'(\vec{x}) \; \gamma_k \widetilde{\nabla}_k \, \rho'(\vec{x}) \tag{5.32b}$$

$$\mathcal{O}_t(\vec{x}) = \left\{ \overline{\psi}(y) \left(P_+ \nabla_0^* + \overleftarrow{\nabla}_0^* P_- \right) \psi(y) \right\}_{y=(a,\vec{x})} \tag{5.32c}$$

$$\mathcal{O}'_t(\vec{x}) = \left\{ \overline{\psi}(y) \left(P_+ \nabla_0 + \overleftarrow{\nabla}_0 P_- \right) \psi(y) \right\}_{y=(T-a,\vec{x})} \; . \tag{5.32d}$$

The pure gluon boundary fields that correspond to $\mathcal{B}_1^{(\prime)}$ are

$$\mathrm{tr}\left\{F_{kl}F_{kl}\right\} \quad \text{and} \quad \mathrm{tr}\left\{F_{0k}F_{0k}\right\} \,, \tag{5.33}$$

with $k, l = 1, 2, 3$. One way to account for them in the improved lattice theory is, for instance, to add the boundary gauge counterterms [90]

$$\begin{aligned} \delta S_{G,b}[U] = \frac{1}{2g_0^2}\,(c_s-1)\sum_{\mathcal{C}_s}\mathrm{tr}[\mathbb{1}-U(\mathcal{C}_s)] \\ +\frac{1}{g_0^2}\,(c_t-1)\sum_{\mathcal{C}_t}\mathrm{tr}[\mathbb{1}-U(\mathcal{C}_t)]\,, \end{aligned} \tag{5.34}$$

where $\mathcal{C}_s$ and $\mathcal{C}_t$ are space- and time-like oriented plaquettes at the boundaries, respectively. However, as we will see below, discretizations in terms of plaquettes *and* rectangles are also useful when these are present in the original action, which is the case for the Lüscher–Weisz gauge action (cf. sec. 3.8).

While the fermion counterterms in (5.32) are considered in the Wilson–Dirac operator, the boundary improvement terms $\delta S_{G,b}$ can be incorporated in the gauge action S_G by introducing weight factors $w_0(\mathcal{C})$ and $w_1(\mathcal{C})$ for the oriented plaquettes $\mathcal{S}_0$ and rectangles $\mathcal{S}_1$, respectively. The boundary improved SF *Wilson gauge action* (cf. (3.39)) then becomes

$$S_G^{\mathrm{W,SF}}[U] = S_G^{\mathrm{W}}[U] + \delta S_{G,b}[U] = \frac{1}{g_0^2}\sum_{\mathcal{C}\in\mathcal{S}_0} w_0(\mathcal{C})\mathrm{tr}[\mathbb{1}-U(\mathcal{C})]\,, \tag{5.35}$$

where under consideration of (5.34), the weight factor depends on the location of the plaquette $\mathcal{C}$ as follows:

$$w_0(\mathcal{C}) = \begin{cases} \frac{1}{2}c_s(g_0^2) & \text{if } \mathcal{C}\in \text{boundary} \\ c_t(g_0^2) & \text{if exactly 1 link of } \mathcal{C}\in\text{boundary} \\ 1 & \text{otherwise .} \end{cases} \tag{5.36}$$

For $c_s = c_t = 1$, one obtains the unimproved SF Wilson gauge action $S_G^{\mathrm{W}}[U]$. In the case of the Abelian boundary gauge fields (5.7), the spatial counterterms in (5.32) and (5.34) do not contribute, i.e. improvement is compatible with $c_s = \tilde{c}_s = 1$. Concerning the temporal improvement coefficients, 2-loop results [90, 94, 96, 97] for c_t and 1-loop results [50, 53] for $\widetilde{c}_t$ are available:

$$\begin{aligned} c_t(g_0^2) &= 1 + \left[-0.08900(5) + 0.0191410(1)\ N_f\right] g_0^2 \\ &\qquad + \left[-0.0294(3) + 0.002(1)\ N_f + 0.0000(1)\ N_f^2\right] g_0^4 + O(g_0^6) \qquad (5.37a) \\ \widetilde{c}_t(g_0^2) &= 1 - 0.01795(2)\ g_0^2 + O(g_0^4)\ . \qquad (5.37b) \end{aligned}$$

We consider the boundary improved SF *tree-level improved Lüscher–Weisz gauge action* next:

$$S_G^{\mathrm{LW,SF}}[U] = S_G^{\mathrm{LW}}[U] + \delta S_{G,b}[U] = \frac{1}{g_0^2} \sum_{i=0,1} c_i \sum_{\mathcal{C}\in\mathcal{S}_i} w_i(\mathcal{C}) \mathrm{tr}[\mathbb{1} - U(\mathcal{C})]\ . \tag{5.38}$$

If the counterterms are again expressed solely in terms of plaquettes, (5.34), one finds

$$w_0(\mathcal{C}) = \begin{cases} \frac{1}{2}c_s(g_0^2) & \text{if } \mathcal{C} \in \text{boundary} \\ c_t(g_0^2) & \text{if exactly 1 link of } \mathcal{C} \in \text{boundary} \\ 1 & \text{otherwise}\ , \end{cases} \tag{5.36}$$

as for the Wilson gauge action, while the rectangles remain unchanged:

$$w_1(\mathcal{C}) = \begin{cases} 0 & \text{if } \mathcal{C} \in \text{boundary} \\ 1 & \text{if exactly 2 links of } \mathcal{C} \in \text{boundary} \\ 1 & \text{otherwise}\ . \end{cases} \tag{5.39}$$

Referring to [51], this amounts to *Choice A*[5]. Another option is *Choice B*, defined by the weight factors [51]

$$w_0(\mathcal{C}) = \begin{cases} \frac{1}{2}c_s(g_0^2) & \text{if } \mathcal{C} \in \text{boundary} \\ c_t(g_0^2) & \text{if exactly 1 link of } \mathcal{C} \in \text{boundary} \\ 1 & \text{otherwise} \end{cases} \tag{5.40}$$

---

[5] Note that we have defined $c_s(g_0^2)$ in a slightly different way compared to [51], for the purpose of similarity to the case of the Wilson gauge action. The difference is a factor $2c_0$, which is however irrelevant here, as will become clear from the discussion below.

and

$$w_1(\mathcal{C}) = \begin{cases} 1/2 & \text{if } \mathcal{C} \in \text{boundary} \\ 3/2 & \text{if exactly 2 links of } \mathcal{C} \in \text{boundary} \\ 1 & \text{otherwise} \end{cases} \tag{5.41}$$

Note that in comparison to its original definition [51], the weights for the spatial loops at the boundaries have been altered here, following the implementation of the SF in the openQCD software package [77, 98] (documented also e.g. in [22]). Since the boundary fields are fixed in the SF, this merely adds an irrelevant constant term to the action. For the same reason, and the use of Abelian boundary gauge fields, $c_s(g_0^2)$ is simply set to one. The only significant difference between Choice A and B is the weight of the rectangles with 2 spatial links at the boundary. Accordingly, the boundary improvement coefficients $c_t(g_0^2)$ that concern the corresponding plaquettes are different (already at tree-level). The advantage of Choice B, on which we will focus from now on, is that it allows to express the SF background field analytically, which is useful for perturbative computations. The tree-level value of $c_t(g_0^2)$ is 1 [51]. The corresponding 1-loop coefficient for Choice B has in principle first been worked out in [99] under use of the N_f-dependent part of the 1-loop coefficient [96] as given in (5.37a), but for a variation for which an additional improvement associated with the rectangles appears. The 1-loop formula which corresponds directly to Choice B as it is defined above reads[6]

$$c_t(g_0^2) = 1 + \left[-0.001782 + 0.0191410\, \frac{3N_f}{5} \right] \cdot g_0^2 + O(g_0^4) \; . \tag{5.42a}$$

The fermionic boundary improvement coefficient is not known beyond tree-level for the Lüscher–Weisz gauge action in connection with Choice B:

$$\tilde{c}_t(g_0^2) = 1 + O(g_0^2) \; . \tag{5.42b}$$

Finally, we remark that an important feature of Abelian background fields (cf. (5.7)) is that their use leads to particularly small cutoff effects once $O(a)$ improvement as described above is established [90].

---

[6] The difference in (5.42a) compared to [99] is merely a factor $c_0 = 5/3$ in the N_f-independent part, and a factor $(c_0)^{-1} = 3/5$ in the N_f-dependent part of the 1-loop coefficient.

5.3 Schrödinger functional correlation functions

Fermion fields at the boundaries of the SF are created by functional derivatives with respect to the fermion boundary fields (5.19):

$$\zeta(\vec{x}) = a^{-3}\frac{\delta}{\delta\overline{\rho}(\vec{x})}\,, \qquad \overline{\zeta}(\vec{x}) = -a^{-3}\frac{\delta}{\delta\rho(\vec{x})}\,, \tag{5.43a}$$

$$\zeta'(\vec{x}) = a^{-3}\frac{\delta}{\delta\overline{\rho}'(\vec{x})}\,, \qquad \overline{\zeta}'(\vec{x}) = -a^{-3}\frac{\delta}{\delta\rho'(\vec{x})}\,. \tag{5.43b}$$

As part of an operator $O \equiv O(\zeta,\zeta',\overline{\zeta},\overline{\zeta}')$, these act on the Boltzmann factor in the path integral expectation value

$$\langle O\rangle = \left[\frac{1}{Z}\int D[\psi,\overline{\psi},U]\; O(\zeta,\zeta',\overline{\zeta},\overline{\zeta}')\;\exp\left\{-S[\psi,\overline{\psi},U]\right\}\right]_{\rho=\rho'=\overline{\rho}=\overline{\rho}'=0}\,. \tag{5.44}$$

Of particular importance are

$$\mathcal{O}^{ji} = a^6\sum_{\vec{y},\vec{z}}\overline{\zeta}^{j}(\vec{y})\;\gamma_5\;\zeta^{i}(\vec{z}) \tag{5.45a}$$

and

$$\mathcal{O}'^{ji} = a^6\sum_{\vec{y},\vec{z}}\overline{\zeta}'^{j}(\vec{y})\;\gamma_5\;\zeta'^{i}(\vec{z})\,, \tag{5.45b}$$

which represent a pseudoscalar quark-antiquark state at the boundaries $x_0 = 0$ and $x_0 = T$, respectively. These states are of definite, zero momentum, due to the summation over the boundaries in (5.45) in connection with the gluon field boundary conditions (5.7) that preserve spatial translation invariance (cf. sec. 5.1).

Under application of $\mathcal{O}^{ji}$, $\mathcal{O}'^{ji}$, and currents which are suitable to create or annihilate pseudoscalar states in the bulk ($x_0 \neq 0, T$) of the SF, namely $\partial_\mu A_\mu^{ij}$ and P^{ij} (cf. (3.73, 3.74, 3.75)), particularly useful, non-vanishing transition amplitudes can be built. The *Schrödinger functional correlation functions (SFCFs)* read

$$f_A^{ij}(x_0) = -\frac{1}{2}\left\langle A_0^{ij}(x)\;\mathcal{O}^{ji}\right\rangle \tag{5.46}$$

$$f_P^{ij}(x_0) = -\frac{1}{2}\left\langle P^{ij}(x)\;\mathcal{O}^{ji}\right\rangle\,, \tag{5.47}$$

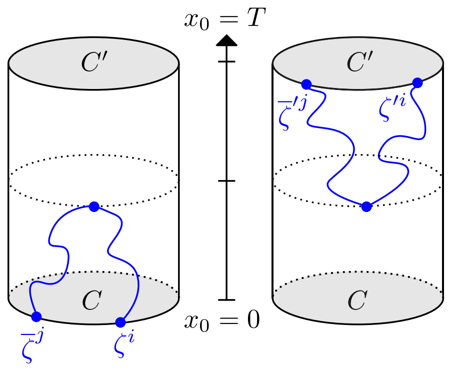


Figure 5.3: Illustration of the states involved in the boundary-to-bulk Schrödinger functional correlation functions and their propagation in time. *Left:* Forward correlation function with $\mathcal{O}^{ji}$. *Right:* Backward correlation function with $\mathcal{O}'^{ji}$.

and

$$f_A'^{ij}(T-x_0) = -\frac{1}{2}\left\langle \mathcal{O}'^{ji}\; A_0^{ij}(x)\right\rangle \;= +\frac{1}{2}\left\langle A_0^{ij}(x)\; \mathcal{O}'^{ji}\right\rangle \tag{5.48}$$

$$f_P'^{ij}(T-x_0) = -\frac{1}{2}\left\langle \mathcal{O}'^{ji}\; P^{ij}(x)\right\rangle \;= -\frac{1}{2}\left\langle P^{ij}(x)\; \mathcal{O}'^{ji}\right\rangle\;. \tag{5.49}$$

The unprimed (primed) correlation functions represent the transition amplitude for a state created at the boundary $x_0 = 0$ ($x_0 = T$), which propagates forward (backward) in time and is annihilated in the bulk at $0 < x_0 < T$. Correspondingly, they are also called *forward* and *backward boundary-to-bulk correlation functions*, respectively. While the forward correlation functions fall off with x_0, the backward counterparts fall off with $T - x_0$, which is the reason for the opposite parametrization in (5.48) and (5.49). The SFCFs are illustrated in fig. 5.3. Due to the aforementioned spatial translation invariance being respected by the SF boundary conditions as well as the boundary operators, they merely depend on x_0. In simulations, one may hence average over all the spatial sites in the evaluation of the correlation functions in order to improve statistics,

$$f_X^{ij}(x_0) \to (a/L)^3 \sum_{\vec{x}} f_X^{ij}(x_0, \vec{x})\;, \qquad X = A, P\;. \tag{5.50}$$

For completeness, we remark that the *boundary-to-boundary correlation function* $f_1^{ij}(x_0)$, also depicted in fig. 5.3, is of particular importance, too. It plays a role for instance in the determination of the renormalization constant Z_P (3.77) of the pseudoscalar current, see e.g. [100]. Moreover, apart from the pseudoscalar operators (5.45), vector operators with the replacement $\gamma_5 \to \gamma_k$, $k = 1, 2, 3$, can be considered.

We now return to the discussion of the bare PCAC mass, which was defined in sec. 3.6 as

$$m^{ij}_{\mathrm{PCAC}}(x) = \frac{\tilde{\partial}_\mu \left\langle A^{ij}_\mu(x)\; O \right\rangle}{2 \left\langle P^{ij}(x)\; O \right\rangle} \,, \tag{3.81}$$

with an arbitrary operator O. From the discussion above, it is clear that the operator $\mathcal{O}^{ji}$ (5.45a) is a particularly useful choice in the Schrödinger functional. It renders the correlation functions in (3.81) space-independent, and therefore also leads to vanishing spatial derivatives in the numerator. The PCAC mass in that case depends only on time, and may be expressed in terms of the axial (5.46) and pseudoscalar (5.47) SFCFs as

$$m^{ij}_{\mathrm{PCAC}}(x_0) = \frac{\tilde{\partial}_0 \left\langle A^{ij}_0(x)\; \mathcal{O}^{ji} \right\rangle}{2 \left\langle P^{ij}(x)\; \mathcal{O}^{ji} \right\rangle} = \frac{\tilde{\partial}_0 f^{ij}_A(x_0)}{2 f^{ij}_P(x_0)} \,. \tag{5.51}$$

Accordingly, the backward PCAC mass is defined in the SF by

$$m'^{ij}_{\mathrm{PCAC}}(x_0) = \frac{\tilde{\partial}_0 \left\langle \mathcal{O}'^{ji}\; A'^{ij}_0(x) \right\rangle}{2 \left\langle \mathcal{O}'^{ji}\; P'^{ij}(x) \right\rangle} = \frac{\tilde{\partial}_0 f'^{ij}_A(T - x_0)}{2 f'^{ij}_P(T - x_0)} \,. \tag{5.52}$$

One may proceed very similarly for the improved PCAC mass (3.108), introduced in sec. 3.7. As it is expressed in terms of the improved axial current $(A_I)^{ij}_0$, cf. (3.103), it is convenient to define the associated improved axial correlation function,

$$f^{ij}_{A,\mathrm{I}}(x_0) = f^{ij}_A(x_0) + a c_A \tilde{\partial}_0 f^{ij}_P(x_0) \,, \tag{5.53}$$

in terms of which the improved PCAC mass reads

$$m^{ij}_{\mathrm{PCAC,I}}(x_0) = \frac{\tilde{\partial}_0 \left\langle (A_I)^{ij}_0(x)\; \mathcal{O}^{ji} \right\rangle}{2 \left\langle P^{ij}(x)\; \mathcal{O}^{ji} \right\rangle} = \frac{\tilde{\partial}_0 f^{ij}_{A,\mathrm{I}}(x_0)}{2 f^{ij}_P(x_0)} = \frac{\tilde{\partial}_0 f^{ij}_A(x_0) + a c_A \tilde{\partial}^2_0 f^{ij}_P(x_0)}{2 f^{ij}_P(x_0)} \,. \tag{5.54}$$

Note that the renormalization pattern of the PCAC mass stays the same in the Schrödinger functional framework. In the improved case and $N_f = 4$, this means

$$m_{\mathrm{R}} = \frac{Z_A(\tilde{g}^2_0)}{Z_P(\tilde{g}^2_0, a\mu)} \, \frac{1 + a b_A(g^2_0)\; m_q}{1 + a b_P(g^2_0)\; m_q} \, m_{\mathrm{PCAC,I}}(x_0) + O(a^2) \,, \tag{5.55}$$

the only difference to (3.107) being the very definition of $m_{\mathrm{PCAC,I}}$.

6 The gradient flow coupling and the topological charge

6.1 The gradient flow coupling

The *gradient flow* (GF) considers the gauge fields in an additional 5th dimension, in which the fundamental fields $A_\mu(x)$ (continuum) or $U_\mu(x)$ (lattice) are evolved in a specific way along the so-called *flow time* t. It has gained in importance in recent years, based on the insight that the gauge fields at positive flow time $t > 0$ do not require renormalization [101–103]. This allows to probe the theory in many interesting ways. Among many applications (see e.g. [104] for a review), it in particular leads to a definition of the renormalized *gradient flow coupling* in terms of the Schrödinger functional [101, 105], which we will address in detail below. The coupling, for instance, together with step scaling [106], became part of the ALPHA Collaboration strategy to determine the Λ parameter in $N_f = 3$ [43, 44]. Important advances were also made in the field of scale setting, where the new gradient flow scales t_0 [101] and w_0 [107] play an important role (see [39] for an overview).

We first discuss general properties of the gradient flow and its coupling in the continuum (sec. 6.1.1), before we move on to the formulation on the lattice (sec. 6.1.2).

6.1.1 The Yang–Mills gradient flow

The *Yang–Mills gradient flow* $B_\mu(x,t)$ of the fundamental $SU(3)$ gauge field A_μ is defined by the non-linear flow equation [101]

$$\partial_t B_\mu(x,t) = D_\nu G_{\nu\mu}(x,t) \ , \qquad B_\mu(x,0) = A_\mu(x) \ , \tag{6.1}$$

where $t \geq 0$ denotes the flow time, and with[1]

$$D_\mu = \partial_\mu + \left[B_\mu, \,\cdot\,\right] \quad \text{and} \quad G_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu + [B_\mu, B_\nu] \,. \tag{6.2}$$

As the right hand side of (6.1) obeys

$$D_\nu G_{\nu\mu} \;\sim\; -\frac{\delta S_G[B]}{\delta B_\mu} \,, \tag{6.3}$$

where S_G is the pure (Yang–Mills) gauge action that corresponds to (2.15), the flow equation evolves the gauge field towards a local minimum of the action. As has been argued in [101], this may be understood by the fact that the flow effectively constitutes a smoothing[2] of the gauge field over a range $\sqrt{8t}$. One remarkable feature of this procedure is the fact that the gauge field B_μ at $t > 0$ is renormalized, rendering expectation values of local, gauge-invariant quantities finite. While this was later proven for the general case of correlation functions of smoothed gauge fields in [102], it was first shown for the particular case of the energy density (cf. (3.12, A.16))

$$\langle E(x,t) \rangle = -\frac{1}{2} \left\langle \mathrm{tr} \left[G_{\mu\nu}(x,t) G_{\mu\nu}(x,t) \right] \right\rangle = \frac{1}{4} \left\langle G^a_{\mu\nu}(x,t) G^a_{\mu\nu}(x,t) \right\rangle \tag{6.4}$$

in [101]. Specifically, in terms of the running coupling $\alpha(\mu) = \bar{g}^2_{\overline{\mathrm{MS}}}(\mu)/(4\pi)$ in the $\overline{\mathrm{MS}}$ scheme at a renormalization scale $\mu = 1/\sqrt{8t}$, the perturbative expansion

$$t^2 \langle E(x,t) \rangle = \frac{3(N^2-1)}{32\pi} \, \alpha(\mu) \left[1 + c_1' \alpha(\mu) + O(\alpha^2) \right] \tag{6.5}$$

with a finite coefficient c_1' has been derived in *infinite volume*.

The energy density at positive flow time may be used to define a non-perturbative *gradient flow coupling* in *finite volume*. In order to be able to study the running of the GF coupling, just as for the SF coupling, finite size scaling (cf. (5.18)) is applied, i.e. the given renormalization scale runs with the size L of the box,

$$\mu = \frac{1}{\sqrt{8t}} = \frac{1}{cL} \,, \tag{6.6}$$

---

[1] Apparent differences with respect to the previously considered covariant derivative and field strength tensor, (3.10), are due to the use of *anti-hermitean generators* $\tilde{T}_a$, see app. A.1, which is standard in the present context [101].

[2] The precise meaning of this statement is given in [101]. Somewhat simplifying, the gauge field $B_\mu(x,t)$ emerges from a term $\sim e^{-(x-y)^2/(2r^2/4)} \, A_\mu(y)$ with $r = \sqrt{8t}$, integrated over y.

where the *smoothing fraction* c represents the fraction of the box over which the gauge field is averaged. This approach was implemented in a periodic box [108] and in the Schrödinger functional [105]. In the latter case, the GF coupling is defined as

$$\overline{g}^2_{\mathrm{GF}}(L) = \mathcal{N}^{-1} \cdot t^2 \left\langle E(x_0, t) \right\rangle \Big|^{x_0=T/2}_{t=c^2L^2/8} . \tag{6.7}$$

Note that due to the broken translational invariance in time direction, the energy density becomes explicitly dependent on x_0. The normalization factor $\mathcal{N}^{-1}$ in (6.7) ensures $\overline{g}^2_{\mathrm{GF}} = g_0^2 + O(g_0^4)$, and may be obtained from the leading order contribution of the perturbative expansion of $\langle E(x_0, t)\rangle$ in the SF, similar to (6.5). Apart from the size L of the SF box, $\mathcal{N}^{-1}$ implicitly depends on the smoothing fraction c, the ratio $\rho = T/L$, the boundary fields ϕ, as well as—since the SF breaks translational invariance in the time direction—the time slice x_0:

$$\mathcal{N} \equiv \mathcal{N}(c, \rho, x_0, \phi) . \tag{6.8}$$

In addition, there is a dependence[3] of the GF coupling on the fermionic phase angle θ, which has been observed to be weak [105], however. Note that as indicated in (6.7), the time slice x_0 is always fixed to the center of the lattice. Hence, in summary, to complete the definition of the GF coupling, it suffices to specify the SF parameters as well as c. We will return to this in the upcoming discussion of the gradient flow coupling on the lattice.

Finally, we mention that it is also possible to define the GF coupling in terms of the spatial components $\left\langle G^a_{ik}(t) G^a_{ik}(t)\right\rangle$ of the energy density only, see e.g. [105].

6.1.2 The gradient flow coupling on the lattice

On the lattice, the flow $V_\mu(x,t)$ associated with the lattice gauge links $U_\mu(x)$ is defined by

$$a^2\, \partial_t V_\mu(x,t) = -g_0^2 \left\{ \tilde{T}^a \partial^a_{x,\mu} S_{\mathrm{flow}}[V] \right\} V_\mu(x,t) , \qquad V_\mu(x,0) = U_\mu(x) , \tag{6.9}$$

which replaces (6.1). The derivative $\partial^a_{x,\mu}$ takes values in the Lie-algebra $\mathfrak{su}(3)$ and is defined to act on an arbitrary function of the link variable $U_\mu(x)$ as

$$\partial^a_{x,\mu}\, f(U_\mu(x)) = \frac{d}{ds}\, f(e^{s\tilde{T}^a} U_\mu(x))|_{s=0} . \tag{6.10}$$

---

[3] Note that in perturbation theory, the dependence arises beyond tree-level order.

For a complete definition of the gradient flow coupling on the lattice, three discretizations need to be chosen:

D1. The gauge action S_G
The gauge action S_G that is used (possibly in connection with the Wilson fermion action S_F^{W} (3.44)) to produce the fundamental gauge field configurations $U_\mu(x)$.

D2. The flow action S_{flow}
The flow action S_{flow} (cf. (6.9)) may be any discretization of the gauge action S_G, independent of the choice in D1. For $S_{\text{flow}} = S_G^{\mathrm{W}}$ (3.36) and $S_{\text{flow}} = S_G^{\mathrm{LW}}$ (3.112), the flow V_μ is referred to as *Wilson flow* or *Lüscher–Weisz flow*, respectively.

D3. The energy density E
Two valid choices for the discretization of the energy density (6.4), that both involve the replacement $U_\mu(x) \to V_\mu(x,t)$, are[4]

$$E^{(\text{plaq})}(x_0,t) = 2\sum_{\mu<\nu} \mathrm{Re}\,\mathrm{tr}[\mathbb{1} - V_{\mu\nu}(x_0,t)] \tag{6.11}$$

$$\text{and} \qquad E^{(\text{clov})}(x_0,t) = -\frac{1}{2}\mathrm{tr}\left[G_{\mu\nu}^{(\text{clov})}(x_0,t)G_{\mu\nu}^{(\text{clov})}(x_0,t)\right] . \tag{6.12}$$

The first option is based on (3.36) with the obvious generalization of the plaquette, $U_{\mu\nu}(x) \to V_{\mu\nu}(x,t)$, while the second option is the (symmetric) clover definition with the straight forward generalization of the clover term (3.102), $F_{\mu\nu}^{(\text{clov})}(x) \to G_{\mu\nu}^{(\text{clov})}(x,t)$.

The three choices determine the intrinsic cutoff effects of the gradient flow coupling, together with the parameters already present in the continuum formulation. Accordingly, the normalization $\mathcal{N}$ (cf. (6.8)) on the lattice depends on those, too. It has been worked out in [105] for the case of

$$S_G = S_G^{\mathrm{W}}, \quad S_{\text{flow}} = S_G^{\mathrm{W}} \quad \text{and} \quad E = E^{(\text{clov})} \tag{6.13}$$

in conjunction with vanishing boundary fields, $\phi = \phi' = 0$ (5.7). In the same work, the gradient flow coupling has been investigated *non-perturbatively* for the above three choices, (6.13), on $N_f = 2$ ensembles at $L \sim 0.4$ fm. Smoothing fractions in the range $c \in [0.3, 0.5]$ have turned out to be convenient, as their use leads to *high statistical precision* at affordable cost (especially for $c \lesssim 0.3$) and *modest cutoff effects* (especially

---

[4] The expressions are understood to be averaged over the spatial components.

for $c \gtrsim 0.5$). These properties make the GF coupling particularly interesting as an alternative to the traditional SF coupling.

Furthermore, the gradient flow coupling was shown to exhibit a rather weak *mass dependence* [105]. In accordance with the fact that the quark masses enter only as sea quarks, the gradient flow coupling depends to a fairly good approximation only on the trace of the mass matrix. While this is suggested by data displayed for instance in [39], it has also been shown to be true for flow observables up to NLO in chiral perturbation theory [109]. Finally, we mention that the Symanzik improvement program to obtain $O(a^2)$ cutoff effects may be applied for flow observables. With regard to S_{flow}, the *Zeuthen flow* [110, 111] has been developed to serve this purpose.

6.2 The topological charge

6.2.1 The topological charge in the continuum

The winding number for a simple model

For simplicity, we consider a $U(1)$ gauge field A in space $\mathbb{R}$,

$$A : \mathbb{R} \to \mathfrak{u}(1) \cong \mathbb{R} \tag{6.14}$$

and restrict ourselves to pure gauge configurations first, i.e. configurations that are related to the vacuum by gauge transformations (cf. (2.13)),

$$A(x) = \frac{i}{e} \left(\partial_x \Omega(x)\right) \Omega^\dagger(x) = \frac{-1}{e} \alpha'(x) \qquad \text{with} \qquad \Omega(x) = e^{i\alpha(x)} \in U(1) \, . \tag{6.15}$$

Without loss of generality, we impose $\lim_{x\to\pm\infty} \Omega(x) = 1$, which by (6.15) is equivalent to

$$\lim_{x\to\pm\infty} \alpha(x) = 2\pi n_\pm \quad \text{with } n_\pm \in \mathbb{Z} \, . \tag{6.16}$$

One then finds

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dx \; A(x) \sim \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \; \alpha'(x) = n_+ - n_- \equiv n \in \mathbb{Z} \, . \tag{6.17a}$$

For the graphic interpretation of this result, it is convenient to use a *stereographic projection* of $\mathbb{R}$ to the unit circle $\mathbb{S}^1$, which may be parametrized by the angle ϕ, see fig. 6.1. The gauge field $A(x)$ is then fully specified by $\alpha(\phi)$, and (6.17a) can be written as

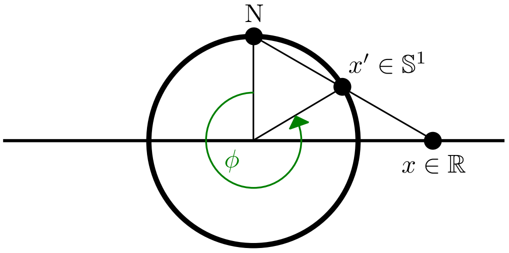


Figure 6.1: Stereographic projection of $\mathbb{R} \to \mathbb{S}^1$, $x \mapsto x'$. The image of this mapping equals $\mathbb{S}^1$ without the north pole (N). It may be compactified by identification of N with $x \to \pm\infty$ and parametrized by the angle $\phi \in [0, 2\pi]$.

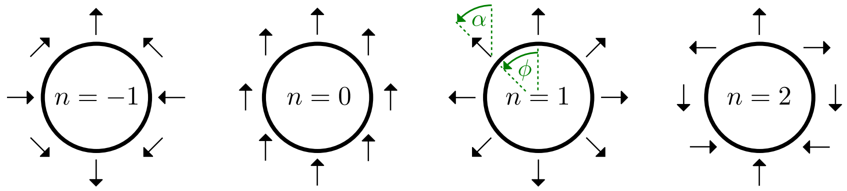


Figure 6.2: Illustration of homotopy classes of different winding number n. As indicated in green, the position of the displayed unit vectors is given by ϕ, whereas their orientation is determined by α. The above representatives obey $\alpha = n \cdot \phi$.

$$\frac{1}{2\pi} \int_0^{2\pi} d\phi \, \alpha'(\phi) = n_+ - n_- \equiv n \in \mathbb{Z} \, . \tag{6.17b}$$

In this formulation, it is obvious that n counts the number of rotations the unit vector field that represents α receives as one moves around the circle (from $\phi = 0$ to $\phi = 2\pi$). For this reason, it is called *winding number*. It is a topological invariant that divides the space of all possible vacuum gauge field configurations into topological equivalence classes, so-called homotopy classes. Elements of the same class may be continuously transformed into each other, while this is not true for elements of two distinct classes. The simplest representatives of the homotopy classes with winding number n are those with $\alpha = n \cdot \phi$ as shown in fig. 6.2.

The generalization to non-vacuum gauge fields A is straight forward, as these satisfy (6.15) in the limit $x \to \pm\infty$, which is a corollary from the finiteness of the (Euclidean) action. Consequently, they may be classified by the winding number n as well. However, away from the boundaries $x \to \pm\infty$ of the compactified $\mathbb{R}$, or equivalently away from the north pole N of $\mathbb{S}^1$, they cannot be specified merely in terms of α, in contrast to the pure gauge fields. The general space of gauge field configurations is in fact connected, but two pure gauge configurations of different n can not be continuously transformed into each other without leaving pure gauge (with zero action).

The topological charge in continuum QCD

The concept of the winding number may be generalized to the case of QCD[5]. Instead of (6.14), the gauge fields correspond to a mapping

$$A_\mu : \mathbb{R}^4 \to \mathfrak{su}(3) \; . \tag{6.18}$$

The equivalent to the winding number n (6.17a) is the *topological charge*

$$Q_{\text{top}} = \int d^4x \; q(x) = \frac{1}{16\pi^2} \int d^4x \; \text{tr} \left[F_{\mu\nu} \widetilde{F}_{\mu\nu} \right] \; , \tag{6.19a}$$

where

$$\widetilde{F}_{\mu\nu} = \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} F_{\rho\sigma} \tag{6.19b}$$

denotes the *dual field strength tensor* and q is the *topological charge density*. In analogy to (6.17a), the topological charge can be expressed as the spacetime integral over a total derivative, namely of the *Chern–Simons current* [112]

$$K_\mu(A) = \frac{1}{16\pi^2} \epsilon_{\mu\nu\rho\sigma} \left(A^a_\nu \partial_\rho A^a_\sigma + \frac{1}{3} f^{abc} A^a_\nu A^b_\rho A^c_\sigma \right) \; , \tag{6.20}$$

which subsequently may be reformulated under use of Gauss' law as a surface integral over the boundary of $\mathbb{R}^4$, or equivalently the infinitely large three-sphere S^3_∞,

$$Q_{\text{top}} = \int d^4x \; \partial_\mu K_\mu(A) = \int_{S^3_\infty} d^3\sigma \; K_\perp(A) \in \mathbb{Z} \; . \tag{6.21}$$

Instantons and the instanton bound

Instantons are non-vacuum gauge field configurations of topological charge $Q_{\text{top}} = \pm 1$ that minimize the (Euclidean) action. As such, they mediate tunneling processes through the finite action barrier between the topologically distinct pure gauge sectors described above. To obtain them, one may reformulate the gauge action (3.12) as (see e.g. [113])

$$S_G = \frac{1}{2g_0^2} \int d^4x \left\{ \text{tr} \left[\left(F_{\mu\nu}(x) \mp \widetilde{F}_{\mu\nu}(x) \right)^2 \right] \pm \text{tr} \left[F_{\mu\nu}(x) \widetilde{F}_{\mu\nu}(x) \right] \right\} \; . \tag{6.22}$$

---

[5] Note that—in contrast to the continuum QCD discussion in sec. 2.1—we work in Euclidean space here to facilitate the transition to the lattice.

The non-negative first term vanishes for an *(anti-)dual* field strength tensor

$$F^I_{\mu\nu} = \pm\widetilde{F}^I_{\mu\nu} \,. \tag{6.23}$$

The self-dual solutions [114, 115] satisfy the classical Yang–Mills equation of motion, $D_\mu F^I_{\mu\nu} = 0$, can be shown to be strongly located in spacetime[6] (see e.g. [116]) and are identified with pseudo-particles, the *(anti-)instantons*. Their action is given by (see e.g. [117])

$$S^I_G = \frac{8\pi}{g_0^2} \,. \tag{6.24}$$

On the other hand, the action (6.22) under use of (6.19a) and (6.23) becomes

$$S^I_G = \pm\frac{8\pi^2}{g_0^2} Q^I_{\mathrm{top}} = \frac{8\pi^2}{g_0^2} |Q^I_{\mathrm{top}}| \,, \tag{6.25}$$

where in the last step the topological charge

$$Q^I_{\mathrm{top}} = \pm 1 \tag{6.26}$$

of the (anti-)instanton, which follows from comparison with (6.24), has been used. A general gauge configuration may contain several (anti-)instantons. Note, however, that these can compensate each other with regard to the (global) topological charge. An equal number of instantons and anti-instantons, for instance, yields $Q_{\mathrm{top}} = 0$. This leads to the inequality

$$S_G \geq \frac{8\pi^2}{g_0^2}\left[\#(\text{instantons}) + \#(\text{anti-instantons})\right] \geq \frac{8\pi^2}{g_0^2} |Q_{\mathrm{top}}| \,, \tag{6.27}$$

which is called *instanton bound* for the action S_G of a gauge configuration with topological charge Q_{top}.

Distribution of the topological charge and the chiral anomaly

The QCD path integral is symmetric with respect to the topological charge, which expresses itself in a vanishing expectation value,

$$\langle Q_{\mathrm{top}} \rangle = 0 \,. \tag{6.28}$$

---

[6] Accordingly, they represent an "instantaneous" event in spacetime.

The width of its distribution increases with the physical volume, which is linked to an increasing opportunity for (anti-)instantons to appear. This circumstance is considered in the *topological susceptibility*

$$\chi_{\text{top}} = \frac{\left\langle Q_{\text{top}}^2 \right\rangle}{V} = \frac{1}{V} \int d^4x \int d^4y \, \langle q(x) q(y) \rangle \ , \tag{6.29}$$

which in general depends on the number N_f of quark flavors, their masses $\mathcal{M}$ (cf. (3.13)), and the considered volume[7] V.

The topological charge plays an important role in the context of the chiral anomaly (cf. sec. 2.2). This is most prominently expressed by the *(Atiyah–Singer) index theorem* [119], which states that, in analogy to (6.17), the topological charge is given by the *index* of the Dirac operator, defined as the difference of the number of its left- and right-handed zero modes:

$$\text{index}(D) = n_L - n_R = Q_{\text{top}} \ . \tag{6.30}$$

The possibility of the index to be different from zero—and hence in the end the topological structure of the QCD vacuum—can be shown to lead to the chiral anomaly (see e.g. [120]). Another demonstration of the intertwining of topological charge and the chiral anomaly is the *Witten–Veneziano formula* [121, 122], that attributes the exceptionally large η' mass to the non-vanishing of the topological susceptibility, $\chi_{\text{top}} \neq 0$.

6.2.2 The topological charge on the lattice

Discretization in terms of the flow

The naive discretization of the topological charge Q_{top} and susceptibility χ_{top} merely consists of the choice of a discretization of the field strength tensor $F_{\mu\nu}$. This, however, leads to non-integrable short-distance singularities, which prevent a well-defined continuum limit for χ_{top}, see e.g. [123].

While different approaches to solve this issue have been developed, see e.g. [123, 124], the gradient flow provides a particularly elegant remedy by defining the topo-

---

[7] Note that the infinite volume limit is often considered (finite size effects for affordable physical volumes are usually found to be small compared to statistical errors). Quenched simulations lead to $\lim_{V\to\infty} \chi_{\text{top}} \approx (190 \text{ MeV})^4$, see e.g. [118].

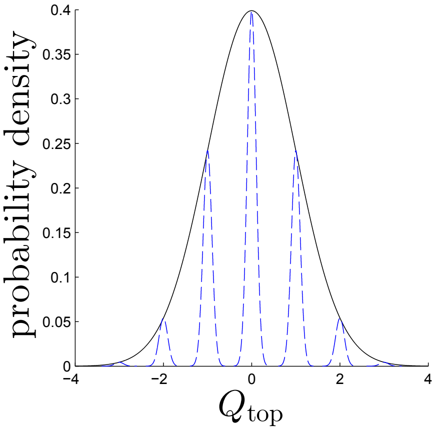


Figure 6.3: Qualitative sketch of the expected distribution of the topological charge. The physical volume V determines the width of the enfolding normal distribution (solid black), while the width of the distributions[9] that belong to a certain sector (in the sense of (6.32)) is set by the lattice spacing a (dashed blue).

logical charge[8] in terms of fields at positive flow time [101]:

$$Q_{\text{top}}(t) = -\frac{a^4}{16\pi^2}\sum_x \operatorname{tr}\left[G^{(\text{clov})}_{\mu\nu}(x,t)\widetilde{G}^{(\text{clov})}_{\mu\nu}(x,t)\right] , \tag{6.31}$$

where $G^{(\text{clov})}_{\mu\nu}$ is the smoothed counterpart of $F^{(\text{clov})}_{\mu\nu}$ known from (6.12). This definition has led to a smooth continuum extrapolation of χ_{top} and results similar to other approaches [125]. Furthermore, the topological charge is expected to approximate integer values as the continuum is approached. Together with the aforementioned increase of $\left\langle Q^2_{\text{top}}\right\rangle$ with the physical volume V, this leads to the expected distribution of Q_{top} on the lattice illustrated in fig. 6.3.

The flow also provides an understanding of how the disjoint topological sectors (cf. sec. 6.2.1) emerge in the continuum limit [101], while on the lattice the gauge field space is connected and the topological charge is to some extent ambiguous. First, it is known that fields which satisfy

$$h < 0.067 , \tag{6.32a}$$

where

$$h = \max_{x\mu\nu} s_{\mu\nu}(x) \qquad \text{with} \qquad s_{\mu\nu}(x) = \operatorname{Re}\operatorname{tr}\left[1 - U_{\mu\nu}(x)\right] \tag{6.32b}$$

is a measure for the smoothness of the field U, belong to disjoint topological sectors

---

[8] Note that we again adopt the convention of anti-hermitean generators in the framework of the gradient flow (cf. footnote 1), which explains the relative minus sign between (6.19a) and (6.31).

[9] Note that these distributions are affected by cutoff effects and neither need to be normal nor centered around an integer. The situation displayed in the figure is idealized.

also on the lattice [126, 127]. Moreover, the gradient flow defines a one-to-one mapping between $U_\mu(x)$ and $V_\mu(x,t)$ [128], which allows to transfer the topological sector classification of U to V and vice versa. Finally, as has been shown for $N_f = 0$ [101] and $N_f = 2$ [129], the probability $P_{h>0.067}$ for a gauge configuration V in fixed volume to violate the bound (6.32) scales as

$$P_{h>0.067} \sim a^6 \ . \tag{6.33}$$

Hence, altogether, the limit $a \to 0$ entails the division of the topological sectors.

The topological charge and the HMC

The observed suppression of the relative weight of configurations between the sectors, (6.33), is a consequence of the increasing action barriers between the topological sectors as the lattice spacing is reduced. This also has severe implications in regard to the use of the Hybrid Monte Carlo algorithm. Transitions from one sector to another need to pass through a configuration with large action, which corresponds to a large error ΔH in the numerical integration (cf. (4.53)) that is unlikely to be accepted. In that case, tunneling between different sectors is strongly suppressed, and Monte Carlo simulations may get stuck in a topological sector, which is known as *topology freezing*. Furthermore, the autocorrelation of the topological charge tends to become very large. This phenomenon, which is referred to as *critical slowing down*, is in principle expected for any observable as the continuum limit is approached. However, the topological charge is notorious for the severity by which it is affected, as it couples strongly to slow modes of the Dirac operator. This is quantitatively captured by the *dynamical critical exponent* z, defined by[10]

$$\tau_{\text{int}} \sim a^{-z} \ , \tag{6.34}$$

which was found to be[11] $z = 5$ for the topological charge and $z \approx 1$ for Wilson loops [89], and is expected to be $z = 2$ for flow observables [130].

The effects of topology freezing and critical slowing down may to some extent be compensated by very large ensembles. However, at some point in the limit $a \to 0$, they inevitably lead to a violation of ergodicity (cf. (4.24)), and an insufficiency of the HMC to correctly sample the different topological sectors. In that case, the distribution of the

---

[10] Considering a factor a^{-4} for the increase of lattice sites, and a factor a^{-1} for the decrease in Monte Carlo step size, the cost of the simulation at constant volume scales as a^{-z-5}.

[11] Note that z depends not only on the observable, but also on the theory and the update algorithm.

topological charge observed in Monte Carlo simulations does not reflect the weights of the different topological sectors in the path integral. This, however, may lead to biased Monte Carlo estimates, in particular for observables which correlate with the topological charge. We will see this explicitly for the gradient flow coupling in sec. 7. In addition, the topological charge, in the usual capacity of the evaluated observable with the largest autocorrelation, is suited to give an estimate for the global exponential autocorrelation time $\tau_{\mathrm{exp}} = \max_O \tau_{O,\mathrm{exp}}$ (cf. (4.111)). This serves as a check for the sufficiency of the thermalization N_{th}, as well as the size N_{cf} of an ensemble needed to correctly estimate autocorrelations, and hence errors for any observable. In fact, τ_{exp} may also be directly applied to improve the latter [89].

Finally, we remark that there is a way to reduce the problem of critical slowing down, by use of open boundary conditions [98] to compute e.g. the hadron spectrum in large volume simulations. However, since the present work focuses on Schrödinger functional boundary conditions, we do not go into details here.

7 Critical slowing down and the gradient flow coupling in the Schrödinger functional

In [105], the GF coupling has been investigated on the lattice for $N_f = 2$ ensembles at a physical volume with $L \sim 0.4$ fm. The path integral is in that case largely dominated by the trivial topological sector ($Q_{\rm top} = 0$) and contributions from other sectors can be considered negligible. In contrast, according to the discussion in sec. 6.2, sectors of non-vanishing topological charge are expected to contribute in larger volumes, which brings up the question about the well-known problems of topology freezing and critical slowing down. The aim of this section is to investigate whether and how the determination of the gradient flow coupling is affected by these phenomena. To this end, at fixed volume with $L \sim 0.8$ fm, we investigate both the gradient flow coupling and the topological charge on the lattice, analyze the dependence of the former on the latter, and extrapolate to the continuum to obtain $\bar{g}^2_{\rm GF}(L)$. Most of the results appeared in [131], which the presentation here follows quite closely.

7.1 Setup of the numerical simulations

In order to be able to produce large statistics, we perform simulations in pure $SU(3)$ Yang–Mills theory (i.e. $N_f = 0$) with the Wilson gauge action S_G (3.38). We choose SF boundary conditions with vanishing boundary fields, $\phi = \phi' = 0$ (5.7), and fix the physical volume in terms of the Sommer scale r_0 (cf. sec. 3.4):

$$L = r_0/0.563 \sim 0.8 \text{ fm} . \tag{7.1}$$

This defines the line of constant physics at which we simulate with the lattice sizes $L/a \in \{8, 12, 16, 20, 24\}$. The corresponding bare couplings ($\beta = 6/g_0^2$) are determined under use of [32]

$$\log\left(a/r_0\right) = -1.6804 - 1.7331(\beta-6) + 0.7849(\beta-6)^2 - 0.4428(\beta-6)^3\,, \tag{7.2}$$

and can be found along with further parameters of the simulation in tab. 7.1. We use the HMC of the `openQCD` package[1] [77, 98] and the contained implementation of Schrödinger functional boundary conditions, as well as the 2-loop value of c_t (cf. (5.37a)) for $O(a)$ improvement. Each produced configuration is evolved by adaptive integration (see e.g. [105]) of the flow equation (6.9) with $S_{\text{flow}} = S_G$, including flow times t that correspond to $c \in \{0.3, 0.5\}$. Afterwards, on the smoothed configurations, the topological charge

$$Q_{\text{top}}(t) = -\frac{a^4}{16\pi^2}\sum_x \text{tr}\left[G^{(\text{clov})}_{\mu\nu}(x,t)\widetilde{G}^{(\text{clov})}_{\mu\nu}(x,t)\right] \tag{6.31}$$

and the GF coupling

$$\bar{g}^2_{\text{GF}}(L) = \mathcal{N}^{-1}\cdot t^2\left\langle E(x_0,t)\right\rangle\Big|^{x_0=T/2}_{t=c^2L^2/8} \tag{6.7}$$

$$\text{with}\qquad E(x_0,t) \equiv E^{(\text{clov})}(x_0,t) = -\frac{1}{2}\text{tr}\left[G^{(\text{clov})}_{\mu\nu}(x_0,t)G^{(\text{clov})}_{\mu\nu}(x_0,t)\right] \tag{6.12}$$

are measured using the clover discretization (3.102) for the field strength $G_{\mu\nu}$. Statistical errors are computed throughout under use of the gamma method (cf. sec. 4.10) with help of the `UWerr` package [88].

7.2 Histories and (auto)correlations

7.2.1 The topological charge

Histories of the topological charge are shown in fig. 7.1 for $c = 0.3$. As expected from the discussion in sec. 6.2, the topological charge tends to assume values that are closer to integers for smaller lattice spacings. For lattices up to $L/a = 16$, one observes that non-trivial configurations appear to cluster more and more as the lattice gets finer. This goes together with an increasing integrated autocorrelation time that corresponds to severe critical slowing down with a dynamical critical exponent of $z > 2$ (cf. (6.34)), displayed in fig. 7.2. Moreover, for the largest lattices $L/a \in \{20, 24\}$, configurations from non-trivial sectors appear less often, see tab. 7.2. However, the autocorrelations

---

[1] Version 1.2.

L/a	β	$N_{\rm cf}$	$\tau_{\rm meas}$
8	5.9032	80000	6
12	6.1410	80000	6
16	6.3413	40000	6
20	6.5119	15000	12
24	6.6552	7000	12

Table 7.1: Parameters of the numerical simulations. $N_{\rm cf}$ is the amount of configurations used for the measurements, and $\tau_{\rm meas}$ denotes the Monte Carlo time between two consecutive of those configurations in MDU.

L/a	$c = 0.3$	$c = 0.5$
8	98.66(8)	98.94(8)
12	98.19(20)	98.45(19)
16	98.46(62)	98.56(61)
20	*99.91(3)	*99.96(2)
24	*99.52(36)	*99.54(36)

Table 7.2: Fraction $\delta^{\epsilon}_{Q_{\rm top},0}$ of configurations with topological charge $Q_{\rm top} \leq \epsilon = 0.5$ (cf. (7.5)), given in percent. The values denoted with an asterisk are biased.

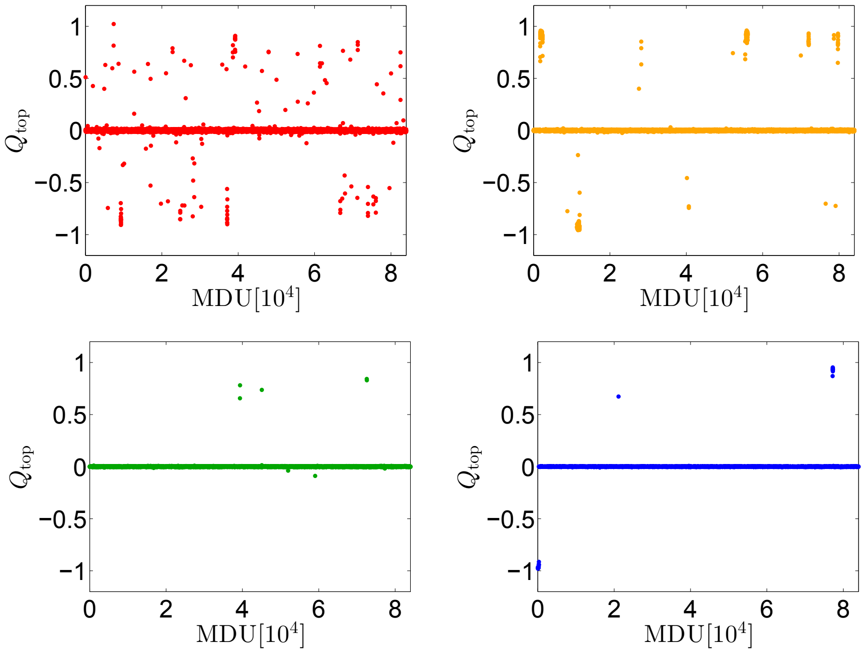


Figure 7.1: Histories (in excerpts) of $Q_{\rm top}$, for $c = 0.3$. The plots in red, orange, green and blue correspond to $L/a = 12, 16, 20, 24$, respectively.

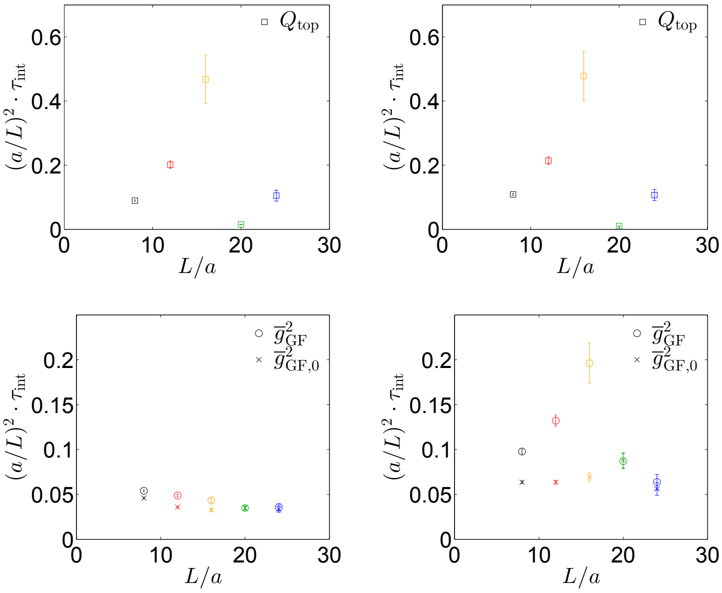


Figure 7.2: Integrated autocorrelation time of the topological charge Q_{top} *(top)* and the (modified) gradient flow coupling *(bottom)*, multiplied by $(a/L)^2$ and in units of 2 MDU. The *left (right)* panel corresponds to $c = 0.3$ $(c = 0.5)$.

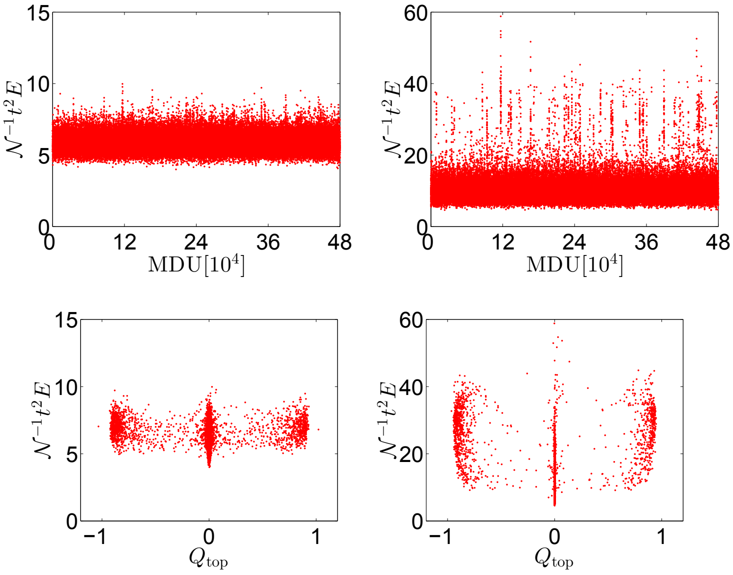


Figure 7.3: *Top:* Histories of $\overline{g}^2_{\mathrm{GF}}$. *Bottom:* Correlation of $\overline{g}^2_{\mathrm{GF}}$ and Q_{top}. All plots are shown exemplarily for $L/a = 12$. The *left (right)* panel corresponds to $c = 0.3$ $(c = 0.5)$.

of Q_{top} are obviously largely underestimated (cf. fig. 7.2), from which we infer that the ensembles are not large enough to sample the topological sectors correctly.

7.2.2 The gradient flow coupling

The histories of the quantity $\mathcal{N}^{-1}t^2E$, whose expectation value is the gradient flow coupling $\bar{g}^2_{\text{GF}}$, exhibit a certain amount of exceptionally large values, see fig. 7.3. This phenomenon is more pronounced the larger c is chosen, and goes along with an increasing correlation between the gradient flow coupling and the (absolute value of the) topological charge, also displayed in fig. 7.3, and expressed in the correlation coefficient (cf. (H.14))

$$\operatorname{cor}(\bar{g}^2_{\text{GF}}, |Q_{\text{top}}|) = \begin{cases} 0.26 & \text{for } c = 0.3 \\ 0.62 & \text{for } c = 0.5\,. \end{cases} \tag{7.3}$$

In particular, the large values of $\mathcal{N}^{-1}t^2E$ stem to a high amount from configurations of non-vanishing topological charge. In turn, the correct sampling of the topological sectors becomes a necessity in order to obtain correct results. However, as we have seen in sec. 7.2.1, this requirement is not fulfilled for the two largest lattices, which means that the results for $\bar{g}^2_{\text{GF}}$ on these lattices are biased.

7.2.3 The modified gradient flow coupling

In order to assess the impact of the non-trivial topological sectors and their insufficient sampling on the determination of $\bar{g}^2_{\text{GF}}$, we consider a modified GF coupling, which has the same perturbative expansion but takes into account only gauge configurations from the trivial sector:

$$\bar{g}^2_{\text{GF},0} = \mathcal{N}^{-1}\, t^2\, \left.\frac{\langle E(t)\ \delta_{Q_{\text{top}},0}\rangle}{\langle \delta_{Q_{\text{top}},0}\rangle}\right|_{t=c^2L^2/8} . \tag{7.4}$$

On the lattice, where we have non-integer values of Q_{top}, all configurations with $|Q_{\text{top}}| \leq \epsilon$ and $\epsilon \in \{0.1, 0.2, \ldots, 0.5\}$ are considered to belong to the trivial sector, i.e. we replace

$$\delta_{Q_{\text{top}},0} \to \delta^{\epsilon}_{Q_{\text{top}},0} = \Theta(Q_{\text{top}} + \epsilon)\Theta(\epsilon - Q_{\text{top}}) \tag{7.5}$$

in (7.4). The results for $L/a = 12$ can be seen in fig. 7.4. The contributions from non-trivial sectors do make a difference, and the effect is stronger for large c due

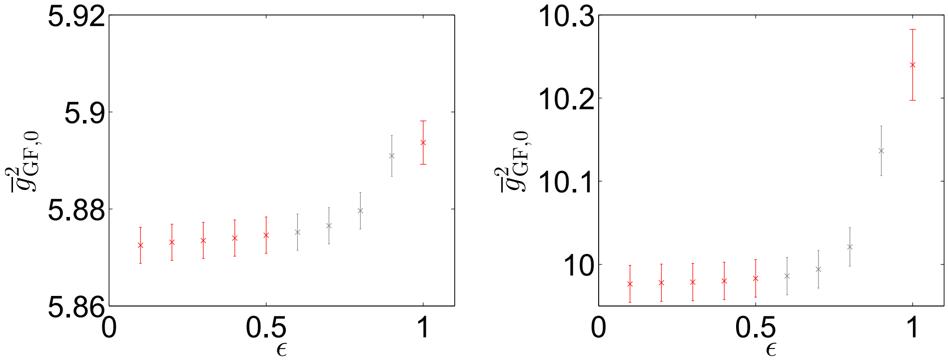


Figure 7.4: Modified gradient flow coupling $\overline{g}^2_{\mathrm{GF},0}$ for $L/a = 12$ against the range ϵ, where $\epsilon \leq 0.5$ can serve as definition of the trivial sector on the lattice. Results for larger ϵ take into account configurations which are considered non-trivial, but are nevertheless shown in gray for completeness. The point at $\epsilon = 1$ corresponds to the original definition $\overline{g}^2_{\mathrm{GF}}$ of the gradient flow coupling. *Left:* $c = 0.3$. *Right:* $c = 0.5$.

to the larger correlation discussed in sec. 7.2.2. Moreover, we see that the particular choice of ϵ has no big influence on the modified gradient flow coupling. We will use $\epsilon = 0.5$ in the following. In fig. 7.2, we compare the integrated autocorrelation time for the gradient flow coupling in its original and modified form. We find that the original coupling is affected by the bad sampling towards the continuum, which is particularly pronounced in the case of $c = 0.5$. In contrast, the modified coupling does suffer less severely from critical slowing down and its autocorrelation scales with $z \approx 2$ as expected (cf. (6.34)). In that sense, the modified gradient flow coupling can be considered to be safer.

7.3 Results

The full set of results for the two couplings $\overline{g}^2_{\mathrm{GF}}$ and $\overline{g}^2_{\mathrm{GF},0}$ is listed in tab. 7.3. On the coarser lattices ($L/a \in \{8, 12, 16\}$), the simulations show a clear difference between the two definitions. This suggests that in the studied volume ($L \sim 0.8\,\mathrm{fm}$), topologically non-trivial configurations play a role in accurately determining the value of $\overline{g}^2_{\mathrm{GF}}$. On the two finer lattices ($L/a \in \{20, 24\}$), we do not observe a difference due to the critical slowing down that affects the determination of the original coupling.

Since the results for $\overline{g}^2_{\mathrm{GF}}$ on the finer lattices ($L/a \in \{20, 24\}$) are biased, we conduct the continuum extrapolation only for $\overline{g}^2_{\mathrm{GF},0}$. To compare data of different lattice spacings we have to take into account an additional error being introduced by the way

	$c = 0.3$		$c = 0.5$	
L/a	$\overline{g}^2_{\rm GF}$	$\overline{g}^2_{\rm GF,0}$	$\overline{g}^2_{\rm GF}$	$\overline{g}^2_{\rm GF,0}$
8	5.647(3)(52)	5.631(3)(51)	9.313(21)(146)	9.140(14)(126)
12	5.894(5)(52)	5.875(4)(51)	10.243(43)(146)	9.983(23)(126)
16	5.924(10)(52)	5.908(8)(51)	10.487(117)(146)	10.255(54)(126)
20	*5.845(10)(52)	5.845(10)(51)	*10.135(74)(146)	10.132(74)(126)
24	*5.818(28)(52)	5.810(17)(51)	*10.128(123)(146)	10.021(96)(126)
∞	*5.818(62)	5.820(58)	*10.156(258)	10.179(210)

Table 7.3: Results for the (modified) gradient flow coupling and its continuum extrapolation. The first error is statistical, the second one stems from the uncertainty associated with the fixing of the physical volume. Biased values are denoted with an asterisk.

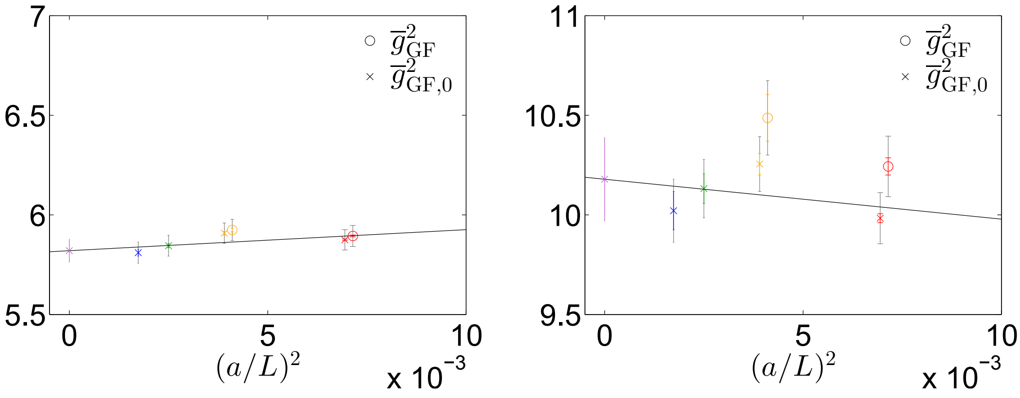


Figure 7.5: Continuum extrapolation of the modified gradient flow coupling (crosses). The unbiased data for the gradient flow coupling is also shown (open circles), the points being slightly shifted to the right for convenience. Statistical errors are displayed in color, errors from the line of constant physics in gray. *Left:* $c = 0.3$. *Right:* $c = 0.5$.

the physical volume was fixed, (7.2)[2]. This uncertainty turns out to be larger than the statistical errors (see tab. 7.3). We find that the data with $L/a > 8$ is well described by a fit linear in $(a/L)^2$, see fig. 7.5.

Hence, in conclusion, the modified gradient flow coupling $\overline{g}^2_{\rm GF,0}$, unlike its original counterpart $\overline{g}^2_{\rm GF}$, is not affected by the bad topology sampling in the studied volume.

---

[2] The error on $a/r_0(\beta)$ depends on β, but for simplicity we propagate its maximum value of 1% [32] globally.

8 Symanzik improvement for $N_f = 4^*$ massive, non-degenerate Wilson fermions

Before we begin the discussion of Symanzik improvement, we compile the main formulas which define the unimproved $N_f = 4^*$ theory of non-degenerate quark masses.

The Wilson fermion action (3.44) can be rewritten as

$$S_F^{\mathrm{W}}[\psi, \overline{\psi}, U] = a^4 \sum_x \overline{\psi}(D_W + M + m_{\mathrm{crit}})\psi \; , \tag{8.1}$$

where D_W is the massless Wilson–Dirac operator,

$$D_W = \frac{1}{2} \sum_\mu \left\{ \gamma_\mu \left(\nabla_\mu^* + \nabla_\mu \right) - a \nabla_\mu^* \nabla_\mu \right\} \; , \tag{3.43}$$

while the matrix

$$M = \mathrm{diag}(m_{q,u}, m_{q,d}, m_{q,s}, m_{q,c}) \tag{3.64}$$

incorporates the bare subtracted quark masses

$$m_{q,i} = m_{0,i} - m_{\mathrm{crit}} = \frac{1}{2a} \left[\frac{1}{\kappa_i} - \frac{1}{\kappa_{\mathrm{crit}}} \right] \; . \tag{3.65, 3.70}$$

For completeness, we recall the Wilson and the tree-level improved Lüscher–Weisz gauge action,

$$S_G^{\mathrm{W}}[U] = \frac{1}{g_0^2} \cdot \sum_{\mathcal{C} \in \mathcal{S}_0} \mathrm{tr}[\mathbb{1} - U(\mathcal{C})] \tag{3.39}$$

$$\text{and} \quad S_G^{\mathrm{LW}}[U] = \frac{1}{g_0^2} \sum_{i=0,1} c_i \sum_{\mathcal{C} \in \mathcal{S}_i} \mathrm{tr}[\mathbb{1} - U(\mathcal{C})] \; , \tag{3.112}$$

whose SF boundary improved counterparts, $S_G^{\mathrm{W,SF}}$ and $S_G^{\mathrm{LW,SF}}$, are given in (5.35) and (5.38), respectively.

The discretization errors of the different actions are of order $O(a)$ for S_F^{W}, $O(a^2)$ for S_G^{W}, and $O(a^2 g_0^2)$ for S_G^{LW}, as we have seen in sec. 3.3, sec. 3.2 and sec. 3.8, respectively. Following the lines of sec. 3.7 (which refers to $N_f = 4$), one needs to improve both the Wilson fermion action and the bilinears in order for cutoff effects in correlation functions with bilinears to be of order $O(a^2)$. The way this can be accomplished for $N_f = 4^*$ massive fermions is the subject of the present section. Although the choice of the gauge action will affect numerical results (in sec. 9), it plays no role in laying the theoretical foundations for improvement. Hence, unless otherwise noted, S_G will be left unspecified here. The full lattice action is given by $S = S_F^{\mathrm{W}} + S_G$.

We begin with the construction of mass-dependent improvement terms in the Symanzik effective theory, sec. 8.1. Subsequently, in sec. 8.2, we consider the corresponding improvement in the underlying lattice theory, in a mass-independent renormalization and improvement scheme. After addressing the drawbacks of such a scheme in conjunction with the presence of a dynamical charm quark, we discuss generic mass-dependent schemes in sec. 8.3 and how they circumvent the problems. Finally, in sec. 8.4, we specify the concrete mass-dependent scheme in which the determination of c_{sw}, to be described in sec. 9, will take place.

8.1 Symanzik effective theory

8.1.1 Spurionic symmetry

As described in sec. 3.7, the terms in the Symanzik effective theory are constructed such that they obey all the symmetries of the original lattice theory, in particular gauge symmetry, time reversal, parity and charge conjugation. An aspect which adds up to this in the case of the non-degenerate mass matrix in $N_f = 4^*$, is that the associated explicit $SU(4)_{fV}$ flavor symmetry breaking should happen in the Symanzik effective theory exactly in the way it does in the original theory. In other words, $SU(4)_{fV}$ should be recovered for the special case of $N_f = 4$ degenerate masses. This may be achieved by imposing the spurionic vector symmetry (D.1) described in app. D, written here in terms of the bare subtracted mass matrix:

$$M \;\rightarrow\; M' = VMV^{\dagger}\,, \qquad V \in SU(4)_{fV}\,. \tag{8.2}$$

The most general mass-dependent improvement terms may then be constructed for both the action and bilinears by requiring them to maintain the behavior of the original terms under all symmetry transformations, in particular $SU(4)_{fV}$. Concerning the action, this means *invariance*,

$$S \overset{V}{\to} S \,, \tag{8.3}$$

while bilinears in general transform *covariantly*.

$$O \overset{V}{\to} O' \equiv O[\psi', \overline{\psi}', U] \,, \tag{8.4}$$

where $\psi' = V\psi$ and $\overline{\psi}' = \overline{\psi}V^{\dagger}$ are the transformed fields. In that case, the Green functions of operators O as path integral averages,

$$\langle O \rangle = \frac{1}{Z} \int \mathcal{D}[\psi, \overline{\psi}, U] \; e^{-S[\psi,\overline{\psi},U]} O[\psi, \overline{\psi}, U] \,, \tag{3.46}$$

also transform covariantly:

$$\big\langle O \big\rangle \overset{V}{\to} \big\langle O' \big\rangle \;. \tag{8.5}$$

A similar approach has been used in the context of chiral perturbation theory [132]. We stress again, however, that (8.3)-(8.5) is *not* the actual transformation behavior for non-degenerate masses. Rather, it *would* be the behavior if the mass matrix transformed according to (8.2). This is the case for degenerate quarks only, however.

In the following, we construct the improvement terms for $N_f = 4^*$ explicitly. Note that, as this is a straight extension to the case of $N_f = 4$ described in sec. 3.7, many comments on the procedure found there hold for the present case without changes and will thus not be repeated here. The changes from $N_f = 4$ to $N_f = 4^*$ concern the mass-dependent improvement terms only. They go back to the fact that apart from the bare subtracted mass matrix M, terms that incorporate the trace $\mathrm{Tr}[M]$ may appear as well. Obviously, this term is also invariant under (8.2). For the previously considered case of $N_f = 4$ degenerate masses, however, the terms with and without trace were proportional to each other,

$$\mathrm{Tr}[M]\mathbb{1}_4 = N_f m_q \mathbb{1}_4 = N_f M \,, \tag{8.6}$$

so that there was no need to consider them independently. In contrast, regarding non-degenerate masses now, we have to account for this difference. This leads to quite a few additional terms, as we will see below. Note that most of the results have been published in a slightly different form before [133].

8.1.2 Symanzik effective action

We begin with the construction of the Symanzik effective action. Splitting the effective Lagrangian into a mass-dependent and a mass-independent part like in $N_f = 4$ (cf. sec. 3.7),

$$\mathcal{L}_{\text{Sym}} = \mathcal{L}^0_{\text{Sym}} + \mathcal{L}^M_{\text{Sym}} \,, \tag{8.7a}$$

it reads

$$\begin{aligned} \mathcal{L}^0_{\text{Sym}} &= \frac{1}{4} F_{\mu\nu} F_{\mu\nu} + \overline{\psi} \slashed{D} \psi \\ &+ a\rho_1 \, \overline{\psi} D_\mu D_\mu \psi \\ &+ a\rho_2 \, \overline{\psi} i\sigma_{\mu\nu} F_{\mu\nu} \psi \end{aligned} \tag{8.7b}$$

and

$$\begin{aligned} \mathcal{L}^M_{\text{Sym}} &= \overline{\psi} M \psi \\ &+ \rho \, \text{Tr}[M] \, \overline{\psi}\psi \\ &+ a\sigma_1 \, \overline{\psi} M^2 \psi \\ &+ a\sigma_2 \, \text{Tr}[M] \, \overline{\psi} M \psi \\ &+ a\sigma_3 \, \text{Tr}[M^2] \, \overline{\psi}\psi \\ &+ a\sigma_4 \, (\text{Tr}[M])^2 \, \overline{\psi}\psi \\ &+ a\sigma_5 \, \text{Tr}[M] \, \text{tr}[F_{\mu\nu} F_{\mu\nu}] \\ &+ a\sigma_6 \, \overline{\psi} \slashed{D} M \psi \\ &+ a\sigma_7 \, \text{Tr}[M] \, \overline{\psi} \slashed{D} \psi \,. \end{aligned} \tag{8.7c}$$

Note that the $N_f = 4$ expression (3.89) is recovered under use of (8.6). In particular, the terms with σ_{1-4} and those with σ_{6-7} reduce to one term, respectively. The terms with ρ_1 and σ_{6-7} can be eliminated in the case of on-shell improvement under use of $\slashed{D}^2 = D^2 + \frac{1}{2}\sigma_{\mu\nu} F_{\mu\nu}$ and the Euclidean space Dirac equation to leading order $\mathcal{O}(a^0)$,

$$(\slashed{D} + M + \rho \, \text{Tr}[M]) \, \psi = 0 \,. \tag{8.8}$$

We show an example explicitly. The ρ_1 term in (8.7b) may be rewritten as

$$\overline{\psi} D_\mu D_\mu \psi = \overline{\psi} \left(\slashed{D}^2 - \frac{1}{2} \sigma_{\mu\nu} F_{\mu\nu} \right) \psi = \overline{\psi} \left((M + \rho \, \text{Tr}[M])^2 - \frac{1}{2} \sigma_{\mu\nu} F_{\mu\nu} \right) \psi \,, \tag{8.9}$$

where the term involving the field strength tensor already appears in (8.7b) with ρ_2, while the others are present in (8.7c) with $\sigma_1, \sigma_2, \sigma_4$. Similarly, one may get rid of the σ_6 and σ_7 terms. Note that alternatively, the latter can be eliminated by a mass-dependent rescaling of the fermion fields:

$$\psi \to (1 - a\sigma_6 M - a\sigma_7 \mathrm{Tr}[M])\ \psi\ . \tag{8.10}$$

This way, the σ_6 and σ_7 terms are cancelled by the rescaling of $\overline{\psi}\not{D}\psi$ in (8.7b), while the additional terms associated with the rescaling of $\overline{\psi}M\psi$ and $\rho\ \mathrm{Tr}[M]\overline{\psi}\psi$ in (8.7c) can be absorbed in the coefficients $\sigma_1, \sigma_2, \sigma_4$. The Symanzik effective Lagrangian (8.7) hence becomes

$$\begin{aligned}
\mathcal{L}_{\mathrm{Sym}} = {} & \frac{1}{4} F_{\mu\nu} F_{\mu\nu} + \overline{\psi}\not{D}\psi \\
& + \overline{\psi} M \psi + \rho\ \mathrm{Tr}[M]\ \overline{\psi}\psi \\
& + a\rho_2\ \overline{\psi} i \sigma_{\mu\nu} F_{\mu\nu} \psi \\
& + a\sigma_1\ \overline{\psi} M^2 \psi \\
& + a\sigma_2\ \mathrm{Tr}[M]\ \overline{\psi} M \psi \\
& + a\sigma_3\ \mathrm{Tr}[M^2]\ \overline{\psi}\psi \\
& + a\sigma_4\ (\mathrm{Tr}[M])^2\ \overline{\psi}\psi \\
& + a\sigma_5\ \mathrm{Tr}[M]\ \mathrm{tr}[F_{\mu\nu} F_{\mu\nu}]\ .
\end{aligned} \tag{8.11}$$

Note that although we have used the notation M here, the mass matrix in the Symanzik effective theory actually differs from (3.64) by a multiplicative renormalization. We will consider the corresponding improvement terms for the lattice action in sec. 8.2.1.

8.1.3 Symanzik effective bilinears

Axial current

We consider the axial current first. For the construction of the improvement terms, we employ the generator basis, see sec. 3.6 and app. B.1, in terms of which the axial current may be written as

$$A^c_\mu(x) = \overline{\psi}(x)\ \gamma_\mu \gamma_5 T^c\ \psi(x)\ , \qquad c \in \{0, \ldots, 15\}\ . \tag{3.73}$$

It is convenient to work in this basis, as it naturally allows to take into account the different behavior of the different components under $SU(4)_{fV}$ flavor transformations. The component with $c = 0$ is the flavor *singlet* current. The 15 other components, the flavor *non-singlet* currents, divide into 3 *neutral* ($c = 3, 8, 15$) and 12 *charged* currents. The names neutral/charged refer to the behavior of those non-singlet currents under charge conjugation. The neutral currents transform trivially due to their diagonal generators, whereas the charged currents with non-diagonal generators do not, see app. B.3. For the transformation of the generators under $V \in SU(4)_{fV}$ flavor transformations, we introduce the notation

$$V^\dagger T^c V = R^{cd}(V) T^d \ . \tag{8.12}$$

Note that in particular $R^{00}(V) = 1$ and $R^{0d}(V) = 0 \ \ \forall\, d \neq 0$. For the transformation of the axial current, we find

$$A_\mu^c = \overline{\psi}\gamma_\mu\gamma_5 T^c \psi \ \xrightarrow{V}\ \overline{\psi}V^\dagger\gamma_\mu\gamma_5 T^c V\psi = R^{cd}(V) A_\mu^d \ , \tag{8.13}$$

i.e. the 15 flavor non-singlet currents A_μ^c with $c \neq 0$ transform according to the adjoint representation of $SU(4)$, whereas A_μ^0 is invariant as stated above. Following the discussion of sec. 8.1.1, the terms in the Symanzik effective operator have to maintain the transformation behavior (8.13) of the axial current under spurionic vector symmetry (8.2).

For the *flavor non-singlet currents* ($c \neq 0$), the Symanzik effective expansion reads

$$(A_{\text{Sym}})_\mu^c = A_\mu^c + a \cdot \Big(\omega_0 \partial_\mu P^c + \omega_1 \mathrm{Tr}[M] A_\mu^c + \omega_2 \overline{\psi}\gamma_\mu\gamma_5 \left\{T^c, M\right\}\psi + \omega_3 \mathrm{Tr}[T^c M] A_\mu^0\Big) \ . \tag{8.14a}$$

In contrast to (3.90), we have immediately dropped the ω_5 term here. The terms with ω_2 and ω_3 are relevant for non-degenerate masses only. In $N_f = 4$, the one with ω_2 can be absorbed in the ω_1 term, while the ω_3 term vanishes. Moreover, the two terms contained in the anticommutator of (8.14a) are a priori independent. They have to contribute with equal weight, however, due to charge conjugation symmetry (see app. B.3). In the case of the *flavor singlet current* ($c = 0$), one ends up with only two mass-dependent improvement terms,

$$(A_{\text{Sym}})_\mu^0 = A_\mu^0 + a \cdot \Big(\overline{\omega}_0 \partial_\mu P^0 + \overline{\omega}_1 \mathrm{Tr}[M] A_\mu^0 + \overline{\omega}_2 \mathrm{Tr}[M A_\mu]\Big) \ . \tag{8.14b}$$

The term with $\overline{\omega}_2$ corresponds directly to the one with ω_2 in (8.14a), and again accounts for non-degeneracy. It has been reformulated in terms of A_μ, which is the 4×4-matrix that incorporates all the currents (cf. (B.2b) in app. B). There is no $\overline{\omega}_3$ term, since the flavor singlet equivalent to the ω_3 term in (8.14a) is already given by the $\overline{\omega}_1$ term. In app. B.2, the required behavior under $SU(4)_{fV}$ is explicitly verified for the $O(a)$ terms in (8.14).

Pseudoscalar current

For the pseudoscalar current P, the mass-dependent $O(a)$ terms (8.14a) and (8.14b) are very similar, the only difference being the replacement

$$\gamma_\mu \gamma_5 \to \gamma_5 \tag{8.15}$$

in all terms. Moreover, as is well known from $N_f = 4$, a mass-independent improvement term (similar to the one with ω_0 in (8.14a)) does not appear in the expression of the Symanzik effective pseudoscalar current, see (3.91). Hence, we find

$$(P_{\text{Sym}})^c = P^c + a \cdot \left(\omega_1' \mathrm{Tr}[M] P^c + \omega_2' \overline{\psi} \gamma_5 \{T^c, M\} \psi + \omega_3' \mathrm{Tr}[T^c M] P^0\right) \tag{8.16a}$$

$$\text{and} \qquad (P_{\text{Sym}})^0 = P^0 + a \cdot \left(\overline{\omega}_1' \mathrm{Tr}[M] P^0 + \overline{\omega}_2' \mathrm{Tr}[MP]\right) \tag{8.16b}$$

for the flavor non-singlets and singlet, respectively.

8.2 $O(a)$ improvement in the mass-independent renormalization scheme

8.2.1 Improvement of the Wilson action

The improved lattice action can be derived from the Symanzik effective action (8.11) by replacing all the continuum terms by discretized counterterms, cf. sec. 3.7. The $O(a)$ improved lattice action reads

$$S = a^4 \sum_x (\mathcal{L}_{F,I} + \mathcal{L}_{G,I}) \tag{8.17a}$$

with

$$\begin{aligned}\mathcal{L}_{F,I} &= \overline{\psi} D_W \psi + \overline{\psi} M \psi \\ &+ ac_{\text{sw}}\ \overline{\psi}\frac{i}{4}\sigma_{\mu\nu}F_{\mu\nu}\psi \\ &+ ac_1\ \overline{\psi}M^2\psi \\ &+ ac_2\ \text{Tr}[M]\overline{\psi}M\psi \\ &+ ac_3\ \text{Tr}[M^2]\overline{\psi}\psi \\ &+ ac_4\ (\text{Tr}[M])^2\overline{\psi}\psi \end{aligned} \qquad (8.17b) \qquad \text{and} \qquad \begin{aligned}\mathcal{L}_{G,I} &= \mathcal{L}_G[U] \\ &+ ac_5\ \text{Tr}[M]\ \text{tr}[F_{\mu\nu}F_{\mu\nu}]\ .\end{aligned} \qquad (8.17c)$$

We stress that the improvement coefficients c_{sw} and c_{1-5} all depend explicitly on the bare coupling g_0^2. Note that $\mathcal{L}_{F,I}$ does not contain an expression which directly corresponds to the ρ term in the Symanzik effective action (8.11). This is because the appearance of the ρ term reflects the difference between the flavor singlet and non-singlet masses, which is taken into account in the renormalization prescription, as we will see below.

Similar to the case of $N_f = 4$, all but the c_{sw} term may be absorbed in the bare parameters and accounted for in the renormalization prescription of those. The absorption of the c_5 term in the bare coupling is completely equivalent to (3.94) and can literally be taken over, except for the replacement $m_q \to \text{Tr}[M]$:

$$\mathcal{L}_{G,I} = \mathcal{L}_G\big|_{g_0^2 \to g_0^2(1-2ac_5 g_0^2\ \text{Tr}[M])}\ . \tag{8.18}$$

With the abbreviation

$$b_g = -2c_5 N_f \cdot g_0^2\ , \tag{8.19}$$

the modified bare coupling (cf. (3.95)) reads

$$g_0^2\ \to\ \tilde{g}_0^2 = g_0^2\left(1 + ab_g(g_0^2)\ \text{Tr}[M]/N_f\right)\ , \tag{8.20}$$

and its renormalization is equivalent to the $N_f = 4$ case (cf. (3.98)):

$$g_R^2 = Z_g(\tilde{g}_0^2, a\mu)\ \tilde{g}_0^2\ . \tag{8.21}$$

The absorption of the fermionic improvement terms c_{1-4} is slightly more complicated in $N_f = 4^*$ than in $N_f = 4$ (cf. (3.97))[1]:

$$m_{q,i} \to \widetilde{m}_{q,i} = m_{q,i} + ac_1 m_{q,i}^2 + ac_2 \mathrm{Tr}[M] m_{q,i} + ac_3 \mathrm{Tr}[M^2] + ac_4 (\mathrm{Tr}[M])^2 \;. \tag{8.22}$$

After introducing a new set of improvement coefficients,

$$b_m = c_1 \;, \qquad d_m = c_1 + N_f c_3 \;, \tag{8.23a}$$

$$\bar{b}_m = c_2 \;, \qquad \bar{d}_m = c_2 + N_f c_4 \;, \tag{8.23b}$$

the reparametrization of the bare quark masses (8.22) becomes

$$\begin{aligned} m_{q,i} \;\to\; \widetilde{m}_{q,i} &= m_{q,i} + ab_m m_{q,i}^2 + a\bar{b}_m \mathrm{Tr}[M] m_{q,i} \\ &\quad + a\left(d_m - b_m\right) \frac{\mathrm{Tr}[M^2]}{N_f} + a\left(\bar{d}_m - \bar{b}_m\right) \frac{(\mathrm{Tr}[M])^2}{N_f} \;. \end{aligned} \tag{8.24}$$

The coefficients $b_m, \bar{b}_m$ and $d_m, \bar{d}_m$ correspond to flavor non-singlet and singlet improvement terms, respectively[2]. In contrast to the $N_f = 4$ case considered in sec. 3.7, in $N_f = 4^*$ one has to differentiate for each of those between dynamical quark loop ($\bar{b}_m, \bar{d}_m$) and valence quark (b_m, d_m) effects, see also [133]. The form (8.24) is more convenient in the context of renormalization, because the flavor non-singlet and singlet improvement terms renormalize differently (cf. (3.63, 3.66, E.1, E.2)). The renormalized improved quark masses thus read [133]

$$\begin{aligned} m_{\mathrm{R}}^i = Z_m(\tilde{g}_0^2, a\mu) \Bigg[& m_{q,i} + \left(r_m(\tilde{g}_0^2) - 1\right) \frac{\mathrm{Tr}[M]}{N_f} + a \cdot \Bigg\{ b_m(g_0^2) m_{q,i}^2 + \bar{b}_m(g_0^2) \mathrm{Tr}[M] m_{q,i} \\ & + \left(r_m(g_0^2) d_m(g_0^2) - b_m(g_0^2)\right) \frac{\mathrm{Tr}[M^2]}{N_f} + \left(r_m(g_0^2) \bar{d}_m(g_0^2) - \bar{b}_m(g_0^2)\right) \frac{(\mathrm{Tr}[M])^2}{N_f} \Bigg\} \Bigg] \;, \end{aligned} \tag{8.25}$$

where we show the dependence on the (modified) bare coupling explicitly again here, restricting ourselves to contributions which affect the order $O(a)$. The above formula is explicitly derived in app. E.2. Note that the appearance of the mass-dependent $O(a^0)$ renormalization term is related to the one of the ρ term in (8.11).

Together with (8.20), (8.21), (8.24) and (8.25), the action for $N_f = 4^*$ in the *mass-independent* renormalization and improvement scheme again requires only the Sheikholeslami–Wohlert term for $O(a)$ improvement. We recall it here for convenience

---

[1] The dependence on the bare coupling g_0^2 is suppressed for a moment.

[2] For the singlet, this can easily be seen by taking the sum over all quark flavors, $\sum_i \widetilde{m}_{q,i}$.

(cf. (3.100, 3.101)):

$$S_{\mathrm{I}}^{\mathrm{W}}[\psi, \overline{\psi}, U] = S_G^{\mathrm{W}}[U] + S_F^{\mathrm{W}}[\psi, \overline{\psi}, U] + \delta S_{\mathrm{sw}}[\psi, \overline{\psi}, U] \;, \tag{8.26}$$

with

$$\delta S_{\mathrm{sw}}[\psi, \overline{\psi}, U] = a^5 c_{\mathrm{sw}}(g_0^2) \; \sum_x \overline{\psi}(x) \frac{i}{4} \sigma_{\mu\nu} F_{\mu\nu}(x) \psi(x) \;. \tag{8.27}$$

The lattice field strength tensor $F_{\mu\nu}$ is chosen to be the clover term $F_{\mu\nu}^{(\mathrm{clov})}$ (3.102) as before.

8.2.2 Improvement of bilinears

The improved bilinears on the lattice, first considered for non-degenerate masses in [133], can easily be derived from the effective bilinears given in (8.14) and (8.16).

Axial current (generator basis)

After replacing the continuum expressions by lattice counterparts and relabelling[3] the coefficients, the renormalized axial currents (cf. (8.14)) read

$$\begin{aligned}(A_{\mathrm{RI}})_\mu^c = Z_A(\tilde{g}_0^2) \Big[A_\mu^c + a \cdot \Big(& c_A(g_0^2) \tilde{\partial}_\mu P^c + \overline{b}_A(g_0^2) \mathrm{Tr}[M] A_\mu^c + b_A(g_0^2) \frac{1}{2} \overline{\psi} \gamma_\mu \gamma_5 \left\{ T^c, M \right\} \psi \\ & + f_A(g_0^2) \mathrm{Tr}[T^c M] A_\mu^0 \Big) \Big] \end{aligned} \tag{8.28a}$$

$$(A_{\mathrm{RI}})_\mu^0 = \overline{Z}_A(\tilde{g}_0^2, a\mu) \Big[A_\mu^0 + a \cdot \Big(\overline{c}_A(g_0^2) \tilde{\partial}_\mu P^0 + \overline{d}_A(g_0^2) \mathrm{Tr}[M] A_\mu^0 + d_A(g_0^2) \mathrm{Tr}[M A_\mu] \Big) \Big] \;. \tag{8.28b}$$

Note that similar to the case of the masses, (8.25), the flavor non-singlet and singlet contributions renormalize differently[4]. The flavor singlet renormalization constant $\overline{Z}_A$ can be expressed in terms of the flavor non-singlet counterpart as (see e.g. [133])

$$\overline{Z}_A(\tilde{g}_0^2, a\mu) = Z_A(\tilde{g}_0^2) \cdot r_A(\tilde{g}_0^2, a\mu) \;. \tag{8.29}$$

We briefly comment on the significance of the improvement terms in (8.28). First

---

[3] We follow the notation employed in [133].

[4] Furthermore, dynamical and valence quark effects are distinguished by the notation of the improvement coefficients, $\overline{b}_A, \overline{d}_A$ and b_A, d_A, as in the case of the quark masses, cf. (8.24).

of all, for the *charged* flavor non-singlet currents ($c \neq 3, 8, 15$), the f_A term does not contribute, because the associated generators have zeros on the diagonal and M is a diagonal matrix. Moreover, the mass dependence of the $\overline{b}_A, \overline{d}_A$ terms is enclosed in the explicit factor $\mathrm{Tr}[M]$, which reflects the fact that these terms represent sea quark (loop) contributions. One may therefore conclude that they contribute only at a higher order in perturbation theory. The mass dependence of the other terms, in contrast, stems from the valence quark propagators involved in the axial current.

Axial current (flavor basis)

For practical use, it is often more convenient to work in the *flavor basis*, see (3.75). The 12 *off-diagonal* currents ($i \neq j$) are linear combinations of the 12 *charged* flavor non-singlet currents in the generator basis (cf. (B.2b)):

$$A_\mu^{ij} = \left(\sum_{c \neq 0,3,8,15} A_\mu^c T^c \right)^{ij} . \tag{8.30}$$

The quantities in the sum on the right hand side share the same renormalization and improvement properties, (8.28a). Hence, in the flavor basis, the *off-diagonal* renormalized and improved flavor current can be written as

$$\begin{aligned}(A_{\mathrm{RI}})_\mu^{ij} =& Z_A(\tilde{g}_0^2) \times \\ & \left[A_\mu^{ij} + a \cdot \left(c_A(g_0^2) \tilde{\partial}_\mu P^{ij} + \overline{b}_A(g_0^2) \mathrm{Tr}[M] A_\mu^{ij} + \frac{b_A(g_0^2)}{2} \overline{\psi} \gamma_\mu \gamma_5 \left\{ \Delta^{ij}, M \right\} \psi \right) \right] .\end{aligned} \tag{8.31}$$

We can simplify the expression under use of

$$\left\{ \Delta^{ij}, M \right\} = \Delta^{ij} M + M \Delta^{ij} = \Delta^{ij} m_{q,j} + m_{q,i} \Delta^{ij} = (m_{q,i} + m_{q,j}) \, \Delta^{ij} , \tag{8.32}$$

and—since it is valid only up to $O(a^2)$ anyway—factorize it according to

$$\begin{aligned}(A_{\mathrm{RI}})_\mu^{ij} =& Z_A(\tilde{g}_0^2) \times \\ & \left[1 + a \overline{b}_A(g_0^2) \mathrm{Tr}[M] + a \frac{b_A(g_0^2)}{2} (m_{q,i} + m_{q,j}) \right] \left[A_\mu^{ij} + a c_A(g_0^2) \tilde{\partial}_\mu P^{ij} \right] .\end{aligned} \tag{8.33}$$

In this convenient form, the improved axial current (3.103) as well as the mass effects are factored out, and the renormalized and improvement axial current resembles the one in $N_f = 4$, cf. (3.105). Alternatively, in the b_A term, one may replace the average bare subtracted quark mass by the PCAC mass[5],

$$\frac{1}{2}(m_{q,i} + m_{q,j}) \;\rightarrow\; m^{ij}_{\mathrm{PCAC}} \,. \tag{8.34}$$

In contrast to the simple off-diagonal case, the *diagonal* flavor currents ($i = j$) are linear combinations of uncharged flavor *non-singlet* currents ($c = 3, 8, 15$) and the flavor *singlet* current ($c = 0$), e.g.

$$A^{ii}_\mu = \left(\sum_{c=0,3,8,15} A^c_\mu T^c \right)^{ii} = \sum_{c=0,3,8,15} A^c_\mu \left(T^c\right)^{ii} \,. \tag{8.35}$$

Due to the different renormalization and improvement pattern of the singlet and non-singlet contributions, (8.28a) and (8.28b), the behavior of A^{ii}_μ is more involved. However, we do not pursue its discussion here, as we will only deploy off-diagonal currents.

Pseudoscalar current (generator and flavor basis)

For completeness, we also mention the renormalized and improved pseudoscalar current in the *generator basis* (cf. (8.16)),

$$\begin{aligned}(P_{\mathrm{RI}})^c = Z_P(\tilde{g}_0^2, a\mu)\Big[P^c + a\cdot\Big(&\bar{b}_P(g_0^2)\mathrm{Tr}[M]P^c + b_P(g_0^2)\frac{1}{2}\bar{\psi}\gamma_\mu\gamma_5\left\{T^c, M\right\}\psi \\ &+ f_P(g_0^2)\mathrm{Tr}[T^c M]P^0\Big)\Big]\end{aligned} \tag{8.36a}$$

$$(P_{\mathrm{RI}})^0 = \bar{Z}_P(\tilde{g}_0^2, a\mu)\Big[P^0 + a\cdot\Big(\bar{d}_P(g_0^2)\mathrm{Tr}[M]P^0 + d_P(g_0^2)\mathrm{Tr}[MP]\Big)\Big] \,, \tag{8.36b}$$

and the corresponding *off-diagonal* components in the *flavor basis* (cf. (8.33)),

$$(P_{\mathrm{RI}})^{ij} = Z_P(\tilde{g}_0^2, a\mu)\left[1 + a\cdot\bar{b}_P(g_0^2)\mathrm{Tr}[M] + a\cdot\frac{b_P(g_0^2)}{2}(m_{q,i} + m_{q,j})\right] P^{ij} \,. \tag{8.37}$$

---

[5] This is because they are related by a finite renormalization, see (E.16, E.17), which can be absorbed in the improvement coefficients.

8.2.3 Problems in the presence of a massive charm quark

In the previous sections, we have obtained the expressions for the renormalized and $O(a)$ improved action and bilinears for non-degenerate $N_f = 4^*$ Wilson fermions in a *mass-independent scheme*. The main advantages of such a scheme, responsible for its wide use, are that it facilitates both the determination of renormalization constants and improvement coefficients as well as the renormalization group equations which govern the scale evolution of renormalized parameters (cf. sec. 2.3). Moreover, it is the natural scheme to study explicitly the effects of non-vanishing masses, as these are analytically expressed in the formulas, see e.g. (8.20, 8.21) and (8.25).

We now examine the applicability of a mass-independent scheme in (large volume) lattice simulations at physical quark masses in $N_f = 4^*$, i.e. including the charm quark. The crucial factor here is the size of the mass terms which multiply the improvement coefficients. For typical lattice spacings $a \leq 0.1$ fm, assuming[6] $m_{q,s} \approx 100$ MeV and $m_{q,c} \approx 1$ GeV (cf. tab. 2.1), these terms are roughly of the sizes

$$am_{q,s} \lesssim 0.05 \tag{8.38a}$$

$$\text{and} \qquad am_{q,c} \lesssim 0.5\,. \tag{8.38b}$$

In $N_f \leq 3$ simulations with small lattice spacings and quark masses, where (8.38a) holds for all the dynamical quarks, a viable way to account for the small mass-dependent improvement terms is to treat their coefficients perturbatively[7] (or even neglect them). In $N_f = 4^*$, however, the effects of terms containing $am_{q,c}$ in the improved theory are expected to be numerically much more important[8], due to (8.38b), thus requiring a non-perturbative treatment of the associated improvement coefficients. In consideration of the numerous appearance of those for non-degenerate quark flavors, cf. (8.20), (8.25), (8.31) and (8.37), this, however, seems unfeasible.

There is another, related issue which concerns the propagation of the unavoidable uncertainties in both the improvement coefficients and renormalization constants. These are amplified in $N_f = 4^*$ by the large charm mass and thus are likely to lead to significant errors. The issue arises at the $O(a^0)$ level in the context of quark mass

---

[6] The exact values depend on the renormalization scheme and scale, but this is not of importance for the qualitative statement to be made here.

[7] Note that this does not necessarily hold for observables that include charm valence quarks. An example is given in [133].

[8] Note that it is well-known that $O(a)$-improvement in the mass-independent scheme breaks down if $am_{q,c} \approx 1$, see e.g. [57, 134], so that (8.38b) can already be considered possibly threatening. We assume, however, that this is not the case here.

renormalization in the Wilson formulation already, where we have

$$m_{\mathrm{R}}^{i} = Z_m(g_0^2, a\mu)\left[m_{q,i} + \left(r_m(g_0^2) - 1\right) \mathrm{Tr}\,[M]\,/N_f\right] \,, \tag{3.66}$$

with the ratio of the flavor singlet and non-singlet renormalization constants,

$$r_m(g_0^2) = \frac{Z_{m^0}(g_0^2, a\mu)}{Z_m(g_0^2, a\mu)} \,. \tag{3.63}$$

Although the corrections of $r_m(g_0^2)$ to 1 are of order $O(g_0^4)$, they may be of significant magnitude and require a non-perturbative treatment, see app. E in [95]. In the presence of a heavy dynamical charm quark, they get blown up by a factor

$$\mathrm{Tr}[M] \overset{(8.38)}{\approx} m_{q,c} \,. \tag{8.39}$$

Therefore, even a small error in $r_m(g_0^2)$ may lead to a huge error in m_{R}^{i}.

At the next order in the lattice spacing, a similar situation emerges for the improved quark mass of any flavor i, which may be written as (cf. (8.22)),

$$\begin{aligned} \widetilde{m}_{q,i} = m_{q,i} + ac_1\, m_{q,i}^2 + ac_2\, m_{q,i} m_{q,c} + a\,(c_3 + c_4)\, m_{q,c}^2 \\ + O(am_{q,j} m_{q,k}) \ \text{ with } j \neq k;\ j,k \in \{u,d,s,c\} \,. \end{aligned} \tag{8.40}$$

Here, the charm mass appears quadratically in association with two ($i \neq c$) or four ($i = c$) improvement coefficients c_i. The relative corrections are particularly large for the light masses with $i = u, d$:

$$\widetilde{m}_{q,u} \overset{(8.40)}{=} m_{q,u} + \ldots + a\,(c_3 + c_4)\, m_{q,c}^2 \overset{(8.38b)}{\approx} (c_3 + c_4)\, m_{q,c} \,. \tag{8.41}$$

Consequently, for observables in the light sector (that do not involve charm valence quarks), the large charm mass cutoff and renormalization effects and assigned uncertainties in the mass-independent scheme would likely surpass the small physical effects of a massive dynamical charm quark (see e.g. [135]) that one is interested in.

In summary, a *mass-independent* renormalization and improvement scheme is inappropriate to describe the effects of a dynamical charm quark in the framework of the $N_f = 4^*$ theory, due to the large number of improvement coefficients for which a non-perturbative treatment is needed, as well as the amplification of uncertainties that concerns light observables in particular. In the next section, we will discuss a *mass-*

dependent renormalization and improvement scheme which avoids these problems.

8.3 $O(a)$ improvement in the mass-dependent renormalization scheme

In the *mass-independent renormalization scheme* used so far, the renormalization conditions are imposed at vanishing masses, i.e. in the massless theory. Correspondingly, apart from a possible dependence on the renormalization scale $a\mu$, all the renormalization constants depend only on g_0^2, or $\widetilde{g}_0^2$ in the improved theory:

$$Z \equiv Z(\widetilde{g}_0^2)\ . \tag{8.42}$$

In the *mass-dependent renormalization scheme*, the mass-dependent improvement terms are shifted from an explicit appearance in the renormalization prescriptions to an implicit appearance in the renormalization constants, for which the renormalization conditions are imposed at finite masses:

$$\widetilde{Z} \equiv \widetilde{Z}(g_0^2, aM)\ . \tag{8.43}$$

The absorption of all mass effects is reflected in the notation $\widetilde{Z}$, and makes the use of $\widetilde{g}_0^2$ as its argument obsolete. It is instructive to reconsider explicitly the improvement of the Wilson action and the bilinears in the *mass-dependent* scheme, to see how it avoids the problems discussed in sec. 8.2.3.

8.3.1 Improvement of the coupling

In the unimproved theory, the renormalization of the coupling reads

$$g_R^2 = Z_g(g_0^2, a\mu)\ g_0^2\ , \tag{3.61}$$

irrespective of the renormalization scheme. In the improved theory and the *mass-independent* scheme, we found

$$\begin{aligned} g_R^2 &\overset{(8.21)}{=} Z_g(\widetilde{g}_0^2, a\mu)\ \widetilde{g}_0^2 \\ &\overset{(8.20)}{=} Z_g(\widetilde{g}_0^2, a\mu)\ (1 + ab_g(g_0^2)\mathrm{Tr}[M]/N_f)\ g_0^2\ . \end{aligned} \tag{8.44}$$

We stress that Z_g is the same function (determined by the same renormalization condition) as in (3.61), only the argument at which it is evaluated changes. In contrast, in the *mass-dependent* scheme, (8.44) becomes

$$g_R^2 = \widetilde{Z}_g(g_0^2, a\mathrm{Tr}\,[M]\,, a\mu)\; g_0^2 \,. \tag{8.45}$$

Note that the masses enter the renormalization constant via *sea quarks* (cf. (8.44)). The comparison of these rather simple renormalization prescriptions illustrate the advantages and disadvantages of the two different schemes. First, the *mass-dependent* scheme involves only one constant. Second, the renormalization constant $\widetilde{Z}_g$ may be employed directly at the non-vanishing masses $a\mathrm{Tr}\,[M]$ one is interested in. In this way, one avoids the situation in the *mass-independent* scheme, where the need to connect the massive theory with the massless theory (by the functional mass dependence as it is explicitly contained e.g. in (8.44)) may introduce large uncertainties, as discussed in sec. 8.2.3. On the other hand, the explicit information on the mass-dependence gets lost in the mass-dependent scheme[9]. Moreover, the imposition of the renormalization conditions is certainly more involved at finite masses.

8.3.2 Improvement of bilinears

In the *mass-independent* scheme, the renormalized and improved off-diagonal axial current has been derived in sec. 8.2.2:

$$\begin{aligned}(A_{\mathrm{RI}})_\mu^{ij} =& Z_A(\widetilde{g}_0^2)\times \\ & \left[1 + a\cdot \overline{b}_A(g_0^2)\mathrm{Tr}[M] + a\cdot \frac{b_A(g_0^2)}{2}(m_{q,i}+m_{q,j})\right]\left[A_\mu^{ij} + a\cdot c_A(g_0^2)\partial_\mu P^{ij}\right]\;.\end{aligned} \tag{8.33}$$

In the *mass-dependent* scheme, the term with $\overline{b}_A$ is absorbed in the renormalization constant, such that (8.33) becomes

$$(A_{\mathrm{RI}})_\mu^{ij} = \widetilde{Z}_A(g_0^2, a\mathrm{Tr}\,[M])\left[1 + a\cdot \frac{b_A(g_0^2)}{2}(m_{q,i}+m_{q,j})\right]\left[A_\mu^{ij} + a\cdot c_A(g_0^2)\partial_\mu P^{ij}\right]\;. \tag{8.46a}$$

---

[9] This statement refers to the situation where $\widetilde{Z}_g$ is known only for one specific $a\mathrm{Tr}\,[M]$.

This form is convenient especially for light valence quarks ($i, j \neq c$), for which the b_A term is rather small and may be treated as before in $N_f \leq 3$. Alternatively, also the b_A term could be absorbed, at the cost of the renormalization constant becoming (valence quark) flavor dependent:

$$(A_{\mathrm{RI}})^{ij}_{\mu} = \widetilde{Z}^{ij}_{A}(g_0^2, aM) \left[A^{ij}_{\mu} + a \cdot c_A(g_0^2) \partial_{\mu} P^{ij} \right] . \tag{8.46b}$$

The corresponding equations for the pseudoscalar current read

$$(P_{\mathrm{RI}})^{ij} = \widetilde{Z}_P(g_0^2, a\mathrm{Tr}\,[M]) \left[1 + a \cdot \frac{b_P(g_0^2)}{2} (m_{q,i} + m_{q,j}) \right] P^{ij} \tag{8.47a}$$

$$\text{and} \quad (P_{\mathrm{RI}})^{ij} = \widetilde{Z}^{ij}_P(g_0^2, aM) P^{ij} . \tag{8.47b}$$

8.3.3 Improvement of the quark mass

We finally consider the renormalization of the quark masses, and recall that the unimproved expression in the *mass-independent* scheme reads

$$m^i_{\mathrm{R}} = Z_m(g_0^2, a\mu) \left[m_{q,i} + \left(r_m(g_0^2) - 1 \right) \mathrm{Tr}\,[M] \,/ N_f \right] . \tag{3.66}$$

In the *mass-dependent* scheme, the dimensionful term that comes with $\mathrm{Tr}[M]$ can—for instance—be attributed to the critical mass, which in that case becomes mass-dependent:

$$m^M_{\mathrm{crit}}(g_0^2, \mathrm{Tr}\,[M]) = m_{\mathrm{crit}}(g_0^2) - \left(r_m(g_0^2) - 1 \right) \mathrm{Tr}\,[M] \,/ N_f . \tag{8.48}$$

The mass renormalization (3.66) may then be written as

$$m^i_{\mathrm{R}} = Z_m(g_0^2, a\mu) \left[m_{0,i} - m^M_{\mathrm{crit}}(g_0^2, \mathrm{Tr}\,[M]) \right] . \tag{8.49}$$

The improved version in the mass-independent scheme is obtained from (3.66) by the replacements $g_0^2 \to \widetilde{g}_0^2$ and $m_{q,i} \to \widetilde{m}_{q,i}$, and given in (8.25). The improvement terms may, like in the previously discussed case of the coupling, be absorbed in the renormalization constants (see app. E.4 for an explicit derivation):

$$m^i_{\mathrm{R}} = \widetilde{Z}_m(g_0^2, a\mathrm{Tr}\,[M], a\mu) \left(1 + a b_m(g_0^2) m_{q,i} \right) \left[m_{0,i} - \widetilde{m}^M_{\mathrm{crit}}(g_0^2, \mathrm{Tr}\,[M]) \right] . \tag{8.50a}$$

The above form might be useful in the case of *light quarks*, if the $b_m(g_0^2)$ term can be handled properly. Its advantage is that it preserves the flavor-independence of the renormalization constants, as well as the mere dependence of those on the trace of the quark masses, $\mathrm{Tr}\,[M]$. Alternatively, in particular with $i = c$ in mind, one may absorb the $b_m(g_0^2)$ term as well (cf. (8.46b)),

$$m_{\mathrm{R}}^{i} = \widetilde{Z}_m^{i}(g_0^2, aM, a\mu)\left[m_{0,i} - \widetilde{m}_{\mathrm{crit}}^{M}(g_0^2, \mathrm{Tr}\,[M])\right] , \tag{8.50b}$$

which however goes along with a loss of the aforementioned features. Note that the mass-dependent (improved) critical mass in (8.49, 8.50) reduces to the ordinary critical mass in the case of massless quarks, (cf. (8.48) and app. E.4):

$$\widetilde{m}_{\mathrm{crit}}^{M}(g_0^2, 0) = m_{\mathrm{crit}}^{M}(g_0^2, 0) = m_{\mathrm{crit}}(g_0^2) \ . \tag{8.51}$$

With the bilinear currents in the mass-dependent scheme at hand, see sec. 8.3.2, the renormalization of the quark mass may also be formulated by means of the PCAC mass (cf. (5.55)),

$$m_{\mathrm{R}}^{ij} = \frac{\widetilde{Z}_A(g_0^2, a\mathrm{Tr}\,[M])}{\widetilde{Z}_P(g_0^2, a\mathrm{Tr}\,[M], a\mu)} \left[1 + a\cdot\left(\widehat{b}_A(g_0^2) - \widehat{b}_P(g_0^2)\right) m_{\mathrm{PCAC,I}}^{ij}\right] m_{\mathrm{PCAC,I}}^{ij}(x_0) + O(a^2) \ , \tag{8.52}$$

where the replacement (8.34) has been exercised[10] to eliminate the bare masses. The renormalized strange quark mass, for instance, may then be obtained from the linear combination

$$\begin{aligned} m_{\mathrm{R}}^{s} &= m_{\mathrm{R}}^{us} - \frac{1}{2} m_{\mathrm{R}}^{ud} \\ &= \frac{\widetilde{Z}_A(g_0^2, a\mathrm{Tr}\,[M])}{\widetilde{Z}_P(g_0^2, a\mathrm{Tr}\,[M], a\mu)} \Bigg\{ \quad \left[1 + a\cdot\left(\widehat{b}_A(g_0^2) - \widehat{b}_P(g_0^2)\right) m_{\mathrm{PCAC,I}}^{us}\right] m_{\mathrm{PCAC,I}}^{us} \\ &\qquad - \frac{1}{2}\left[1 + a\cdot\left(\widehat{b}_A(g_0^2) - \widehat{b}_P(g_0^2)\right) m_{\mathrm{PCAC,I}}^{ud}\right] m_{\mathrm{PCAC,I}}^{ud} \Bigg\} + O(a^2) \ , \end{aligned} \tag{8.53}$$

where the up and down quark are assumed to be degenerate.

---

[10] The notation $\widehat{b}$ of the improvement coefficients indicates that these differ from the ones which multiply the bare subtracted masses, cf. (8.46b, 8.47b) and footnote 5. Moreover, note that the difference between the use of m_{PCAC}^{ij} and $m_{\mathrm{PCAC,I}}^{ij}$ in the improvement terms merely affects the order $O(a^2)$.

8.3.4 Improvement coefficients in the mass-dependent scheme

So far, we have considered only the renormalization constants in the mass-dependent scheme. This is in accordance with the fact that we have always examined renormalized and improved quantities up to order $O(a)$ only. As announced in sec. 3.7, the improvement coefficients c_{sw}, c_A, as well as e.g. b_g, b_m, b_A, b_P, do in the massive $N_f = 4^*$ theory in fact depend on the modified bare coupling $\tilde{g}_0^2 = (1 + ab_g(g_0^2)\mathrm{Tr}[M]/N_f)\, g_0^2$ instead of g_0^2, which has been disregarded as the difference affects merely the order $O(a^2)$. For the improvement coefficient c_A as it appears in (8.46), for instance, one finds

$$\underbrace{\tilde{c}_A(g_0^2, aM)}_{\hat{=}\; c_A(\tilde{g}_0^2)} = \underbrace{\tilde{c}_A(g_0^2, 0)}_{\hat{=}\; c_A(g_0^2)} + a\sum_i m_{q,i}\, \frac{\partial \tilde{c}_A}{\partial(am_{q,i})}(g_0^2, 0) + O(a^2)\ , \tag{8.54}$$

where the dependence on the bare coupling and the masses has been separated in the notation. In the presence of a massive charm quark, the size of the $O(a^2)$ mass effects may be large, very similar to the $O(a)$ mass effects considered in the framework of renormalization before. Hence, although strictly speaking not necessary for $O(a)$ improvement, it is a self-evident approach to avoid the explicit introduction of large charm mass cutoff effects to the order $O(a^2)$ that would go along with the use of $c_A(g_0^2, 0)$, and to treat also the improvement coefficients in a *mass-dependent* scheme. The same argumentation holds for the only remaining improvement coefficient in the action,

$$\tilde{c}_{\mathrm{sw}} \equiv \tilde{c}_{\mathrm{sw}}(g_0^2, aM)\ , \tag{8.55}$$

which we consider at finite masses as well. Hence, in summary, all the renormalization constants and the improvement coefficients become dependent on the quark masses aM in the mass-dependent scheme:

$$Z \to \tilde{Z}(g_0^2, a\mathrm{Tr}\,[M]\,, (a\mu)) \tag{8.56a}$$

$$c \to \tilde{c}(g_0^2, a\mathrm{Tr}\,[M])\ . \tag{8.56b}$$

Note that the above tilde notation will be dropped in the following.

8.4 Determination of improvement coefficients in $N_f = 3+1$

The mass-dependent scheme in $N_f = 4^*$ described in the previous section schedules the determination of the renormalization constants and improvement coefficients as functions of *all* the quark masses, i.e.

$$aM = \mathrm{diag}(am_{q,u}, am_{q,d}, am_{q,s}, am_{q,c}) \tag{8.57}$$

in (8.56). Recall, however, that the main objective of the mass-dependent scheme was to avoid the expansion in the large *charm mass* $am_{q,c}$. In order to simplify things, one might consider intermediate schemes in which some of the lighter masses are set to zero again. For instance, the *massive charm scheme*, where the renormalization constants and improvement coefficients depend on the charm mass only:

$$aM \to \mathrm{diag}(0, 0, 0, am_{q,c}) \ . \tag{8.58}$$

Another option is the *massive charm and strange scheme*, where

$$aM \to \mathrm{diag}(0, 0, am_{q,s}, am_{q,c}) \ . \tag{8.59}$$

Concerning the renormalization constants $\widetilde{Z}(g_0^2, aM, (a\mu))$, these choices are equivalent to neglecting all the improvement terms with $am_{q,u}, am_{q,d}$, and, in the case of (8.58), $am_{q,s}$. Consequently, $O(a)$ improvement would not be fulfilled strictly anymore, although the effects would be rather small. Of course, they could also be treated separately (e.g. in a perturbative manner).

In the case of the improvement coefficients, where (as discussed before) mass effects concern the overall order $O(a^2)$ only, we decide to impose the improvement conditions in $N_f = 3+1$, where the up, down and strange quark are mass-degenerate light quarks,

$$aM \to \mathrm{diag}(am_{q,l}, am_{q,l}, am_{q,l}, am_{q,c}) \ . \tag{8.60}$$

In other words, restricting ourselves to c_{sw} in the following, we will determine

$$c_{\mathrm{sw}} = c_{\mathrm{sw}}(g_0^2, am_{q,l}, am_{q,c}) \ . \tag{8.61}$$

In comparison to (8.57), which we understand to correspond to $N_f = 2+1+1$ under neglect of isospin breaking, the number of mass parameters is reduced by one, which—as will become clear in sec. 9—will simplify the determination of $c_{\rm sw}$ significantly.

In principle, it would be desirable to vary all the bare parameters $g_0^2, am_{q,l}, am_{q,c}$ and find an interpolating formula for $c_{\rm sw}$ in this 3-dimensional space. However, as this is not practicable, instead we fix the light and charm mass to (approximately) physical values[11] indicated by an asterisk,

$$am_{q,l} \to am_{q,l}^{\star} \;, \qquad am_{q,c} \to am_{q,c}^{\star} \;, \tag{8.62}$$

and interpolate only in g_0^2:

$$c_{\rm sw} = c_{\rm sw}(g_0^2, am_{q,l}^{\star}, am_{q,c}^{\star}) \;. \tag{8.63}$$

With these two approximations at hand,

$$c_{\rm sw}(g_0^2, aM) \xrightarrow{N_f=3+1} c_{\rm sw}(g_0^2, am_{q,l}, am_{q,c}) \xrightarrow[\text{masses}]{\text{fix}} c_{\rm sw}(g_0^2, am_{q,l}^{\star}, am_{q,c}^{\star}) \;, \tag{8.64}$$

the existing $O(a^2)$ cutoff effects are modified explicitly by terms proportional to

$$a^2(m_{q,u} - m_{q,l}^{\star}) \;, \quad a^2(m_{q,d} - m_{q,l}^{\star}) \;, \quad a^2(m_{q,s} - m_{q,l}^{\star}) \;, \qquad a^2(m_{q,c} - m_{q,c}^{\star}) \;. \tag{8.65}$$

Obviously, the relevant point here is to ensure that the charm mass is fixed accurately, so that the difference $m_{q,c} - m_{q,c}^{\star}$ does not become too large in (large volume) simulations at or close to the physical point.

---

[11] The exact meaning of this statement will be clarified in sec. 9.1.

9 Determination of c_{sw} for $N_f = 3+1$ massive Wilson fermions

The improvement coefficient c_{sw} has been determined as a function of the bare coupling g_0^2 for various numbers of massless, dynamical quark flavors and different actions before [16–22]. The improvement condition by which it is defined is set up in the framework of the Schrödinger functional (cf. sec. 5), and has essentially been deployed in the same way[1] it was originally proposed in [46]. The associated Sheikholeslami–Wohlert term has found successful application in the improvement of simulations with *massless* quarks (e.g. [136] with $N_f = 4$) or *light* quarks (e.g. [38, 137] with $N_f = 3$). As explained in the previous section, the *mass-dependent* improvement terms of order $O(am_{q,l})$ can in these cases be treated perturbatively or even be neglected due to the smallness of $am_{q,l}$, cf. (8.38a).

The present work, designed to be applied in simulations with a massive charm quark, differs from the previous cases in two fundamental aspects. First of all, the determination is performed with *massive*, dynamical quarks for the reasons explained in sec. 8. Second, the size of the SF box is kept constant. In the aforementioned previous simulations (see app. F for a review), in contrast, it was the lattice resolution that was held fixed. These two new aspects go hand in hand, and are taken into account together by a *line of constant physics (LCP)*. While this approach has substantial advantages to be discussed further below, its technical implementation is considerably more involved. Numerical results will be obtained under use of the tree-level improved Lüscher–Weisz gauge action,

$$S_G = S_G^{\mathrm{LW}} \,. \tag{9.1}$$

In the next section, sec. 9.1, we specify the line of constant physics, which in particular fixes the quark masses $am^{\star}_{q,l}$ and $am^{\star}_{q,c}$ (cf. (8.62)). Its definition is completed in sec. 9.2 by numerical input from the $N_f = 2$ theory, and the strategy to tune the bare

---

[1] This holds up to small deviations of technical nature.

parameters to it is described in sec. 9.3. Subsequently, in sec. 9.4, the improvement condition which determines $c_{\mathrm{sw}}(g_0^2)$ along this LCP is discussed. Afterwards, we summarize the procedure in sec. 9.5 and specify some technical details of our numerical simulations. Finally, in sec. 9.6, we present results. A short account of the subject has been published in [138] already.

9.1 Line of constant physics

In sec. 8.4, it has been outlined that c_{sw} will be determined in $N_f = 3+1$ at fixed quark masses,

$$c_{\mathrm{sw}} = c_{\mathrm{sw}}(g_0^2, am_{q,l}^\star, am_{q,c}^\star) \ . \tag{8.63}$$

This will be given a specific meaning in the following, by relating the bare masses to the relevant physical scales. In the $N_f = 3+1$ theory, and in a finite volume like in the case of the Schrödinger functional, these are the temporal and spatial extents T, L of the lattice and the RGI masses M_l and M_c of the light and charm quark, respectively. *Until further notice, we restrict the discussion to lattices with $T = L$ only.* For reasons which will become clear in sec. 9.4 (when the improvement condition for c_{sw} is discussed), we will—somewhat unconventionally—use the notation T instead of L in the upcoming discussion.

9.1.1 Fixing of the physical scales

On the lattice, the bare parameters $g_0^2, am_{q,l}$ and $am_{q,c}$ may be considered functions of the three continuum scales T, M_l and M_c. These functions are governed by the theory (N_f), the lattice resolution T/a, and renormalization constants Z. Under the reasonable assumption that the mapping

$$(T \otimes M_l \otimes M_c) \xrightarrow{N_f,\ T/a;\ Z} \left(g_0^2 \otimes am_{q,l} \otimes am_{q,c}\right) \tag{9.2}$$

is injective[2], the scales uniquely determine the bare parameters. Hence, the light and the charm bare subtracted masses, $am_{q,l}$ and $am_{q,c}$, may be assigned values $am_{q,l}^\star$ and

---

[2] This only needs to be ensured in a relevant area in the vicinity of the physical masses.

$am^{\star}_{q,c}$ that correspond to (approximately) physical RGI masses $M^{\star}_{l}$ and $M^{\star}_{c}$:

$$\left.\begin{array}{l} T \\ TM^{\star}_{l} \\ TM^{\star}_{c} \end{array}\right\} \xrightarrow{T/a} \left\{\begin{array}{l} g_0^2 \\ am^{\star}_{q,l} \\ am^{\star}_{q,c} \quad . \end{array}\right. \tag{9.3}$$

This is the exact meaning of (8.62) in sec. 8.4. In order to fix M_l and M_c in physical units, however, the setting of the scale T is required (cf. sec. 3.4).

Fixed lattice resolution approach

In principle, one could choose g_0^2 at one's discretion for some fixed T/a, and adjust κ_l and κ_c such that—under knowledge of the scale T (as well as renormalization constants)—the dimensionless products TM_l and TM_c assume the aspired values $TM^{\star}_{l}$ and $TM^{\star}_{c}$:

$$\left.\begin{array}{l} TM^{\star}_{l} \\ TM^{\star}_{c} \end{array}\right\} \xrightarrow{T/a,\ g_0^2} \left\{\begin{array}{l} am^{\star}_{q,l} \\ am^{\star}_{q,c} \quad . \end{array}\right. \tag{9.4}$$

Apart from issues concerning the practical implementation of this *fixed lattice resolution approach* at finite masses, which we will get back to below, it is conceptually advantageous for the determination of c_{sw} to fix *all* the relevant physical scales, including T. The reason[3] for this is related to the fact that c_{sw} (just like any other improvement coefficient) is ambiguous to the order $O(a)$, even at fixed bare parameters[4]. To explain this, first note that so far, we have suppressed the explicit dependence on the lattice resolution a/T. Including it, we get

$$\begin{aligned} c_{\mathrm{sw}} &\equiv c_{\mathrm{sw}}(g_0^2, a/T, am_{q,l}\,, am_{q,c}) \\ &\equiv c^{M}_{\mathrm{sw}}(g_0^2, a/T, aM_l\,, aM_c)\ , \end{aligned} \tag{9.5}$$

where in the second line, we used the physical mass scales as arguments, which is reflected in the superscript M of the new function name. In addition to what is shown in (9.5), c_{sw} depends on the chosen improvement condition and its kinematic parameters (cf. sec. 9.4). In perturbation theory, however, where the improvement coefficient c_{sw}

---

[3] The upcoming discussion is based on [92].

[4] See (8.54) for the impact of different masses.

is expanded in g_0^2,

$$c_{\mathrm{sw}}^{M}(g_0^2, a/T, aM_l, aM_c) = c_{\mathrm{sw}}^{M(0)}(a/T, aM_l, aM_c) + g_0^2\, c_{\mathrm{sw}}^{M(1)}(a/T, aM_l, aM_c) + O(g_0^4)\ , \tag{9.6}$$

the coefficients $c_{\mathrm{sw}}^{M(0)}(0,0,0)$ and $c_{\mathrm{sw}}^{M(1)}(0,0,0)$ are unique. Consequently, if (and only if) *all* the scales and the improvement condition are kept constant, c_{sw} is unique in the limit $g_0^2 \to 0$ (or $a \to 0$). In that case, as opposed to the fixed lattice resolution approach, one avoids the explicit introduction of $O(a)$ cutoff effects in c_{sw} already at tree level in perturbation theory, which may possibly get large in the non-perturbative regime ($g_0^2 \gg 0$).

Line of constant physics approach

For these reasons, we introduce the conditions

$$T = T^\star \tag{9.7a}$$

$$TM_l = T^\star M_l^\star \tag{9.7b}$$

$$TM_c = T^\star M_c^\star \tag{9.7c}$$

to fix *all* the physical scales to particular values $T^\star, M_l^\star$ and $M_c^\star$, which will be specified later. The above equations define the *line of constant physics (LCP)* at which the improvement condition for c_{sw} will be imposed. This terminology refers to the trajectory in bare parameter space defined by the scales, on which the corresponding physics remains unchanged as the lattice resolution is varied. A direct implication of (9.7a) is that the lattice spacing a changes with[5] $T^\star/a$. As $T^\star/a$ is bound to integer values, and is even further constrained due to practical considerations (see sec. 9.5.1), the lattice spacing may only assume specific, discrete values, see fig. 9.1. The same holds for the bare parameters:

$$\left.\begin{array}{l} T^\star \\ T^\star M_l^\star \\ T^\star M_c^\star \end{array}\right\} \xrightarrow{T^\star/a} \left\{\begin{array}{l} g_0^{2\star} \\ am_{q,l}^\star \\ am_{q,c}^\star \end{array}\right. . \tag{9.8}$$

---

[5] Here and in the following, if the volume is fixed by the line of constant physics, this will be emphasized by the notation $T^\star/a$ for the lattice resolution, instead of T/a.

[6] I would like to thank Patrick Fritzsch, who provided me with templates for this figure.

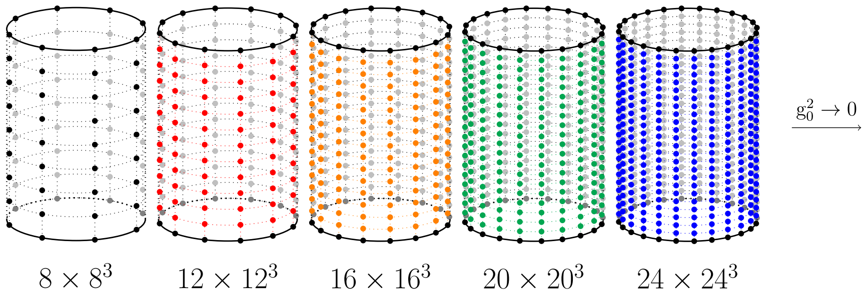


Figure 9.1: The LCP (9.7) is imposed at fixed physical volume $T = T^\star \sim 0.8$ fm with $L = T$ and vanishing boundary conditions indicated by the transparent sides[6].

In particular, the determination of c_{sw} in terms of a LCP can only take place at specific, discrete values $g_0^{2\star} \equiv g_0^{2\star}(T^\star/a)$ of g_0^2. This may be considered an additional restriction on top of those described in sec. 8.4, by which (8.64) becomes

$$c_{\text{sw}}(g_0^2, aM) \xrightarrow{N_f=3+1} c_{\text{sw}}(g_0^2, am_{q,l}, am_{q,c}) \xrightarrow{\text{LCP}} c_{\text{sw}}(g_0^{2\star}, am_{q,l}^\star, am_{q,c}^\star) \ . \tag{9.9}$$

Note, however, that irrespective of the approach, practicability admits simulations at only a few discrete values of g_0^2 and thus requires a subsequent interpolation of c_{sw} in g_0^2 anyway. The new aspect brought in by the LCP is merely that these values $g_0^{2\star}$ are predetermined by the physical scales, as well as the feasible lattice resolutions $T^\star/a$. This aspect should be taken into account in the choice of $T^\star$ though, to ensure that a reasonable range of g_0^2 is covered for the interpolation of c_{sw}. It is the only scale where one has a substantial freedom to choose. In contrast, the RGI masses, especially $M_c^\star$, are bound by the requirement to be close to the physical point in order to minimize the explicitly introduced cutoff effects in (8.65).

We still need to assign values to the RGI masses, $M_l^\star$ and $M_c^\star$. While the choice for $M_c^\star$ is obvious from the discussion in sec. 8.4,

$$M_c^\star = M_{c,\text{phys}} \ , \tag{9.10a}$$

we opt for

$$M_l^\star = M_{s,\text{phys}}/3 \ , \tag{9.10b}$$

in the case of the light RGI mass, which to a very good approximation is the average of the physical up, down and strange RGI masses. Hence, the trace of the mass matrix will also be close to its correct physical value, which may be of particular importance in large scale simulations with massive fermions (see e.g. [38]).

9.1.2 Implicit definition of the LCP in $N_f = 2$

There are two qualifications to be added to the definition of the LCP. First, the quantities $M_{s,\text{phys}}$ and $M_{c,\text{phys}}$ are unknown in the $N_f = 3+1$ theory, which is why we have to make use of the estimates from the $N_f = 2$ theory. Hence, (9.10) gets replaced by

$$M_l^\star = M_{s,\text{phys}}^{(N_f=2)}/3 \tag{9.11a}$$

$$M_c^\star = M_{c,\text{phys}}^{(N_f=2)} \, . \tag{9.11b}$$

Second, the direct realization of the line of constant physics in terms of the quantities $T^\star$ (not yet specified), $M_l^\star$ (cf. (9.11a)) and $M_c^\star$ (cf. (9.11b)) on the lattice, see (9.8), requires the knowledge of scale setting as well as the mass renormalization constants. Since again none of this is available in $N_f = 3+1$, we need to resort to $N_f = 2$ results here as well, which leads to the bare parameters[7] $g_0^{2\star}, \kappa_l^\star, \kappa_c^\star$ (within $N_f = 2$). For the implementation of the LCP in $N_f = 3+1$, it is convenient to define three independent lattice quantities Φ_1, Φ_2, Φ_3, that are easily accessible functions of the bare parameters only[8]:

$$\Phi_i \equiv \Phi_i(g_0^2, \kappa_l, \kappa_c) \, , \qquad i = 1,2,3 \, . \tag{9.12}$$

In particular, their computation should require no external input like e.g. renormalization constants[9]. In the $N_f = 2$ theory, one may translate the LCP to a specific point $\Phi_1^\star, \Phi_2^\star, \Phi_3^\star$ in the image of the functions Φ_i. To this end, one evaluates $\Phi_i(g_0^{2\star}, \kappa_l^\star, \kappa_c^\star)$ at different lattice resolutions $T^\star/a$ and extrapolates to the continuum. The obtained values $\Phi_i^\star$ directly correspond to $T^\star, M_l^\star$ and $M_c^\star$ (within $N_f = 2$), and do in particular not depend on the specifics of the lattice discretization due to universality. The implementation of the described procedure is subject of sec. 9.2.

In $N_f = 3+1$, the LCP conditions (9.7) then get replaced by

$$\Phi_1 = \Phi_1^\star \, , \quad \Phi_2 = \Phi_2^\star \, , \quad \Phi_3 = \Phi_3^\star \, , \tag{9.13}$$

---

[7] In the present context, the hopping parameters are frequently used instead of the bare subtracted masses.

[8] Note that the dependencies in (9.12) hold for both $N_f = 2$ and $N_f = 3+1$, although in the former case the masses enter only via valence quark masses.

[9] Recall that these are not available in $N_f = 3+1$.

$$\begin{tikzcd}
T^\star \arrow[r, "N_f=2"] \arrow[dr, dotted] \arrow[ddr, dotted] & g_0^{2\star}(a/T^\star) \arrow[r, "N_f=2"] \arrow[dr, dotted] \arrow[ddr, dotted] & \Phi_1(a/T^\star) \arrow[r, squiggly, "\text{cont. lim.}"] & \Phi_1^\star \arrow[r, "N_f=3+1"] \arrow[dr, dotted] \arrow[ddr, dotted] & g_0^{2\star}(a/T^\star) \\
M_l^\star \arrow[r] & \kappa_l^\star(a/T^\star) \arrow[r] \arrow[dr, dashed] & \Phi_2(a/T^\star) \arrow[r, squiggly] & \Phi_2^\star \arrow[r] \arrow[ur, dotted] \arrow[dr, dotted] & \kappa_l^\star(a/T^\star) \\
M_c^\star \arrow[r] & \kappa_c^\star(a/T^\star) \arrow[r] & \Phi_3(a/T^\star) \arrow[r, squiggly] & \Phi_3^\star \arrow[r] \arrow[ur, dotted] \arrow[uur, dotted] & \kappa_c^\star(a/T^\star)
\end{tikzcd}$$

Figure 9.2: Illustration of the strategy to establish an implicit LCP in $N_f = 3+1$. The first three steps leading to $\Phi_i^\star$ are conducted in $N_f = 2$, the last one in $N_f = 3+1$. The solid arrows show the main dependencies of the parameters, whereas the dotted arrows denote unavoidable minor dependencies. The dashed arrow is a weak dependence minimized by the chosen definition of Φ_3 (see (9.23b) and fig. G.2). The wiggly arrows symbolize continuum limits.

which implicitly defines the bare parameters (within $N_f = 3+1$), depending on $T^\star/a$. The whole strategy and the dependencies of the involved quantities are illustrated in fig. 9.2. We emphasize again that the implicit definition of the LCP avoids scale setting and mass renormalization in $N_f = 3+1$. The price to pay for this is that the resulting LCP will not exactly correspond to the values of the (physical) input quantities $T^\star$, $M_l^\star$ and $M_c^\star$, as the translation to $\Phi_i^\star$ was conducted in $N_f = 2$. Nevertheless, we expect this effect, which—as was explained before (cf. sec. 8.4 and (8.65))—concerns c_{sw} only at the order $O(a)$, to be reasonably small. Having described the inevitable detour by means of the $N_f = 2$ theory, we are now ready to understand why the fixed lattice resolution approach, in addition to its conceptual drawbacks, is also highly unfeasible, as has been alleged above. At fixed T/a, one would have to deal with different values of T while keeping $M_l^\star$ and $M_c^\star$ at their physical values. The quite elaborate (cf. sec. 9.2) $N_f = 2$ translation of $T^\star$, $M_l^\star$ and $M_c^\star$ to $\Phi_i^\star$ would have to be performed once for each T (or equivalently g_0^2), which is certainly very cumbersome to implement.

The missing ingredients to complete the LCP are the definitions of Φ_i and $T^\star$, which we are going to address in the remainder of this section. For the selection of suitable Φ_i, the following criteria should be taken into account:

C1. They should be finite volume quantities defined by means of the *Schrödinger functional,* since this is computationally cheap (cf. sec. 5.1) and convenient, in particular as the SF is also the framework in which the improvement condition on c_{sw} will be formulated.

As such, they will (like $\Phi_i^\star$ in the continuum) generally depend on the details of the Schrödinger functional, in particular the boundary conditions ϕ, ϕ' for the gauge fields (5.7) and the fermionic boundary angle θ (5.21). We here anticipate our choices. These are vanishing boundary conditions,

$$\phi = \phi' = 0\,, \tag{9.14}$$

while

$$\theta = \begin{cases} 0.5 & \text{for } N_f = 2 \\ 0 & \text{for } N_f = 3+1 \end{cases} \tag{9.15}$$

is chosen differently for the different theories. This is because we will reuse existing $N_f = 2$ ensembles that correspond to (9.14) and (9.15), see sec. 9.2, while we employ a vanishing fermionic boundary angle for the determination of c_{sw}, as before in [16, 17, 21, 22]. The alteration of the LCP by the change in θ is expected to be negligible compared to the aforementioned effects on the LCP that result from the involvement of $N_f = 2$ results in its implicit definition.

C2. Since on the lattice the Φ_i are determined by the bare parameters of the theory, see (9.12), they should preferably be chosen such that they are *independent of one another*, with each of them depending strongly on one of the bare parameters but only weakly on the others, see fig. 9.2. The better this requirement is fulfilled, the simpler will be the tuning of the bare parameters to the LCP in $N_f = 3+1$.

C3. The quantities should have *good statistical properties* (good signal-to-noise ratio, small autocorrelations, ...) and mild cutoff effects.

Keeping these aspects in mind, we now turn to the pending definitions of Φ_i and the specification of $T^\star$.

- Φ_1

As the first quantity to implicitly fix T, M_l and M_c, we choose the gradient flow coupling introduced in sec. 6.1:

$$\Phi_1 = \bar{g}^2_{\mathrm{GF}}\,. \tag{9.16}$$

As explained there, and indicated in (6.7), it mainly depends on T and is expected to be only weakly dependent on M_l and M_c (cf. criterion C2) due to the findings in [39, 105, 109]. Moreover, it has proven to provide excellent statistical precision, while its cutoff effects are modest [105] (cf. criterion C3). Note that in accordance with the discussion in sec. 6.1, $\overline{g}^2_{\mathrm{GF}}$ is only weakly sensitive to θ and therefore expected not to be significantly affected by the ambiguity (9.15). To complete its definition, several parameters need to be specified. The smoothing fraction c which links the flow time t to the size $T = L$ of the box, see (6.6) and (6.7), is taken to be

$$c = 0.3\,, \tag{9.17}$$

which puts the emphasis slightly more on statistical precision than small cutoff effects (cf. sec. 6.1). Next, the discretized action S_{flow} that governs the evolution of the gauge fields in flow time, see (6.9), needs to be chosen. We stress again that this choice is independent from the gauge action S_G that is used in the Boltzmann factor of the path integral (4.1), as the links $U_\mu(x)$ of the gauge configurations sampled with S_G merely provide the starting point $V_\mu(x,0)$ of the flow. We use the Wilson flow,

$$S_{\mathrm{flow}} = S_G^{\mathrm{W}}\,, \tag{9.18}$$

for which the the normalization $\mathcal{N}^{-1}$ (cf. (6.7)) in the Schrödinger functional is known[10] in combination with vanishing boundary conditions from [105]. Finally, the discretization of the field strength tensor $G_{\mu\nu}$ which enters $\overline{g}^2_{\mathrm{GF}}$ via the energy density (6.4) must be specified. We use the symmetric clover definition (6.12),

$$E = E^{(\mathrm{clov})}\,. \tag{9.19}$$

- **Φ_2 and Φ_3**

We define *effective meson masses (EMM)* by

$$\Gamma^{ij} \equiv \Gamma^{ij}_A = -\tilde{\partial}_0 \log\left(f^{ij}_{A,\mathrm{I}}(x_0)\right)\big|_{x_0=T/2}\,, \quad i,j \in \{l,c\}\,, \tag{9.20}$$

---

[10] In the meantime, the normalization has also been calculated for other options, see for instance [110, 111, 139], but neither this nor the implementation of other flow actions than (9.18) in the utilized openQCD code had been available at the time this work was started.

where $f_{A,\mathrm{I}}^{ij}$ is the improved Schrödinger functional correlation function of the axial current given in (5.53). Its spectral decomposition (cf. (3.51)) reads

$$\begin{aligned} f_{A,\mathrm{I}}^{ij}(x_0) &\overset{(5.46)}{\sim} \left\langle A_0^{ij}(x)\ \mathcal{O}^{ji} \right\rangle \\ &\sim \left\langle 0|A_0^{ij}|\mathrm{PS}\right\rangle \left\langle \mathrm{PS}|\mathcal{O}^{ji}|i_0\right\rangle\ e^{-x_0 m_{\mathrm{PS}}^{ij}} \cdot \left(1 + O(e^{-x_0 \Delta E_{PS}^{ij}})\right) + O(a)\ , \end{aligned} \tag{9.21}$$

where PS stands for the pseudoscalar ground state of valence quark flavors ij and i_0 denotes the SF boundary state with vacuum quantum numbers, see e.g. [140, 141]. Employing (9.21) in (9.20), one finds

$$\Gamma_{ij} = m_{\mathrm{PS}}^{ij} + O(e^{-T/2\cdot\Delta E_{PS}^{ij}}) + O(a)\ . \tag{9.22}$$

If the volume T is sufficiently large for contributions ΔE_{PS}^{ij} from excited states to be negligible, the EMM are close to (infinite volume) masses of pseudoscalar mesons with valence (anti-)quark flavors i and j. As such, they may well serve to fix M_l, M_c on the lattice. In order to obtain quantities that depend primarily on either M_l or M_c, but only weakly on the respective other RGI mass (cf. criterion C2), we choose the combinations

$$\Phi_2 = T\cdot\Gamma^{su} \tag{9.23a}$$

$$\Phi_3 = T\cdot\left(\Gamma^{sc} - \frac{1}{2}\Gamma^{su}\right)\ . \tag{9.23b}$$

The subtraction of $\frac{1}{2}\Phi_2$ in (9.23b) is aimed at cancelling the leading influence of M_l in Φ_3 (cf. fig. 9.2). The effectiveness of this measure is studied in $N_f = 2$, see app. G.3. Note that for dimensional reasons, however, the explicit appearance of the scale T (which is already captured by Φ_1) in Φ_2 and Φ_3 is unavoidable. The notation Γ^{su} is employed to emphasize that the axial current in (9.20) is defined as the *non-singlet* axial current of two distinct (but otherwise arbitrarily chosen from u, d, s) flavors with equal, light mass.

Note that the above argumentation, in particular (9.21) and (9.22), holds equally well for Γ_P^{ij} with the pseudoscalar correlation function f_P. As both operators have the same pseudoscalar quantum numbers, the corresponding effective meson masses approach the same meson masses m_{PS}^{ij} in the large volume $\left(O(e^{-T/2\cdot\Delta E_{PS}^{ij}}) \xrightarrow{T\to\infty} 0\right)$ and continuum $\left(O(a) \xrightarrow{a\to 0} 0\right)$ limit. The advantage of Γ_A^{ij} is that in the chiral limit,

$m_{\mathrm{PS}}^{ij} \to 0$, the excited state contributions vanish also for finite T:

$$\lim_{a\to 0} \Gamma_A^{ij}|_{\text{chiral limit}} = 0 \tag{9.24a}$$

$$\lim_{a\to 0} \Gamma_P^{ij}|_{\text{chiral limit}} \neq 0\ . \tag{9.24b}$$

The underlying reason for this is the PCAC relation (3.79), which renders Γ_A^{ij} proportional to the quark masses[11],

$$\Gamma_A^{ij} = \frac{\partial_0 \left\langle A_0^{ij}(x)\ \mathcal{O}^{ji} \right\rangle}{\left\langle A_0^{ij}(x)\ \mathcal{O}^{ji} \right\rangle} = \left(m_{0,i} + m_{0,j}\right) \cdot \frac{\left\langle P^{ij}(x)\ \mathcal{O}^{ji} \right\rangle}{\left\langle A_0^{ij}(x)\ \mathcal{O}^{ji} \right\rangle}\ , \tag{9.25}$$

which is not the case for Γ_P^{ij}. Consequently, if the involved quark masses are small as in the case of Φ_2, Γ_A^{ij} is more sensitive to the bare masses than Γ_P^{ij}, which allows to tune Φ_2 with an improved accuracy. This may be seen in app. G.3, where a detailed numerical discussion of the mass dependence of the effective meson masses is given.

There are several choices to be made which have impact on the size of the discretization effects in (9.22). The first one is the use of the *improved* correlation function $f_{A,\mathrm{I}}^{ij}$ already indicated in the above equations. Although this is not a necessity for the determination of c_{sw}, as the difference with respect to the use of the *unimproved* correlation function f_A^{ij} is merely an $O(a^2)$ effect, the knowledge of the 1-loop coefficient $c_A^{(1)}$ (cf. (3.115)) may be applied, while some non-perturbative results on c_A are also available (cf. app. F.2). We will come back to the use of c_A, discussed separately for $N_f = 2$ and $N_f = 3+1$, in sec. 9.2.4 and sec. 9.3.

Second, different discretizations of Γ^{ij} and the involved logarithmic derivative (cf. (9.20)) are possible, covered in detail in app. G.1. We opt for

$$\Gamma^{ij} = \Gamma_{(1)}^{ij} = -\frac{1}{2a} \log\left(\frac{f_{A,\mathrm{I}}^{ij}(x_0+a)}{f_{A,\mathrm{I}}^{ij}(x_0-a)}\right)\Bigg|_{x_0=T/2} \tag{9.26}$$

over $\Gamma_{(2)}^{ij}$ as given in (G.1), to avoid an explicit introduction of cutoff effects, see (G.5) and (G.6).

---

[11] We consider the continuum here.

- $T^\star$

Having specified the LCP in its explicit form (9.7) with the associated mass scales (9.11), and its implicit form (9.13) in terms of the three quantities Φ_i (9.16, 9.23a, 9.23b), the only missing ingredient to complete its definition is the scale $T^\star$.

It is advantageous to establish the LCP in a rather large volume to minimize excited state contributions to the EMM, cf. (9.21). However, if the volume gets too large, the signal strength of the Φ_i gets impaired. Moreover, regarding the determination of Φ_1, the effects of critical slowing down and topology freezing (cf. sec. 6.2.2) possibly become threatening (cf. sec. 7), and the size of the cutoff effects becomes significant [105]. A good compromise is

$$T^\star \sim 0.8 \text{ fm} \,. \tag{9.27}$$

This particular value is chosen because corresponding $N_f = 2$ ensembles (9.27) already exist and may be reused. These will provide the exact definition of $T^\star$, to be described in detail in sec. 9.2. We close this section with the remark that the LCP as defined above replaces the improvement point $M^{ud}(T/2, T/4) \overset{!}{=} 0$ in previous determinations of c_{sw} with massless quarks at fixed lattice resolution, cf. (F.1) in app. F. In the upcoming section, we will determine the values $\Phi_i^\star$ in $N_f = 2$ by which the LCP is characterized.

9.2 Determination of $\Phi^\star$ in $N_f = 2$

The determination of the values $\Phi_i^\star$ has been outlined in sec. 9.1, cf. (9.13) and fig. 9.2. They are defined as the continuum extrapolated values of the quantities Φ_i (cf. (9.16, 9.23a, 9.23b)) in massless $N_f = 2$ with the scales[12] T, M_l, M_c set to $T^\star$ (9.27) and the (partly) physical $N_f = 2$ RGI masses $M_l^\star$ (9.11a) and $M_c^\star$ (9.11b), respectively:

$$\Phi_1^\star = \lim_{a/T\to 0} \left[\, \overline{g}^2_{\text{GF}} \, \right]_{N_f=2,\; T=T^\star} \tag{9.28a}$$

$$\Phi_2^\star = \lim_{a/T\to 0} \left[\, T \cdot \Gamma^{s\bar{s}} \, \right]_{N_f=2,\; T=T^\star,\; M_l=M_l^\star} \tag{9.28b}$$

$$\Phi_3^\star = \lim_{a/T\to 0} \left[\, T \cdot \left(\Gamma^{sc} - \frac{1}{2}\Gamma^{s\bar{s}} \right) \right]_{N_f=2,\; T=T^\star,\; M_l=M_l^\star,\; M_c=M_c^\star} \,. \tag{9.28c}$$

---

[12] We remind the reader that all the Schrödinger functional parameters employed here were specified in sec. 9.1. In particular, the spatial size of the lattice is equal to its temporal size, $T = L$.

#	notation	dynamics	mass
1	u	dynamical	massless
2	d	dynamical	massless
3	s or l	quenched	light
4	c	quenched	charm

#	notation	dynamics	mass
1	u or l	dynamical	light
2	d or l	dynamical	light
3	s or l	dynamical	light
4	c	dynamical	charm

Table 9.1: Basic properties and notation used for the different quark flavors present in the $N_f = 2$ *(left)* and $N_f = 3 + 1$ *(right)* theories. The definition of *light* and *charm* mass in $N_f = 2$ is straight forward, expressed by (9.11a) and (9.11b), respectively. In $N_f = 3 + 1$, however, the exact meaning of this terminology is given by the tuning of the masses to the LCP (9.13), which implies that the light and charm quarks are to be understood as *light- and charmlike* quarks, respectively. Recall that in this context *light* always means one third of the strange mass, cf. (9.11a).

Note that in contrast to $N_f = 3 + 1$, only the (quenched) strange quark corresponds to a light flavor with index l in $N_f = 2$, while the (dynamical) up and down quarks are massless. Accordingly, the EMM with two light valence quarks are labelled with the superscript $s\tilde{s}$ instead of su (cf. (9.23)). The tilde notation here indicates the unaltered consideration of *non-singlet* axial current components. An overview on the notation of the quark flavors in the different theories and their basic properties is given in tab. 9.1. In the following, we present the $N_f = 2$ ensembles which provide the exact definition of $T^\star$. These will be deployed for the determination of $\Phi_i^\star$ subsequently.

9.2.1 Scale setting and ensembles in $N_f = 2$

In [95, 142], a typical hadronic scale L_1 has been implicitly defined through a LCP in $N_f = 2$, formulated in terms of the Schrödinger functional coupling (cf. (5.15)):

$$\bar{g}^2_{\mathrm{SF}}(L_1) = 4.484 \;, \quad L_1 m^{ud}_{\mathrm{PCAC}} = 0 \;. \tag{9.29}$$

Here, m^{ud}_{PCAC} refers to the PCAC mass of the degenerate up and down quark[13], cf. (5.51). In the course of [142], the bare parameters (β, κ) were tuned to the LCP (9.29) for a couple of lattice resolutions, leading to a set of ensembles that correspond to

$$L_1 \quad \text{with} \quad L_1/a \in \{6, 8, 10, 12, 16\} \;. \tag{9.30}$$

---

[13] Note that the PCAC mass without explicit declaration of a time argument always corresponds to the center of the lattice, i.e. $m^{ud}_{\mathrm{PCAC}} \equiv m^{ud}_{\mathrm{PCAC}}(x_0)|_{x_0=T/2}$.

L_2/a	β	κ_{crit}	τ_{meas}	$\bar{g}^2_{\mathrm{SF}}(L_1)$	am^{ud}_{PCAC}
12	5.2638	0.135985	6	4.423(75)	-0.01154(83)
16	5.4689	0.136700	6	4.473(83)	-0.00424(24)
20	5.6190	0.136785	6	4.49(10)	-0.00257(11)
24	5.7580	0.136623	6	4.501(91)	+0.00067(7)
32	5.9631	0.136422	10	4.40(10)	-0.00096(4)

Table 9.2: Properties of the $N_f = 2$ ensembles corresponding to the lattice extent L_2 created in the course of [142]. The numbers are taken from the same reference. τ_{meas} denotes the Monte Carlo time between two consecutive measurements, given in MDU. Note that the observables $\bar{g}^2_{\mathrm{SF}}$ and am^{ud}_{PCAC} both refer to the related ensembles with $L_1 = L_2/2$, cf. (9.30) and (9.31).

At the very same bare parameters, a second set of ensembles with twice the volume, i.e.

$$L_2 = 2L_1 \quad \text{with} \quad L_2/a \in \{12, 16, 20, 24, 32\} \tag{9.31}$$

was created, which corresponds to a scaling step with respect to (9.30). The extent T_1 or T_2 of the SF box in time is the same as the one in space for all of the above ensembles,

$$T_1 = L_1 \quad \text{and} \quad T_2 = L_2\,. \tag{9.32}$$

The main properties of the ensembles in the larger volume are compiled in tab. 9.2. The fact that they were generated with vanishing boundary conditions and under use of $\theta = 0.5$ is reflected in (9.14) and (9.15). Moreover, the Wilson gauge action

$$S_G = S_G^{\mathrm{W}} \tag{9.33}$$

was employed. While this does not directly affect the continuum extrapolated values $\Phi_i^\star$ due to universality, it is relevant for $O(a)$ improvement, which is implemented for the action $S = S_F^{\mathrm{W}} + S_G^{\mathrm{W}}$ by the use of the non-perturbative result for c_{sw} (cf. tab. F.1 in app. F), as well as the perturbative 2-loop and 1-loop expressions for c_t and $\tilde{c}_t$, respectively (cf. (5.37)).

The results of [95] directly relate a physical value to L_1. In this work, the scale in $N_f = 2$ has been set for a range of inverse bare parameters $\beta' \in \{5.2, 5.3, 5.5\}$ under use of the kaon decay constant f_K, based on the principle described in sec. 3.4. The knowledge of the scale f_K can be transferred to the knowledge of the scale L_1 by building the product $L_1 f_K = \frac{L_1}{a} \cdot a f_K$ and extrapolating to the continuum. However, since the lattice resolutions at the LCP are known for slightly different values of β (cf. tab. 9.2), the values L_1/a first need to be interpolated to the β' at which af_K has been computed[14]. In doing so, one finds [95]

$$L_1 f_K = 0.315(8)(2)\ , \tag{9.34}$$

where the first and second error refer to statistical and systematic uncertainties, respectively. Together with $f_{K,\mathrm{phys}} = 155$ MeV (as used in [95]), this allows to identify the scale L_1, which approximates to

$$L_1 \sim 0.4\ \mathrm{fm}\ . \tag{9.35}$$

It is convenient to reuse the existing $N_f = 2$ ensembles. For the reasons described at the end of sec. 9.1, we opt for the ones in the larger volume. The previously stated value $T^\star \sim 0.8$ fm (cf. (9.27)) thus results from

$$T^\star = L_2 = 2L_1\ , \tag{9.36}$$

the exact definition of these scales being provided by the LCP (9.29). *In order to clearly distinguish between the different theories, we use the notation $T^\star$ in the case of $N_f = 3+1$, whereas the same physical scale will be denoted by L_2 in the framework of $N_f = 2$.* Note that apart from the need to produce new configurations, our approach of reusing the ensembles also avoids the interpolation of the bare coupling g_0^2 as a function of the lattice spacing, which due to the restrictions on the lattice resolution would be necessary if a value slightly different from L_2 was chosen.

9.2.2 Fixing the RGI masses on the lattice

While $\Phi_1^\star$ (9.28a) may be computed directly with the help of the $N_f = 2$ ensembles (sec. 9.2.3), the observables $\Phi_2^\star$ (9.28b) and $\Phi_3^\star$ (9.28c) involve valence quarks at fixed

---

[14] Actually, in advance of this step, the hopping parameters were retuned in [95] for the lattice resolutions $L_1/a \in \{6, 8, 12\}$ (cf. tab. 9.2) and $\bar{g}^2_{\mathrm{SF}}$ was computed with higher statistics in order to improve the matching of the LCP and the overall precision. The resulting small deviations of the Schrödinger functional coupling from the aspired LCP were then translated to values of L_1/a that differ slightly from the integer values used in [142] (cf. tab. 9.2).

RGI masses, which we recall here for convenience:

$$M_l^\star = M_{s,\mathrm{phys}}^{(N_f=2)}/3 \tag{9.11a}$$

$$M_c^\star = M_{c,\mathrm{phys}}^{(N_f=2)} \ . \tag{9.11b}$$

After extracting the values of the dimensionless combinations

$$z_i^\star = L_2 M_i^\star \ , \qquad i \in \{l, c\} \ , \tag{9.37}$$

from available results in the literature in the following, we will relate those to the associated hopping parameters $\kappa_i^\star$ of the aforementioned valence quarks, which enter the evaluation of $\Phi_2^\star$ (sec. 9.2.4) and $\Phi_3^\star$ (sec. 9.2.5) in practice.

The *RGI strange quark mass* in $N_f = 2$ is given in units of the kaon decay constant by [95]

$$\frac{M_{s,\mathrm{phys}}^{(N_f=2)}}{f_K} = 0.887(19)(7) \ , \tag{9.38}$$

which together with (9.11a) and (9.37), as well as the values of $L_1 f_K$ (9.34) and $L_1 = L_2/2$ (9.36) leads to

$$z_l^\star = \frac{L_2 M_{s,\mathrm{phys}}^{(N_f=2)}}{3} = \frac{2}{3} \cdot L_1 f_K \cdot \frac{M_{s,\mathrm{phys}}^{(N_f=2)}}{f_K} = 0.186(7) \ . \tag{9.39}$$

The dimensionless *RGI charm quark mass* in $N_f = 2$, in contrast, is derived indirectly according to

$$z_c^\star = z_b^\star \cdot \frac{z_c^\star}{z_b^\star} \ , \tag{9.40}$$

from the dimensionless RGI bottom quark mass, that is readily available from an $N_f = 2$ HQET analysis of the ALPHA Collaboration [143],

$$z_b^\star = L_2 M_{b,\mathrm{phys}}^{(N_f=2)} = 2 L_1 M_{b,\mathrm{phys}}^{(N_f=2)} = 26.50(52) \ . \tag{9.41}$$

The converting ratio

$$\frac{z_b^\star}{z_c^\star} = \frac{M_{b,\mathrm{phys}}}{M_{c,\mathrm{phys}}} \tag{9.42a}$$

can either be obtained from quenched simulations of the ALPHA Collaboration [144, 145],

$$\left.\begin{aligned} r_0 M_{b,\mathrm{phys}}^{(N_f=0)} &= 17.38(28) \\ r_0 M_{c,\mathrm{phys}}^{(N_f=0)} &= 4.055(58) \end{aligned}\right\} \Rightarrow \left.\frac{z_b^\star}{z_c^\star}\right|_{\mathrm{ALPHA}} = 4.29(9)\ , \tag{9.42b}$$

or taken from the HPQCD Collaboration [146],

$$\left.\frac{z_b^\star}{z_c^\star}\right|_{\mathrm{HPQCD}} = 4.51(4)\ . \tag{9.42c}$$

Employing the average value of

$$\frac{z_b^\star}{z_c^\star} = 4.40 \tag{9.42d}$$

together with (9.40) and (9.41) leads to

$$z_c^\star = 6.0\ . \tag{9.43}$$

Note that a similar result on $z_c^\star$ is obtained directly under use of (cf. (9.27))

$$L_2 \approx 4\ \mathrm{GeV}^{-1}\ . \tag{9.44a}$$

from [147], where

$$M_{c,\mathrm{phys}}^{(N_f=2)} = 1.51(4)\ \mathrm{GeV} \tag{9.44b}$$

is given in physical units.

As announced earlier, we now derive the relation of $z_i^\star$ to the associated hopping parameters $\kappa_i^\star(g_0^2)$ for the bare couplings $g_0^2 = 6/\beta$ given in tab. 9.2. $M_i^\star$ is related to the renormalized mass $m_R^{i\star}$ by (E.12) in terms of the scale-dependent but flavor-independent conversion factor h, see (E.11). The renormalized mass $m_R^{i\star}$ itself can be expressed in terms of the bare subtracted quark mass $m_{q,i}^\star$ via the flavor-independent renormalization constant Z_m, see (E.14). Taking into consideration that we work in massless $N_f = 2$ now, such that one may replace $\tilde{g}_0^2 \to g_0^2$ in all renormalization constants (cf. the discussion in app. E), and finite size scaling (cf. (5.18))

$$\mu = 1/L_1 \tag{9.45}$$

associated with the Schrödinger functional is employed, the corresponding equations, (E.11) and (E.14), may be written as

$$h(1/L_1) = M/m_R(1/L_1) \tag{9.46}$$

and

$$M_i^\star = h(1/L_1)\; Z_m(g_0^2, a/L_1)\left[1 + ab_m(g_0^2)m_{q,i}^\star\right] m_{q,i}^\star\,. \tag{9.47}$$

Note that while the particular scale (9.45) is in principle irrelevant in the sense that the product hZ_m is scale-independent, it is self-evident to use L_1 with regard to the evaluation of the single factors, in particular Z_m, as it is directly intertwined with the bare couplings in question via the LCP (9.29). After solving (9.47) for $m_{q,i}^\star$,

$$m_{q,i}^\star = \frac{-1 \pm \sqrt{1 + 4ab_m(g_0^2)\frac{M_i^\star}{h(1/L_1)Z_m(g_0^2,a/L_1)}}}{2ab_m(g_0^2)}\,, \tag{9.48}$$

and using the positive solution $m_{q,i}^\star > 0$, the hopping parameters (cf. (3.70)) become

$$\begin{aligned}\kappa_i^\star &= \left\{\frac{1}{\kappa_{\mathrm{crit}}} + 2am_{q,i}^\star\right\}^{-1} \\ &= \left\{\frac{1}{\kappa_{\mathrm{crit}}(g_0^2)} - \frac{1}{b_m(g_0^2)}\left(1 - \sqrt{1 + 4ab_m(g_0^2)\frac{M_i^\star}{h(1/L_1)Z_m(g_0^2,a/L_1)}}\right)\right\}^{-1}\,. \end{aligned} \tag{9.49}$$

Since there are no direct results for $Z_m(g_0^2, a/L_1)$ available, one uses

$$Z(g_0^2) = \frac{Z_m(g_0^2, a/L_1)\; Z_P(g_0^2, a/L_1)}{Z_A(g_0^2)} \tag{9.50}$$

to replace it by the renormalization constants Z_A, Z_P and Z that appear in connection with the PCAC mass (cf. (E.17)). In terms of the dimensionless RGI masses $z_i^\star$ (9.37),

L_2/a	β	κ_{crit}	b_m	Z	Z_A	Z_P	h	$\kappa_l^\star$	$\kappa_c^\star$
12	5.2638	0.135985	-0.5240	0.9710	0.7754	0.5068	1.308	0.135689	0.125521
16	5.4689	0.136700	-0.6532	1.0530	0.7909	0.5110		0.136496	0.129520
20	5.6190	0.136785	-0.7087	1.0854	0.8010	0.5125		0.136628	0.131342
24	5.7580	0.136623	-0.7422	1.1048	0.8096	0.5186		0.136495	0.132221
32	5.9631	0.136422	-0.7715	1.1223	0.8211	0.5188		0.136329	0.133278

Table 9.3: Improvement coefficient b_m and renormalization constants associated with the LCP in the volume L_1, taken from the references mentioned in the main text, either directly or by application of interpolating formulas given therein. The last two columns show the resulting hopping parameters associated with the RGI masses $M_l^\star$ (9.11a) and $M_c^\star$ (9.11b), respectively, according to (9.51).

the final formula to be applied then reads

$$\kappa_i^\star = \left\{ \frac{1}{\kappa_{\text{crit}}(g_0^2)} - \frac{1}{b_m(g_0^2)} \left(1 - \sqrt{1 + 4 b_m(g_0^2) \frac{Z_P(g_0^2, a/L_1)}{Z_A(g_0^2) \cdot Z(g_0^2)} \frac{a}{L_2} \frac{z_i^\star}{h(1/L_1)}} \right) \right\}^{-1} . \tag{9.51}$$

In $N_f = 2$, non-perturbative results for a range of bare couplings exist for $b_m(g_0^2)$ and $Z(g_0^2)$ [58], as well as $Z_A(g_0^2)$ [148]. Moreover, specifically for the LCP, i.e. the relevant bare couplings $g_0^2 = 6/\beta$, the volume L_1 and the lattice resolutions L_1/a, we know $\kappa_{\text{crit}}(g_0^2)$ from [142], while $Z_P(g_0^2, a/L_1)$ and the conversion factor

$$h(1/L_1) = 1.308(16) \tag{9.52}$$

are given in [95]. This way, from (9.51), we get the hopping parameters $\kappa_l^\star, \kappa_c^\star$ associated with the central values of the RGI masses $z_l^\star, z_c^\star$, given in (9.39) and (9.43), respectively. The results are listed in tab. 9.3.

9.2.3 Determination of $\Phi_1^\star$

We now turn to the determination of $\Phi_1^\star$, whose definition we recall for convenience:

$$\Phi_1^\star = \lim_{a/T \to 0} \left[\bar{g}^2_{\text{GF}} \right]_{N_f=2,\ T=L_2} \tag{9.28a}$$

L_2/a	ens.	N_{cf}	$\overline{g}^2_{\mathrm{GF}}$	τ_{int}	Q_{fit}
12	A_{12}	301	8.59(6)	1.4(2)	
16	A_{16}	100	7.85(8)	1.8(6)	
20	A_{20}	19	7.64(25)	3(2)	
24	B_{24}	1000	7.54(7)	14(4)	
32	B_{32}	700	7.45(10)	29(9)	
∞	without	$L_2/a = 12$	7.31(10)		0.99
∞	with	$L_2/a = 12$	7.19(7)		0.45

Table 9.4: Results for the GF coupling in $N_f = 2$, determined in the SF at fixed $L_2 \sim 0.8$ fm. τ_{int} is given in units of 2 MDU, where τ_{meas} as given in tab. 9.2 for the different lattice sizes has been taken into account. The values of τ_{int} for $L_2/a = 12, 16$ are to be considered upper bounds, following from the fact that τ_{int} in units of $\tau_{\mathrm{meas}} = 6$ MDU is compatible with no autocorrelation at all.

Given some specific subsets[15] (cf. tab. 9.2) of the respective existing $N_f = 2$ ensembles, labelled by $A_{L_2/a}$ and $B_{L_2/a}$ (cf. tab. 9.4), we evaluate the GF coupling $\overline{g}^2_{\mathrm{GF}}$ (6.7) by integrating the flow equation[16] (6.9). The results are listed in tab. 9.4 and illustrated in fig. 9.3. Due to the implementation of $O(a)$ improvement (cf. sec. 9.2.1), we expect the leading cutoff effects in $\overline{g}^2_{\mathrm{GF}}$ to be quadratic in the lattice spacing, and thus employ the fit[17] ansatz

$$\overline{g}^2_{\mathrm{GF}} = \Phi_1^\star + s_1 \cdot (a/L_2)^2 \tag{9.53}$$

to extract $\Phi_1^\star$. The value for $L_2/a = 12$ seems to deviate from this behavior, suggesting that it is affected by higher order cutoff effects. This has also been observed in [105] for the same lattice spacing, in half the physical volume fixed by the very same LCP (9.30), i.e. $L_1/a = 6$. It is reasonable to exclude this point from the continuum extrapolation, which then leads to

$$\Phi_1^\star = 7.31(10)\ . \tag{9.54}$$

---

[15] For $L_2/a = 12, 16, 20$, MC histories of $\overline{g}^2_{\mathrm{GF}}$ for ensemble subsets A_{12}, A_{16}, A_{20} with somewhat peculiar numbers of configurations N_{cf} - see tab. 9.4 - were readily available from work done in the course of [105] These have been used here. Regarding $L_2/a = 24, 32$, we employed relatively large subsets B_{24}, B_{32} to achieve sufficient statistics, compared to the ensemble subsets $C_{12}, C_{16}, C_{20}, C_{24}, C_{32}$ which will be used in connection with the determination of $\Phi_2^\star, \Phi_3^\star$, see sec. 9.2.4 and sec. 9.2.5, respectively.

[16] An adaptive integrator (see e.g. [105]) was deployed here.

[17] Fit details like the fit parameters, their covariance matrix and the quality of the fit are listed in tab. H.1, see app. H.3.

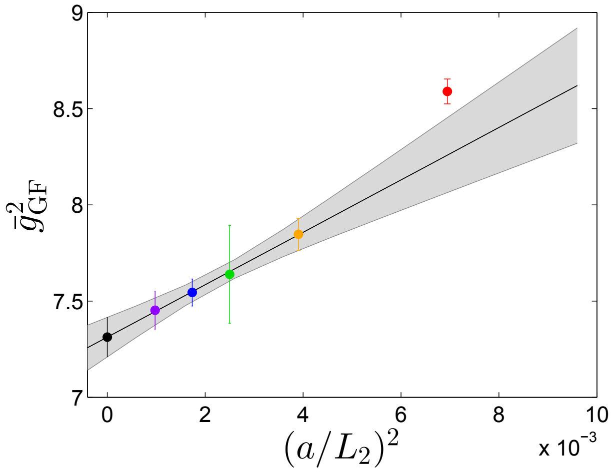


Figure 9.3: The GF coupling and its continuum extrapolation (9.53), which leads to $\Phi_1^\star$ ((9.54), black). The fit excludes the data for $L/a = 12$ (red).

Since we do not know about the GF coupling and its cutoff effects in the $N_f = 3+1$ theory, a rough determination like (9.54) is sufficient for our purposes, and we take its central value for the definition of the LCP.

9.2.4 Determination of $\Phi_2^\star$

Next, we consider

$$\Phi_2^\star = \lim_{a/T\to 0} \left[T \cdot \Gamma^{s\tilde{s}} \right]_{N_f=2,\ T=L_2,\ M_l=M_l^\star} . \tag{9.28b}$$

For the determination of $L_2\Gamma^{s\tilde{s}}$, we again use certain subsets of the $N_f = 2$ ensembles, labelled by $C_{L_2/a}$ (cf. tab. 9.5). The incorporated improved axial current correlation function $f_{A,\mathrm{I}}$ (cf. (9.20)) is evaluated under use of the valence quark hopping parameters $\kappa_l^\star$, listed in tab. 9.3, as well as the improvement coefficient c_A that is available non-perturbatively for the relevant case of $N_f = 2$ and the Wilson gauge action S_G^{W}, see (3.116). This ensures full $O(a)$ improvement, i.e. cutoff effects in $L_2\Gamma^{s\tilde{s}}$ are again expected to be of order $O(a^2)$.

However, we observe that the results for $L_2\Gamma^{s\tilde{s}}$, which can be found in tab. 9.5 and fig. 9.4, can not be reconciled with this behavior, making a reasonable continuum extrapolation impossible. A very similar pattern appears for $L_2\Gamma^{ud}$, calculated with $\kappa = \kappa_{\mathrm{crit}}$ in the chiral limit, from which we know that it is supposed to vanish in the continuum, see (9.24a). We attribute this behavior to the imprecise tuning of am^{ud}_{PCAC}

L_2/a	ens.	$N_{\rm cf}$	$L_2\Gamma^{s\tilde{s}}$	$L_2\Gamma^{ud}$	$L_2\Gamma^{s\tilde{s}}_{\rm cor}$	$\tau_{\rm int}$	$Q_{\rm fit}$
12	C_{12}	500	1.41(7)	0.76(6)	0.65(1)	4(1)	
16	C_{16}	500	0.79(6)	0.11(5)	0.68(1)	7(2)	
20	C_{20}	180	0.51(4)	-0.15(4)	0.66(2)	9(4)	
24	C_{24}	500	0.79(3)	0.19(3)	0.60(1)	17(6)	
32	C_{32}	180	0.48(2)	-0.14(2)	0.62(1)	8(3)	
∞ without $L_2/a = 12$					0.59(1)		0.01
∞ with $L_2/a = 12$					0.61(1)		0.00

Table 9.5: Results of the $N_f = 2$ runs to determine $L_2\Gamma^{s\tilde{s}}_{\rm cor}$ at $z_l^\star = L_2 M_l^\star$. The parameters β, $\kappa_{\rm crit}$ and $\tau_{\rm meas}$ associated with the ensembles are listed in tab. 9.2, while $\kappa_l^\star$ can be found in tab. 9.3. $\tau_{\rm int}$ corresponds to the quantity $L_2\Gamma^{s\tilde{s}}_{\rm cor}$, and is given in units of 2 MDU under consideration of $\tau_{\rm meas}$ (cf. tab. 9.2).

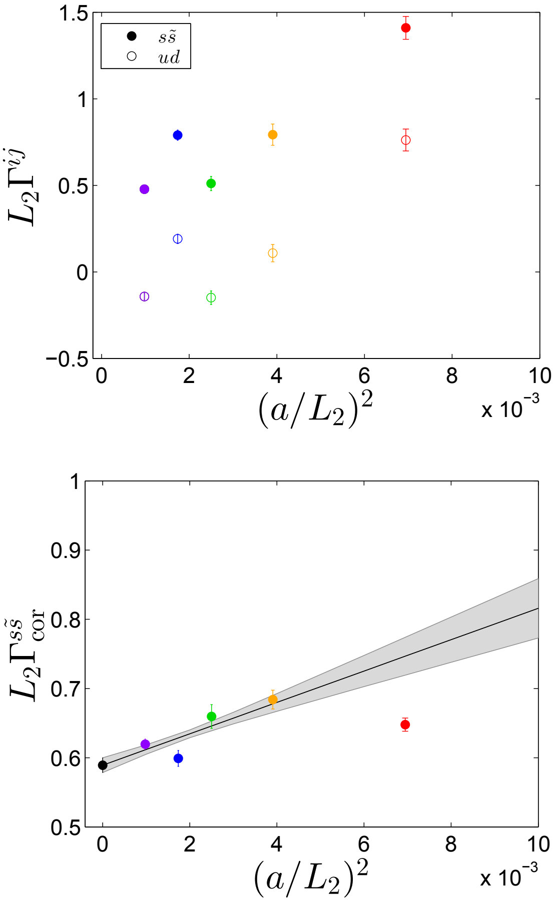


Figure 9.4: *Top:* $L_2\Gamma^{s\tilde{s}}$ at $z_l^\star = L_2 M_l^\star$ (filled circles) and $L_2\Gamma^{ud}$ at $z_u = z_d = 0$ (open circles). *Bottom:* $L_2\Gamma^{s\tilde{s}}_{\rm cor}$ and its continuum extrapolation (9.56), which leads to $\Phi_2^\star$ ((9.57), black). The point for $L_2/a = 12$ (red) is excluded.

to zero, see tab. 9.2, and define a corrected quantity

$$L_2\Gamma^{s\tilde{s}}_{\text{cor}} = L_2\left(\Gamma^{s\tilde{s}} - \Gamma^{ud}\right) \tag{9.55}$$

that differs from $L_2\Gamma^{s\tilde{s}}$ only by cutoff effects. As one can see in fig. 9.4, $L_2\Gamma^{s\tilde{s}}_{\text{cor}}$ shows an improved behavior, and is therefore used for the continuum extrapolation to $\Phi_2^\star$. A fit[18] quadratic in a/L_2,

$$L_2\Gamma^{s\tilde{s}}_{\text{cor}} = \Phi_2^\star + s_2 \cdot (a/L_2)^2 \,, \tag{9.56}$$

gives[19]

$$\Phi_2^\star = 0.59(1) \,. \tag{9.57}$$

Note that again the $L_2/a = 12$ value is not taken into account. As a side remark, we mention that the alternative discretization $\Gamma^{ij}_{(2)}$ of the light EMM gives very similar results to the case of $\Gamma^{ij} = \Gamma^{ij}_{(1)}$ (cf. (9.26)), both on the lattice ($L_2\Gamma^{s\tilde{s}}_{\text{cor}}$) and in the continuum ($\Phi_2^\star$). Details can be found in app. G.2. The same holds if the *unimproved* axial correlation function ($c_A \to 0$) is used for the effective meson masses, see app. G.4.

9.2.5 Determination of $\Phi_3^\star$

Finally,

$$\Phi_3^\star = \lim_{a/T\to 0} \left[\, T \cdot \left(\Gamma^{sc} - \frac{1}{2}\Gamma^{s\tilde{s}}\right) \right]_{N_f=2,\ T=L_2,\ M_l=M_l^\star,\ M_c=M_c^\star} \,, \tag{9.28c}$$

is determined very similarly to the case of $\Phi_2^\star$, on the ensembles $C_{L_2/a}$. We again use the quantities which correct for the massless sea quark contributions,

$$L_2\Gamma^{sc}_{\text{cor}} = L_2\left(\Gamma^{sc} - \Gamma^{ud}\right) \tag{9.58}$$

$$L_2\Gamma^{s\tilde{s}}_{\text{cor}} = L_2\left(\Gamma^{s\tilde{s}} - \Gamma^{ud}\right) \,, \tag{9.55}$$

---

[18] See footnote 17.

[19] The quality of the fit, $Q_{\text{fit}} = 0.01$, is quite poor. This is partly due to the obviously underestimated autocorrelation for $L_2/a = 32$ and the associated error, see tab. 9.5. However, as explained before, a rough determination of the continuum extrapolated value $\Phi_2^\star$ is sufficient for our purposes, so that we make do with the situation.

L_2/a	ens.	N_{cf}	$L_2\Gamma^{sc}$	$L_2\left(\Gamma^{sc}_{\mathrm{cor}} - \frac{1}{2}\Gamma^{s\tilde{s}}_{\mathrm{cor}}\right)$	τ_{int}	Q_{fit}
12	C_{12}	500	7.40(7)	6.32(3)	3(1)	
16	C_{16}	500	6.94(10)	6.49(4)	7(2)	
20	C_{20}	180	6.55(7)	6.37(6)	8(3)	
24	C_{24}	500	6.64(6)	6.15(5)	18(6)	
32	C_{32}	180	6.28(5)	6.11(4)	14(6)	
∞ without $L_2/a = 12$				5.96(5)		0.30
∞ with $L_2/a = 12$				6.17(4)		0.00

Table 9.6: Results of the $N_f = 2$ runs to determine $L_2(\Gamma^{sc}_{\mathrm{cor}} - \frac{1}{2}\Gamma^{s\tilde{s}}_{\mathrm{cor}})$ at $z_l^\star = L_2 M_l^\star$ and $z_c^\star = L_2 M_c^\star$. While the parameters β, κ_{crit}, τ_{meas}, $\kappa_l^\star$ and $\kappa_c^\star$ are given in tab. 9.2 and tab. 9.3, the results on $\Gamma^{s\tilde{s}}$ and Γ_{ud} are the same as in tab. 9.5. τ_{int} corresponds to the quantity $L_2(\Gamma^{sc}_{\mathrm{cor}} - \frac{1}{2}\Gamma^{s\tilde{s}}_{\mathrm{cor}})$, and is given in units of 2 MDU under consideration of τ_{meas} (cf. tab. 9.2).

such that in total, the quantity of interest becomes

$$L_2\left(\Gamma^{sc}_{\mathrm{cor}} - \frac{1}{2}\Gamma^{s\tilde{s}}_{\mathrm{cor}}\right) = L_2\left(\Gamma^{sc} - \frac{1}{2}\Gamma^{s\tilde{s}} - \frac{1}{2}\Gamma^{ud}\right) . \tag{9.59}$$

Note that in comparison to (9.55), the correction is only half as large in absolute size. The results are shown in tab. 9.6 and fig. 9.5. The usual fit[20] ansatz,

$$L_2\left(\Gamma^{sc}_{\mathrm{cor}} - \frac{1}{2}\Gamma^{s\tilde{s}}_{\mathrm{cor}}\right) = \Phi_3^\star + s_3 \cdot (a/L_2)^2 , \tag{9.60}$$

without consideration of the $L/a = 12$ result, yields[21]

$$\Phi_3^\star = 5.96(5) . \tag{9.61}$$

In contrast to the case of $\Phi_2^\star$, the use of the alternative discretization $\Gamma^{ij}_{(2)}$ instead of $\Gamma^{ij} = \Gamma^{ij}_{(1)}$ gives quite different results on the lattice. These relative cutoff effects, investigated in detail in app. G.2, become significant as they enter here in combination with the charm mass. However, we find (cf. (G.9)) $\Phi^\star_{3,(2)} = 5.94(5)$ with $\Gamma^{ij}_{(2)}$, i.e. agreement between the two definitions in the continuum, as expected. The same holds if Γ^{ij} is considered with the *unimproved* axial current ($c_A \to 0$). As one can see in detail in app. G.4, this has a small effect on the lattice results, which vanishes as the continuum

---

[20] See footnote 17.

[21] Footnote 19 applies here as well, for $\Phi_3^\star$.

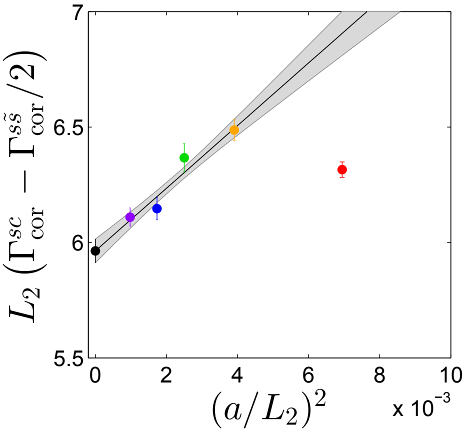


Figure 9.5: $L_2\left(\Gamma^{sc}_{\rm cor} - \Gamma^{s\tilde{s}}_{\rm cor}/2\right)$ and its continuum extrapolation (9.60), which leads to $\Phi_3^\star$ ((9.61), black). The point for $L/a = 12$ (red) is again excluded from the fit.

limit is taken.

9.3 Tuning to the line of constant physics

In the previous two sections, we have defined a LCP,

$$\Phi_1 = \Phi_1^\star\,, \quad \Phi_2 = \Phi_2^\star\,, \quad \Phi_3 = \Phi_3^\star\,, \tag{9.13}$$

in terms of the three quantities

$$\Phi_1 = \bar{g}^2_{\rm GF}(T)\,, \quad \Phi_2 = T\cdot\Gamma^{su}\,, \quad \Phi_3 = T\cdot\left(\Gamma^{sc} - \frac{1}{2}\Gamma^{su}\right)\,, \qquad \text{(9.16, 9.23a, 9.23b)}$$

by fixing the values

$$\Phi_1^\star = 7.31\,, \qquad \Phi_2^\star = 0.59\,, \qquad \Phi_3^\star = 5.96 \qquad \text{(9.54, 9.57, 9.61)}$$

from numerical input of the $N_f = 2$ theory. We are now in the position to implement this LCP in $N_f = 3+1$ with the tree-level improved Lüscher–Weisz gauge action $S_G^{\rm LW}$, which, as discussed in sec. 8.4 and the introduction to the present sec. 9, is the starting point for the determination of

$$c_{\rm sw} = c_{\rm sw}(g_0^2, am^\star_{q,l}, am^\star_{q,c}) \tag{8.63}$$

at fixed, non-vanishing quark masses, to be discussed in sec. 9.4. Affiliated with the LCP is the use of the same SF extent in time and space,

$$T = L \,, \tag{9.62}$$

see (9.32). SF boundary improvement in terms of the improvement coefficients c_t and $\widetilde{c}_t$ is implemented perturbatively like in $N_f = 2$ (see sec. 9.2.1 and (5.37)). However, in the present case of the gauge action S_G^{LW}, these are only known to 1-loop and tree-level order, respectively (cf. (5.42)). The lattice setup is completed by the fermion boundary angle $\theta = 0$ as outlined in (9.15). Details on the algorithmic implementation of the MC simulations will be discussed in sec. 9.5. Once a representative ensemble of gauge configurations is generated, the practical evaluation of the quantities Φ_i in $N_f = 3+1$ on those is straight forward. There are merely two changes with respect to the previously considered $N_f = 2$ case to be accounted for. First, the gradient flow coupling $\bar{g}^2_{\mathrm{GF}}$ acquires a different normalization[22] $\mathcal{N}^{-1}$ due to the use of $S_G = S_G^{\mathrm{LW}}$ (9.1) as gauge action, see (6.13) and the discussion in sec. 6.1.2. Second, while the EMM (cf. (9.26) and (5.53)) that appear in Φ_2 and Φ_3 were defined in terms of the non-perturbatively available c_A (cf. (3.116) and sec. 9.2.4) in $N_f = 2$, we use the perturbative 1-loop expression (3.115) in $N_f = 3+1$.

The task to be addressed in the following is to adjust the bare parameters such that the quantities $\Phi_i \equiv \Phi_i(g_0^2, \kappa_l, \kappa_c)$ assume the values $\Phi_i^\star$ prescribed by the LCP (9.13). This poses a classical optimization problem of a function

$$\vec{\Phi} : \mathbb{R}^3 \to \mathbb{R}^3 \,, \quad \left(g_0^2, \kappa_l, \kappa_c\right) \mapsto \left(\Phi_1, \Phi_2, \Phi_3\right) \,. \tag{9.63}$$

The tuning of the bare parameters, however, needs to (and can only be) carried out up to a certain precision. The tolerated deviations of Φ_i from $\Phi_i^\star$ have to be well defined, which is subject of sec. 9.3.1. After reformulation of the optimization problem in sec. 9.3.2, our strategy to solve it will be discussed in sec. 9.3.3 and sec. 9.3.4. Finally, we discuss the error estimation of the obtained bare parameters at the line of constant physics in sec. 9.3.5, and conclude in sec. 9.3.6.

9.3.1 Imposed precision

We aim to match the physical quantities $T^\star, M_l^\star, M_c^\star$ with a certain precision. Introducing the vector notation $X = (X_1, X_2, X_3) = (T, M_l, M_c)$, they are represented by $X_i^\star$,

---

[22] I am grateful to Alberto Ramos for providing me with the numbers.

while $\delta X_i = X_i - X_i^\star$ are the actual deviations, and $\|\delta X_i\|$ denotes the *(maximal) allowed deviations*:

$$|\delta X_i| \leq \|\delta X_i\| \,, \qquad i \in \{1,2,3\} \,. \tag{9.64}$$

As explained in sec. 8.4, it is particularly important to fix the charm mass on the lattice, i.e. to fix T and M_c rather restrictively, whereas the precision required for M_l is of less importance. The allowed deviations are thus set to

$$\frac{\|\delta T\|}{T^\star}, \frac{\|\delta M_c\|}{M_c^\star} = 5\% \qquad \text{and} \qquad \frac{\|\delta M_l\|}{M_l^\star} = 11\% \,. \tag{9.65}$$

In the following, they will be translated to allowed deviations of Φ_1, Φ_2, Φ_3,

$$(\|\delta T\|, \|\delta M_l\|, \|\delta M_c\|) \longrightarrow (\|\delta \Phi_1\|, \|\delta \Phi_2\|, \|\delta \Phi_3\|) \;. \tag{9.66}$$

Since each Φ_i can be considered as a function $\Phi_i \equiv \Phi_i\,(T, M_l, M_c)$, each deviation $\delta\Phi_i$ gets contributions from *all* the deviations δX_i:

$$\delta\Phi_i = \frac{\partial\Phi_i}{\partial T}\delta T + \frac{\partial\Phi_i}{\partial M_l}\delta M_l + \frac{\partial\Phi_i}{\partial M_c}\delta M_c = \sum_{j=1}^{3}\frac{\partial\Phi_i}{\partial X_j}\delta X_j \,. \tag{9.67}$$

Higher order terms will be neglected, and the derivatives are always understood to be evaluated at the LCP, e.g.

$$\frac{\partial\Phi_i}{\partial X_j} \equiv \left.\frac{\partial\Phi_i}{\partial X_j}\right|_{X=X^\star} \,, \qquad i,j \in \{1,2,3\} \,. \tag{9.68}$$

The maximal allowed deviations, unlike the deviations themselves, can not compensate each other, i.e. one has to assume the most unfavorable addition of deviations:

$$\|\delta\Phi_i\| = \left|\frac{\partial\Phi_i}{\partial T}\right| \|\delta T\| + \left|\frac{\partial\Phi_i}{\partial M_l}\right| \|\delta M_l\| + \left|\frac{\partial\Phi_i}{\partial M_c}\right| \|\delta M_c\| = \sum_{j=1}^{3}\left|\frac{\partial\Phi_i}{\partial X_j}\right| \|\delta X_j\| \,. \tag{9.69}$$

The quantities Φ_i were constructed such that they depend mainly on X_i (see criterion C2 in sec. 9.1), so that

$$\left|\frac{\partial\Phi_i}{\partial X_i}\right| \|\delta X_i\| \gg \left|\frac{\partial\Phi_i}{\partial X_j}\right| \|\delta X_j\| \qquad \text{for } j \neq i \tag{9.70}$$

is supposedly fulfilled. For this reason, and in order to avoid the involvement of off-diagonal derivatives, we propagate only the main error, and approximate the allowed deviations by

$$\|\delta\Phi_i^{\text{approx}}\| = \left|\frac{\partial\Phi_i}{\partial X_i}\right| \|\delta X_i\| \leq \|\delta\Phi_i\| \ . \tag{9.71}$$

The inequality holds because only positive terms are dropped in the approximation, cf. (9.69). We will employ $\|\delta\Phi_i^{\text{approx}}\|$ as upper bounds for the modulus of the deviations $\delta\Phi_i$ from the tuning point. This is a stronger criterion which implies that the tuning target (9.65) for the physical quantities is fulfilled just like the original criterion:

$$|\delta\Phi_i| \leq \|\delta\Phi_i^{\text{approx}}\| \quad \Rightarrow \quad |\delta\Phi_i| \leq \|\delta\Phi_i\| \quad \Leftrightarrow \quad |\delta X_i| \leq \|\delta X_i\| \ . \tag{9.72}$$

We will determine $\|\delta\Phi_i^{\text{approx}}\|$ for $i \in \{1,2,3\}$ via (9.71) in the following.

1) $\delta T \to \delta\Phi_1$

In order to determine $\|\delta\Phi_1^{\text{approx}}\|$ and the involved derivative $\left|\frac{\partial\Phi_1}{\partial T}\right|_{T=T^\star}$, we need the functional relationship of the gradient flow coupling $\Phi_1 = \bar{g}^2_{\mathrm{GF}}$ and T. It is given in perturbation theory at large energies by (2.40), which to lowest order in the Schrödinger functional with finite size scaling $\mu = 1/T$ (cf. (5.18)) becomes

$$\frac{1}{\Phi_1} = -2b_0 \log\left(T\Lambda\right) + \dots \ . \tag{9.73a}$$

The corresponding differential reads

$$d\left(\frac{1}{\Phi_1}\right) = -2b_0 \frac{dT}{T} \ , \tag{9.73b}$$

so that the aspired *relative* maximal allowed deviation of T (cf. (9.65)) translates to an *absolute* maximal allowed deviation of $1/\Phi_1$. Hence, we use this quantity instead of Φ_1 to formulate a precision condition, to ensure that T associated with the outcome of the tuning procedure is symmetrically distributed around $T^\star$ in the statistical limit. From (9.71) and (9.73b), we find an allowed deviation of the function $1/\Phi_1$ that in relative terms amounts to

$$\frac{\|\delta\left(1/\Phi_1\right)^{\text{approx}}\|}{1/\Phi_1^\star} = 2b_0 \ \Phi_1^\star \ \frac{\|\delta T\|}{T^\star} \approx 4\% \ , \tag{9.74}$$

where $b_0 \approx 5.277 \cdot 10^{-2}$ from (2.41) with $N_f = 4$, $\Phi_1^\star = 7.31$ from (9.54) and $\frac{\|\delta T\|}{T^\star} = 5\ \%$ from (9.65) was used. Note that the above use of perturbation theory up to an energy scale as low as $\mu \approx \frac{1}{T^\star} \approx \frac{1}{0.8\ \mathrm{fm}} \approx 250$ MeV is justified by the slow running of the gradient flow coupling, see [149].

2) $\delta M_l \to \delta\Phi_2$ **and** $\delta M_c \to \delta\Phi_3$

In contrast to the previously considered case, the derivatives $\left|\frac{\partial\Phi_2}{\partial M_l}\right|_{M_l=M_l^\star}$ and $\left|\frac{\partial\Phi_3}{\partial M_c}\right|_{M_c=M_c^\star}$ cannot be assessed with analytical methods. Instead, we use our estimates from $N_f = 2$, listed in tab. G.4 of app. G.3:

$$\left|\frac{\partial\Phi_2}{\partial M_l}\right|_{M_l=M_l^\star} \overset{!}{=} \frac{\partial}{\partial M_l}\left(\lim_{a\to 0} L_2\Gamma^{s\tilde{s}}\right)\Bigg|_{N_f=2,\ z_l=z_l^\star} = 2.95(3)\cdot T^\star \tag{9.75a}$$

$$\left|\frac{\partial\Phi_3}{\partial M_c}\right|_{M_c=M_c^\star} \overset{!}{=} \frac{\partial}{\partial M_c}\left(\lim_{a\to 0} L_2\left(\Gamma^{sc} - \frac{1}{2}\Gamma^{s\tilde{s}}\right)\right)\Bigg|_{N_f=2,\ z_l=z_l^\star,\ z_c=z_c^\star} = 0.78(1)\cdot T^\star \tag{9.75b}$$

Together with the target precisions (9.65), this leads to the relative allowed deviations

$$\frac{\|\delta\Phi_2^{\text{approx}}\|}{\Phi_2^\star} = \frac{1}{\Phi_2^\star}\cdot 2.95\ T^\star \cdot 0.11\ M_l^\star \approx 10\% \tag{9.76a}$$

$$\frac{\|\delta\Phi_3^{\text{approx}}\|}{\Phi_3^\star} = \frac{1}{\Phi_3^\star}\cdot 0.78\ T^\star \cdot 0.05\ M_c^\star \approx 4\% \tag{9.76b}$$

where $z_l^\star = T^\star M_l^\star = 0.186$ and $z_c^\star = T^\star M_c^\star = 6.0$ was used, see (9.39) and (9.43), respectively.

9.3.2 Reformulation of the optimization problem

An optimization problem like the one outlined in the introduction of this section, where one is interested in the point where the function $\vec{\Phi}$ assumes a specific value $\vec{\Phi}^\star = (\Phi_1^\star, \Phi_2^\star, \Phi_3^\star)$, is usually trivially reconsidered as the search for the point where the function $\vec{\Phi} - \vec{\Phi}^\star$ vanishes. However, in our case, it should in addition be taken into account that the function $1/\Phi_1$ is to be considered instead of Φ_1. It is convenient in practice to consider the following *distance functions*,

$$D_1(\Phi_1) = -\frac{1/\Phi_1 - 1/\Phi_1^\star}{1/\Phi_1^\star}\,, \qquad D_2(\Phi_2) = \frac{\Phi_2 - \Phi_2^\star}{\Phi_2^\star}\,, \qquad D_3(\Phi_3) = \frac{\Phi_3 - \Phi_3^\star}{\Phi_3^\star}\,, \tag{9.77}$$

which not only vanish at the aspired line of constant physics

$$D_i(\Phi_i^\star) = 0 \, , \tag{9.78}$$

but in terms of which the fulfillment of the relative allowed deviations (9.74) and (9.76) may also neatly be expressed as (cf. (9.72))

$$|D_1| \leq 4\% \, , \qquad |D_2| \leq 10\% \, , \qquad |D_3| \leq 4\% \, . \tag{9.79a}$$

The sign convention in the definition of D_1 is chosen to ensure that D_1 and $\Phi_1 - \Phi_1^\star$ have the same sign. Hence, D_1 monotonically increases with T (cf. (9.73)), just as D_2 and D_3 monotonically increase with M_l and M_c (cf. (9.75)), respectively. An aspect which has not been taken into account so far are the statistical errors of Φ_i that propagate to statistical errors of the distance functions D_i. We do require the tuning criteria (9.79a) to be fulfilled also if those are added to the mean value in an unfavorable way, i.e. we replace

$$|D_i| \;\to\; |D_i| + |\Delta D_i| \tag{9.79b}$$

in (9.79a). The optimization problem is then reformulated in terms of the function[23] (cf. (9.63))

$$\vec{D} : \mathbb{R}^3 \to \mathbb{R}^3 \, , \quad \left(g_0^2, \kappa_l, \kappa_c\right) \mapsto \left(D_1, D_2, D_3\right) \, . \tag{9.80}$$

Although the aspired tuning point is well defined by (9.78) and (9.79), it is crucial for the optimization procedure to assess the *total* distance of a point in bare parameter space that does not yet fulfill the tuning criteria. To achieve this, a (scalar) *global distance function*

$$D : \mathbb{R}^3 \to \mathbb{R} \, , \quad \left(g_0^2, \kappa_l, \kappa_c\right) \mapsto D \tag{9.81}$$

has to be defined, which should

1. be positive semidefinite: $D(\Phi_1, \Phi_2, \Phi_3) \geq 0$.
2. vanish at and only at the LCP: $D(\Phi_1, \Phi_2, \Phi_3) = 0 \quad \Leftrightarrow \quad D_i(\Phi_i) = 0 \ \forall i$.

---

[23] Note that here and in the following, we will frequently use Φ_i and the bare parameters interchangeably as arguments of the distance functions. Although being a slight abuse of notation, the meaning of this is obvious as the mapping $\vec{\Phi}$ (cf. (9.63)) provides a one-to-one correspondence between these two sets of parameters.

3. monotonically increase with the modulus of D_i .

This suggests the basic form

$$D(\Phi_1,\Phi_2,\Phi_3) = f\left(w_1 \cdot (D_1(\Phi_1))^2 + w_2 \cdot (D_2(\Phi_2))^2 + w_3 \cdot (D_3(\Phi_3))^2\right) \ , \tag{9.82}$$

where the factors $w_i > 0$ weigh the individual distance functions in a certain way, and $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is any function that preserves the above criteria. It seems useful to account for the difference in the tolerated relative allowed deviations, cf. (9.79a), i.e. to ensure an equal treatment of the relative deviations of $1/\Phi_1, \Phi_2, \Phi_3$, by considering

$$D(\Phi_1,\Phi_2,\Phi_3) = \sqrt{w \cdot (D_1(\Phi_1))^2 + (D_2(\Phi_2))^2 + w \cdot (D_3(\Phi_3))^2} \tag{9.83a}$$

$$\text{with} \qquad w = (10/4)^2 \ . \tag{9.83b}$$

The chosen overall factor in combination with the square root ensures that D is an upper bound for the individual distance functions, $D_i(\Phi_i) \leq D(\Phi_1,\Phi_2,\Phi_3)\ \forall i$, with equality $D_2 = D$ in the case of $D_1 = D_3 = 0$. Furthermore, D and $\max_i D_i(\Phi_i)$ are of the same order of magnitude. The particular choice of f and w_i will affect the rate of convergence of the iterative tuning procedure in which $D \to 0$ is minimized, to be described in the upcoming section. Note, however, that the actual tuning criteria continue to be expressed in terms of the individual distance functions D_i, via (9.79).

9.3.3 Iterative solution of the optimization problem

Each evaluation of $\vec{\Phi}$ (or equivalently $\vec{D}$) at some point in bare parameter space, from now on denoted by

$$\vec{x} = \left(g_0^2, \kappa_l, \kappa_c\right) \ , \tag{9.84}$$

requires its own MC simulation. Since this is rather expensive in $N_f = 3+1$, an effective iterative optimization technique is desired, to keep the computational effort of the tuning small. This requirement gains even more in importance since, as we will learn in sec. 9.4, the tuning will need to be conducted about

$$\underbrace{\#(\text{lattice sizes})}_{5} \times \underbrace{\#(\text{average no. of } c_{\text{sw}} \text{ per lattice size})}_{O(5)} = O(25) \tag{9.85}$$

times. Our iterative approach schedules the following steps:

S0. Computation of D_i at some initial guess $\vec{x}^{\mathrm{initial}}$ and in its vicinity.

S1. Interpolation of the data for D by a fit function D^{fit}.

S2. Minimization[24] of $\left|D^{\mathrm{fit}}\right|$ leads to bare parameters $\vec{x}^{\mathrm{min}}$.

S3. Computation of D_i at (and possibly in the vicinity of) the bare parameters $\vec{x}^{\mathrm{min}}$.

S4. If tuning criteria are not yet reached, return to step S1.

With regard to step S0, reasonable initial guesses $\vec{x}^{\mathrm{initial}}$ may be obtained by inter- or extrapolation of previous results at other parameters (T/a and c_{sw}). We will return to this in sec. 9.6.1. Furthermore, the suitability and effectiveness of different approaches concerning step S1 and S2 depend on the properties of the function to be optimized. We will discuss our strategy concerning this matter in the following two sections.

9.3.4 Step S1: Interpolation of data

Direct fit $D^{\mathrm{fit}}_{\mathrm{dir}}$ of the global distance function D

As a first, direct approach to the interpolation of D scheduled in step S1, we aim to use 3-dimensional Taylor polynomials of order 1 or 2 around a central point $\vec{x}^{(0)}$ to approximate the data D:

$$D^{\mathrm{fit}}_{\mathrm{dir}}(\vec{x}, F) = F^0 + F' \cdot \left(\vec{x} - \vec{x}^{(0)}\right) + \frac{1}{2}\left(\vec{x} - \vec{x}^{(0)}\right)^T F'' \left(\vec{x} - \vec{x}^{(0)}\right) \ . \tag{9.86}$$

For this scalar function, we have 1 fit parameter F^0 to zeroth order, a gradient vector F' of 3 fit parameters $\left(F^1, F^2, F^3\right)$ to first order, and an additional Hessian matrix F'', i.e. a symmetric matrix of second derivatives with 6 fit parameters $\left(F^4, \ldots, F^9\right)$ to second order. We summarize all these parameters by

$$F = \left\{F^0, F^1, \ldots, F^3, F^4, \ldots, F^9\right\} \ . \tag{9.87}$$

There are two features that make the application of this kind of fit difficult, both related to the form of the global distance function D, (9.83a), which is basically the sum over the squared individual distance functions D_i. First, each D_i, like the associated Φ_i, is strongly dependent and monotonically increasing with exactly one of the bare

---

[24] In principle, one is interested in the one and only zero of D, approximated by D^{fit}. In practice, however, D^{fit} may not possess a zero, which is why we use a more general minimization condition here. Note, furthermore, that a zero of $D^{\mathrm{fit}}_{\mathrm{dir}}$—introduced in (9.86)—may also not be unique. These issues will be discussed further in sec. 9.3.5.

parameters, x_i, while it is by construction expected to rather weakly depend on the others. The data for D_i, however, shows that there is in fact a significant overlap of dependencies on the bare parameters, as a result of criterion C2 (cf. sec. 9.1) not being adequately fulfilled. This leads to a slightly more complicated dependence of D on the bare parameters, and presumably impedes the direct fit of D. Second, and more importantly, there is an issue related to the involved squares. Even if the D_i were perfectly independent, and under the assumption that each of the them is adequately approximated by a Taylor polynomial of first or second order, one automatically needs a Taylor polynomial of higher order to obtain an equally valid fit for D. In turn, a direct fit of D in terms of a Taylor polynomial of first or second order like (9.86) is often not feasible. In practice, this manifests itself as follows. If there are $N^{\mathrm{data}} < 11$ data points and only a linear Taylor polynomial with 4 fit parameters can be used, the quality of the fit is very poor and it serves at most as a very rough approximation. In the case of $N^{\mathrm{data}} \geq 11$ data points, a quadratic Taylor polynomial may be applied, which usually leads to slightly more reasonable fit results, but only if the data is restricted to a relatively small area in bare parameter space. An example will be provided below.

Composite Fit $D^{\mathrm{fit}}_{\mathrm{com}}$ based on the individual distance functions D_i

In view of this, it may be more useful in general to resort to fits of the individual distance functions D_i, which have a much simpler dependence on the bare parameters:

$$D_i^{\mathrm{fit}}(\vec{x}, F_i) = F_i^0 + F_i' \cdot \left(\vec{x} - \vec{x}^{(0)}\right) + \frac{1}{2}\left(\vec{x} - \vec{x}^{(0)}\right)^T F_i'' \left(\vec{x} - \vec{x}^{(0)}\right) . \tag{9.88}$$

These approximations are used to construct a composite fit $D^{\mathrm{fit}}_{\mathrm{com}}$ of the global distance function (cf. (9.83)):

$$D^{\mathrm{fit}}_{\mathrm{com}}(\vec{x}, F_1, F_2, F_3) = \sqrt{w \cdot \left(D_1^{\mathrm{fit}}(\vec{x}, F_1)\right)^2 + \left(D_2^{\mathrm{fit}}(\vec{x}, F_2)\right)^2 + w \cdot \left(D_3^{\mathrm{fit}}(\vec{x}, F_3)\right)^2} , \tag{9.89}$$

where we have summarized the fit parameters of each individual distance function by

$$F_i = \left\{F_i^0, F_i^1, \ldots, F_i^3, F_i^4, \ldots, F_i^9\right\} . \tag{9.90}$$

Practical implementation

In practice, for a given set of data $D_i(\{\vec{x}^1, \ldots, \vec{x}^N\})$, all possible Taylor fits are considered, by varying

- the central point $\vec{x}^{(0)} \in \{\vec{x}^1, \ldots, \vec{x}^N\}$ of the Taylor expansion.
- the number of points N^{fit} included in the fit, with $4 \leq N^{\mathrm{fit}} \leq N^{\mathrm{data}}$.
- the order of the Taylor polynomial (second order only if $N^{\mathrm{fit}} \geq 11$).
- the fit type, $D^{\mathrm{fit}}_{\mathrm{dir}}$ or $D^{\mathrm{fit}}_{\mathrm{com}}$.

Mainly based on χ^2 and the related fit qualities Q_{fit} (cf. app. H), but also under consideration of the range covered by the ingoing data, one of those fits is selected. Consistent with the previous discussion, the best results, expressed by large Q_{fit} and manifest in a fast convergence of the algorithm, are in general obtained under use of the point closest to the LCP as the central point, i.e. $\vec{x}^{(0)} = \vec{x}^m$ with $D(\vec{x}^m) < D(\vec{x}^i)\ \forall\ i \neq m$, a number $N^{\mathrm{fit}} \geq 11$ that allows for second order Taylor polynomials D^{fit}_i employed to approximate the *individual* distance function, which result in $D^{\mathrm{fit}}_{\mathrm{com}}$. As an example, in fig. 9.6, the direct fit $D^{\mathrm{fit}}_{\mathrm{dir}}$ as well as the individual and composite fits D^{fit}_i and $D^{\mathrm{fit}}_{\mathrm{com}}$ are shown, together with $N^{\mathrm{fit}} = N^{\mathrm{data}} = 12$ data points taken from calculations at $T/a = 16, c_{\mathrm{sw}} = 1.7$. As one can see, the individual fits D^{fit}_i of second order describe the respective data rather well, $Q_{\mathrm{fit}}(D^{\mathrm{fit}}_1) = 0.40$, $Q_{\mathrm{fit}}(D^{\mathrm{fit}}_2) = 0.82$, $Q_{\mathrm{fit}}(D^{\mathrm{fit}}_3) = 0.28$. This transfers to $D^{\mathrm{fit}}_{\mathrm{com}}$, which (like the global distance function) is dominated by D_2, even if one takes into account the increased weight of D_1 and D_3 by the factor w (cf. (9.83)). In contrast, the data for D is only moderately well described by the direct fit, $Q_{\mathrm{fit}}(D^{\mathrm{fit}}_{\mathrm{dir}}) = 0.09$.

Although clearly superior in quality concerning the description of data, $D^{\mathrm{fit}}_{\mathrm{com}}$ also has disadvantages in comparison to $D^{\mathrm{fit}}_{\mathrm{dir}}$, to be discussed next.

9.3.5 Step S2: Fit minimization and error estimates

In the subsequent second step of an iteration, the absolute value of the respective fit function $D^{\mathrm{fit}}(\vec{x}) = D^{\mathrm{fit}}_{\mathrm{dir}}(\vec{x}, F)$ or $D^{\mathrm{fit}}(\vec{x}) = D^{\mathrm{fit}}_{\mathrm{com}}(\vec{x}, F_1, F_2, F_3)$ as an approximation of D is minimized, leading to the point

$$\vec{x}^{\mathrm{min}} = \arg\min_{\vec{x}} \left| D^{\mathrm{fit}}(\vec{x}) \right| \,, \tag{9.91}$$

which comes into consideration for fulfilling the tuning criteria (cf. step S3 and S4). $\vec{x}^{\mathrm{min}}$ is obtained numerically under use of standard optimization algorithms. In our case, we apply the (quasi-)Newton method, which is broadly applicable[25] to the mini-

---

[25] Note that in the case of $D^{\mathrm{fit}} = D^{\mathrm{fit}}_{\mathrm{dir}}$, $\vec{x}^{\mathrm{min}}$ is accessible analytically as well, due to the linearity of the fit parameters (cf. (H.5) in app. H.). This, however, is not the case for $D^{\mathrm{fit}} = D^{\mathrm{fit}}_{\mathrm{com}}$.

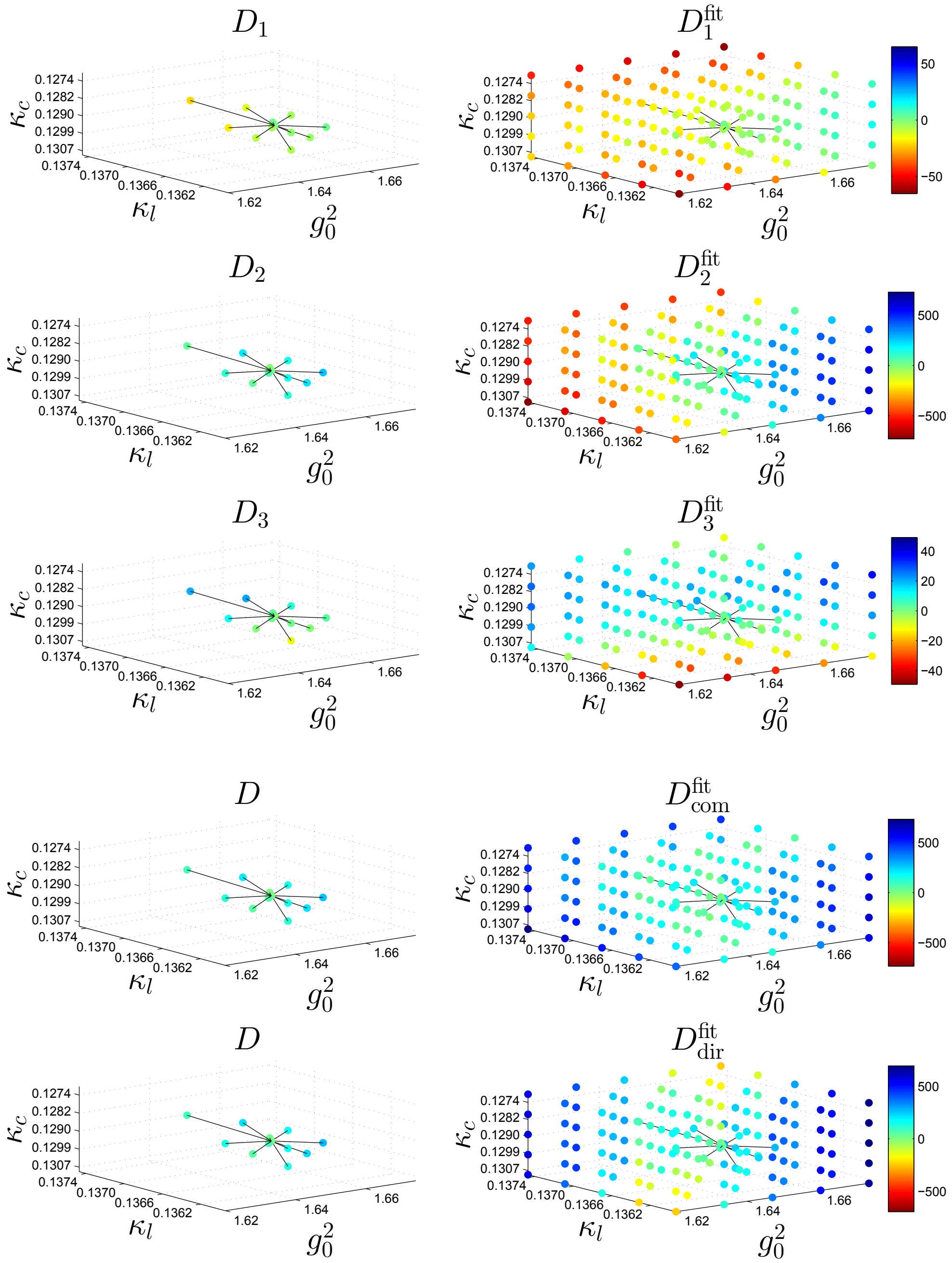


Figure 9.6: Illustration of data *(left panel)* and data embedded in the corresponding fit functions *(right panel)* for the last iteration of the tuning procedure in the case of $T/a = 16$, $c_{\rm sw} = 1.7$. The central point of the Taylor expansion is $\vec{x}^{(0)} = \vec{x}^{\rm min}$, and $N^{\rm fit} = N^{\rm data} = 12$ points are taken into account, all of which are connected to $\vec{x}^{(0)}$ by straight lines to guide the eye. The colormap on the right hand side encodes the function values, respectively. Note that these vary substantially in magnitude between the different D_i and D.

mization of all the fit functions presented in the previous section.

The reason why we refrain from imposing $D^{\mathrm{fit}}_{\mathrm{dir}} = 0$ directly is that this stronger condition can in many cases not be met, i.e. one finds $D^{\mathrm{fit}}_{\mathrm{dir}}(\vec{x}^{\mathrm{min}}) \neq 0$. This holds especially if $D^{\mathrm{fit}} = D^{\mathrm{fit}}_{\mathrm{com}}$ is employed, which may be understood as follows. $D^{\mathrm{fit}}_{\mathrm{com}}$ is positive semidefinite *by definition*, as it is the sum of positive semidefinite terms $\left(D^{\mathrm{fit}}_i\right)^2$. As such, it vanishes if and only if all D^{fit}_i vanish simultaneously. In theory, each (unknown) exact function D_i determines a hyperplane $D_i = 0$ in bare parameter space, and the three hyperplanes defined this way meet at exactly one point, the LCP, which fulfills $D = 0$. However, already the tiniest deviation of a fit D^{fit}_i from the corresponding real functional dependence D_i makes the three hyperplanes $D^{\mathrm{fit}}_i = 0$ not meet in a single point. Hence, the minimum $\vec{x}^{\mathrm{min}}$ of $D^{\mathrm{fit}}_{\mathrm{com}}$ generally leads to[26]

$$D^{\mathrm{fit}}_{\mathrm{com}}(\vec{x}^{\mathrm{min}}) > 0 \,. \tag{9.92}$$

In contrast, in the case of $D^{\mathrm{fit}}_{\mathrm{dir}}$, one usually does find

$$D^{\mathrm{fit}}_{\mathrm{dir}}(\vec{x}^{\mathrm{min}}) = 0 \,, \tag{9.93}$$

at the cost of $\vec{x}^{\mathrm{min}}$ not being unique[27] if a second order fit is employed, in which case it is reasonable to select $\vec{x}^{\mathrm{min}}$ such that the distance to the central point $\vec{x}^{(0)}$ of the expansion is minimal. Note that this behavior, in particular the occurrence of implausible negative values $D^{\mathrm{fit}}_{\mathrm{dir}} < 0$ outside the region covered by data, can also be seen in fig. 9.6. Although (9.92) does not pose a problem in principle, as it still iteratively minimizes the distance to the LCP and may well lead to fulfillment of the tuning criteria, related to this issue, amongst others, is the rather problematic estimation of errors of $\vec{x}^{\mathrm{min}}$, to be discussed next.

Error propagation to $\vec{x}^{\mathrm{min}}$

The error estimates $\Delta\vec{x}^{\mathrm{min}}$ are relevant mainly as they, with regard to step S3 of an iteration, allow to evaluate the D_i in a sensible range of the bare parameters if more than the single computation at the best estimate $\vec{x}^{\mathrm{min}}$ is wanted. Moreover, in principle, if one was interested in matching the LCP exactly, one could continue the iterative process until the proposal $\vec{x}^{\mathrm{min}}$ is compatible with the best available data point, and

---

[26] If the fit qualities of D^{fit}_i are good enough, one may find a value $D^{\mathrm{fit}}_{\mathrm{com}}(\vec{x}^{\mathrm{min}})$ that is compatible with zero within errors though.

[27] Recall that $\vec{x}^{\mathrm{min}}$ is defined as the zero of the *absolute value* of $D^{\mathrm{fit}}_{\mathrm{dir}}$, see (9.91). It is not unique in the case when the quadratic form $D^{\mathrm{fit}}_{\mathrm{dir}}$ itself has a negative minimum.

propagate the associated errors $\Delta\vec{x}^{\mathrm{min}}$ to the results on c_{sw}. We will return to this issue in sec. 9.3.6, after we discuss technical details of the error propagation itself in the following.

Depending on the employed fit D^{fit}, the errors of the data D or D_i are the basic quantities that determine the errors of $\vec{x}^{\mathrm{min}}$. We denote the elements of the data sets by $D^j \equiv D(\vec{x}^j)$ and $D_i^j \equiv D_i(\vec{x}^j)$, respectively, with $j \in \{1,\ldots,N^{\mathrm{data}}\}$. For simplicity, we restrict the discussion to the former case of the *global* distance function D and its direct fit $D^{\mathrm{fit}}_{\mathrm{dir}}$ first. The errors ΔD translate to (correlated) errors ΔF of the fit parameters (cf. (9.86)) according to

$$\begin{aligned}\mathrm{cov}\left(F^\kappa, F^\lambda\right) &= \sum_{j,k} \frac{\partial F^\kappa}{\partial D^j}\ \mathrm{cov}\left(D^j, D^k\right)\ \frac{\partial F^\lambda}{\partial D^k} \\ &= \sum_j \underbrace{\frac{\partial F^\kappa}{\partial D^j}}_{\text{analytical}}\ \left(\Delta D^j\right)^2\ \underbrace{\frac{\partial F^\lambda}{\partial D^j}}_{\text{analytical}}\ , \qquad (9.94\text{a})\end{aligned}$$

where in the second step the independence of the data at different points in bare parameter space,

$$\mathrm{cov}\left(D^j, D^k\right) = \left(\Delta D^j\right)^2\ \delta^{jk}\ , \qquad (9.94\text{b})$$

was used. The partial derivatives in (9.94a) are known analytically for the present case of linear fit parameters, see (H.9) and (H.11) in app. H.2. Hence, also the above covariance matrix of the fit parameters F_i can be obtained analytically. As these covariances are propagated further via the fit function $D^{\mathrm{fit}}_{\mathrm{dir}}$ to the errors[28] $\Delta\vec{x}^{\mathrm{min}}$ sought for,

$$\left(\Delta x^{\mathrm{min}}_\alpha\right)^2 = \sum_{\kappa,\lambda} \underbrace{\frac{\partial x^{\mathrm{min}}_\alpha}{\partial F^\kappa}}_{\text{numerical}}\ \underbrace{\mathrm{cov}\left(F^\kappa, F^\lambda\right)}_{\text{analytical}}\ \underbrace{\frac{\partial x^{\mathrm{min}}_\alpha}{\partial F^\lambda}}_{\text{numerical}}\ , \qquad (9.94\text{c})$$

this does not hold true anymore. In fact, due to the non-analytic nature of the dependence of $\vec{x}^{\mathrm{min}}$ on F_i via the minimization of the corresponding fit $D^{\mathrm{fit}}_{\mathrm{dir}}$, the derivatives in (9.94c) can only be determined numerically. Of course, the nested equations (9.94a)

---

[28] We are only interested in the diagonal entries of the covariance matrix here.

and (9.94c) can also be combined to

$$\left(\Delta x_\alpha^{\min}\right)^2 = \sum_j \underbrace{\left(\frac{\partial x_\alpha^{\min}}{\partial D^j}\right)^2}_{\text{numerical}} \left(\Delta D^j\right)^2 \,, \tag{9.95}$$

where the entire derivatives (that are composed of the derivatives occurring in (9.94)) are treated numerically. We will refer to (9.94) as *semi-analytical* and (9.95) as *numerical* error propagation. The partial derivatives are evaluated numerically in a symmetric way, such that the one that appears in (9.95), for instance, reads

$$\frac{\partial x_\alpha^{\min}}{\partial D^j} = \frac{x_\alpha^{\min}\left(D^1, \ldots, D^j + d\Delta D^j, \ldots, D^N\right) - x_\alpha^{\min}\left(D^1, \ldots, D^j - d\Delta D^j, \ldots, D^N\right)}{2d\Delta D^j} \,. \tag{9.96}$$

The natural scale for the derivatives is given by $d = 1$, but this parameter may be varied to check the numerical results for stability. Inserted in (9.95), we find

$$\left(\Delta x_\alpha^{\min}\right)^2 = \frac{1}{4d^2} \sum_{j=1}^{N} \left[x_\alpha^{\min}\left(D^1, \ldots, D^j + d\Delta D^j, \ldots, D^N\right) - \right.$$
$$\left. x_\alpha^{\min}\left(D^1, \ldots, D^j - d\Delta D^j, \ldots, D^N\right)\right]^2 \,. \tag{9.97}$$

The case of the *individual* fit functions D_i^{fit} is quite similar and emerges from (9.94) and (9.95) after two changes. First, each appearance of D^j is to be replaced by $\sum_i D_i^j$. Second, the equivalent to the covariance matrix (9.94b) is

$$\operatorname{cov}\left(D_i^j, D_l^k\right) = \left(\Delta D_i^j\right)^2 \, \delta^{jk} \, \delta_{il} \,. \tag{9.98}$$

The occurrence of δ_{il} here is equivalent to the neglect of all cross correlations between the different individual distance functions at one and the same $\vec{x}^j$, i.e. $\operatorname{cov}\left(D_i^j, D_l^j\right) \overset{!}{=} 0$. Although in reality, these certainly exist, their computation from $\operatorname{cov}\left(\Phi_i^j, \Phi_l^j\right)$ poses additional effort which seems not to be justified, for reasons which will become clear in sec. 9.3.6, where the discussion will be resumed. We find

$$\left(\Delta x_\alpha^{\min}\right)^2 = \sum_{i,j} \underbrace{\left(\frac{\partial x_\alpha^{\min}}{\partial D_i^j}\right)^2}_{\text{numerical}} \left(\Delta D_i^j\right)^2 \tag{9.99}$$

$$\begin{tikzcd}
\Delta D \arrow[r, "(9.94a)"] \arrow[r, dotted, no head, "(9.95)"'] & \Delta F \arrow[rr, dotted, no head, "\text{semi-analytical}"] \arrow[rr, dotted, no head, "\text{numerical}"'] & & \Delta D^{\mathrm{fit}}_{\mathrm{dir}} \arrow[dr, dotted, "(9.94c)"] \arrow[dr, dotted, shift right] & \\
& & & & \Delta\vec{x}^{\mathrm{min}} \\
\Delta D_i \arrow[r, dotted, no head, "(9.99)"] \arrow[r, shift right] & \Delta F_i \arrow[r, dotted, no head] & \Delta D_i^{\mathrm{fit}} \arrow[r, dotted, no head] & \Delta D^{\mathrm{fit}}_{\mathrm{com}} \arrow[ur, dotted] \arrow[ur, dotted, shift right] &
\end{tikzcd}$$

Figure 9.7: Propagation from the error of the data, either ΔD or ΔD_i, to errors $\Delta\vec{x}^{\mathrm{min}}$ of the iterative estimate $\vec{x}^{\mathrm{min}}$ of the point in bare parameter space that corresponds to the line of constant physics, $D = 0$. For each case, the *semi-analytical* approach to evaluate derivatives is represented by the outer path, whereas the fully *numerical* approach is depicted by the inner path. Solid lines represent analytic error propagation, whereas the error propagation along the dotted lines is conducted numerically. Intermediately occurring quantities do explicitly enter the error analysis only if an arrow points to them, i.e. ΔF or ΔF_i in the semi-analytical approach.

for the *numerical* error propagation.

The propagation of errors from ΔD or ΔD_i to $\Delta\vec{x}^{\mathrm{min}}$ as discussed above is illustrated diagrammatically in fig. 9.7. We use the results of the fully *numerical* evaluation of derivatives, (9.95) and (9.99), while the *semi-analytical* approach (e.g. (9.94)) may serve as a cross check. Note, however, that the use of the individual data D_i^j suffers from

- the neglect of covariances, (9.98)
- $\vec{x}^{\mathrm{min}}$ being the minimum (9.92) instead of the zero (9.93) of the fit function
- a factor 3 in the number of terms summed over, cf. (9.95) and (9.99).

This manifests itself in larger, less reliable, and more unstable (under variation of the approach to estimate partial derivatives, as well as the parameter d, cf. (9.96)) error estimates with respect to the use of D^j. Concerning the error analysis, $D^{\mathrm{fit}}_{\mathrm{dir}}$ is thus advantageous over $D^{\mathrm{fit}}_{\mathrm{com}}$.

9.3.6 Conclusions

The main properties of the two fit approaches $D^{\mathrm{fit}} = D^{\mathrm{fit}}_{\mathrm{dir}}$ and $D^{\mathrm{fit}} = D^{\mathrm{fit}}_{\mathrm{com}}$, discussed in the previous sections, are summarized in tab. 9.7, showing the complementary advantages and disadvantages. While the requirement to accurately describe data is

D^{fit}	$Q_{\mathrm{fit}}(N^{\mathrm{data}} < 10,\ d=1)$	$Q_{\mathrm{fit}}(N^{\mathrm{data}} \geq 10,\ d=2)$	$D^{\mathrm{fit}}(\vec{x}^{\mathrm{min}})$	Error propagation
$D^{\mathrm{fit}}_{\mathrm{dir}}$	poor	moderate	$= 0$	good
$D^{\mathrm{fit}}_{\mathrm{com}}$	moderate	good	$\neq 0$	poor

Table 9.7: Summary of general properties of the direct fit $D^{\mathrm{fit}}_{\mathrm{dir}}$ (9.86) and the composite fit $D^{\mathrm{fit}}_{\mathrm{com}}$ (9.89). N^{data} denotes the number of data points, which allow for the use of Taylor polynomials of order d.

met primarily by $D^{\mathrm{fit}}_{\mathrm{com}}$, as we have learned in sec. 9.3.4, $D^{\mathrm{fit}}_{\mathrm{dir}}$ is conceptually cleaner and allows to properly estimate the errors $\Delta\vec{x}^{\mathrm{min}}$ of the point at the LCP, according to the discussion in sec. 9.3.5.

During the tuning procedure, the fit quality is more important than proper error estimates. Consequently, the use of $D^{\mathrm{fit}}_{\mathrm{com}}$ crucially accelerates its convergence. $D^{\mathrm{fit}}_{\mathrm{dir}}$ is not very useful in this regard, but may serve as a complement especially to estimate the (order of) errors and assess a range in bare parameter space for the subsequent search for the tuning point. If one aims at propagating the error $\Delta\vec{x}^{\mathrm{min}}$ at the end of the tuning procedure, one needs both a satisfying fit quality Q_{fit} and a reliable estimation of the error $\Delta\vec{x}^{\mathrm{min}}$. In principle, both $D^{\mathrm{fit}}_{\mathrm{dir}}$ and $D^{\mathrm{fit}}_{\mathrm{com}}$ (the latter without the neglect of covariances, cf. (9.98)), may be applied. However, one would need to gather a lot of high quality data in the immediate vicinity of the estimated LCP for reliable fits. In comparison to our approach to simply abort the iterative procedure once the tuning criteria (9.79) are fulfilled, this amounts to a huge additional effort. This, however, is not justified because, as we will see in sec. 9.6.4, the obtained errors $\Delta\vec{x}^{\mathrm{min}}$, if propagated to c_{sw}, are presumably small compared to the errors that stem from the determination of c_{sw} (in terms of an improvement condition, cf. sec. 9.4) itself.

9.4 Improvement condition

In this section, we discuss the strategy to determine c_{sw} at the bare parameters $\{g_0^2, aM\}$ determined by the LCP, by formulation of an appropriate improvement condition based on the PCAC relation, established in [46]. Although the presentation is tailored to the massive $N_f = 3+1$ theory, the difference to the massless case is—as will become clear below—insignificant for its implementation.

As we know from sec. 3.6, sec. 3.7 and app. C, the PCAC relation is violated in the Wilson regularized lattice theory due to the breaking of chiral symmetry by the Wilson term (3.43):

$$\tilde{\partial}_\mu \left\langle (A_{\mathrm{RI}})^{ij}_\mu (x)\; O_{\mathrm{R}} \right\rangle = 2m^{ij}_{\mathrm{R}} \left\langle (P_{\mathrm{RI}})^{ij} (x)\; O_{\mathrm{R}} \right\rangle + O(a^z)\;, \tag{9.100}$$

or equivalently

$$m^{ij}_{\mathrm{R}} = \frac{\tilde{\partial}_\mu \left\langle (A_{\mathrm{RI}})^{ij}_\mu (x)\; O_{\mathrm{R}} \right\rangle}{2 \left\langle (P_{\mathrm{RI}})^{ij} (x)\; O_{\mathrm{R}} \right\rangle} + O(a^z)\;, \tag{9.101}$$

which, very similarly to (3.80), amounts to a definition of the *renormalized mass* m^{ij}_{R}. In the above equations, we have $z = 1$ in the unimproved and $z = 2$ in the improved theory. Recall further that $(A_{\mathrm{I}})^{ij}_\mu$ and $(P_{\mathrm{I}})^{ij}$ are the improved flavor *non-singlet* axial and pseudoscalar current (cf. (3.73, 3.74, 3.75)) with quark flavors $j \neq i$, respectively, as known from sec. 8. In contrast to the massless case considered before, there are quite a few more improvement coefficients in the massive $N_f = 3+1$ case, namely $\overline{b}_A, b_A, b_P, b_g$ and c_{sw}, c_A in the mass-independent scheme,

$$(A_{\mathrm{RI}})^{ij}_\mu = Z_A(g_0^2) \left[1 + a \cdot \overline{b}_A(g_0^2)\mathrm{Tr}[M] + a \cdot \frac{b_A(g_0^2)}{2}(m_{q,i} + m_{q,j})\right] \left[A^{ij}_\mu + a \cdot c_A(g_0^2)\partial_\mu P^{ij}\right] \tag{8.33}$$

$$(P_{\mathrm{RI}})^{ij} = Z_P(g_0^2) \left[1 + a \cdot \overline{b}_P(g_0^2)\mathrm{Tr}[M] + a \cdot \frac{b_P(g_0^2)}{2}(m_{q,i} + m_{q,j})\right] P^{ij}\;, \tag{8.37}$$

but only c_{sw}, c_A in the mass-dependent scheme (cf. (8.46b)):

$$(A_{\mathrm{RI}})^{ij}_\mu = \widetilde{Z}^{ij}_A(g_0^2, aM) \left[A^{ij}_\mu + a \cdot c_A(g_0^2, aM)\partial_\mu P^{ij}\right] \tag{9.102}$$

$$(P_{\mathrm{RI}})^{ij} = \widetilde{Z}^{ij}_P(g_0^2, aM, a\mu)\; P^{ij}\;. \tag{8.47b}$$

For the sake of brevity, we use the mass-dependent scheme in the following. However, the particular renormalization scheme does not play a role for the imposition of the improvement condition, as we will see soon. Employing the Schrödinger functional with its boundary operators as well as the associated Schrödinger functional correlation functions, we find

$$m^{ij}_{\mathrm{R}} = \frac{\widetilde{Z}^{ij}_A(g_0^2, aM)}{\widetilde{Z}^{ij}_P(g_0^2, aM, a\mu)} \cdot m^{(\prime)ij}_{\mathrm{PCAC,I}}(x_0) + O(a^z)\;, \tag{9.103}$$

under use of the bare improved PCAC mass in the SF,

$$m^{ij}_{\text{PCAC,I}}(x_0) = \frac{\tilde{\partial}_0 \left\langle (A_{\text{I}})^{ij}_0(x)\ \mathcal{O}^{ji} \right\rangle}{2\left\langle P^{ij}(x)\ \mathcal{O}^{ji} \right\rangle} = \frac{\tilde{\partial}_0 f^{ij}_{A,\text{I}}(x_0)}{2 f^{ij}_P(x)} = \frac{\tilde{\partial}_0 f^{ij}_A(x_0) + a c_A \tilde{\partial}^2_0 f^{ij}_P(x_0)}{2 f^{ij}_P(x)}\ , \qquad (5.54)$$

or its primed (backward) counterpart $m'^{ij}_{\text{PCAC,I}}$ (cf. (5.52)). Up to now, in all equations for the PCAC masses, we have suppressed quite a few of its arguments. In the present case, the improved PCAC mass $m^{ij}_{\text{PCAC,I}}$ inherits a dependence on g_0^2, aM, c_{sw} from the action and on c_A from A_{I}. Moreover, it depends on x_0 like the inserted currents, the choice of the operator $\mathcal{O}^{ji}$ or $\mathcal{O}'^{ji}$ (the latter case being associated with $m'^{ij}_{\text{PCAC,I}}$), as well as all the parameters of the Schrödinger functional, namely the temporal and spatial sizes T and L, the boundary gauge fields C_k and C'_k (cf. (5.7, 5.25)), and the fermion angle θ (cf. (5.21)). If we summarize the latter parameters by

$$P_{\text{SF}} = \{x_0, \mathcal{O}^{ji}, T, L, C_k, C'_k, \theta\}\ , \qquad (9.104)$$

and the renormalization constants by

$$\widetilde{Z}^{ij}_{m_{\text{PCAC}}}(g_0^2, aM, a\mu) = \frac{\widetilde{Z}^{ij}_A(g_0^2, aM)}{\widetilde{Z}^{ij}_P(g_0^2, aM, a\mu)}\ , \qquad (9.105)$$

equation (9.103), with all arguments explicitly listed, becomes[29]

$$m^{ij}_{\text{R}} = \widetilde{Z}^{ij}_{m_{\text{PCAC}}}(g_0^2, aM, a\mu) \cdot\ m^{ij}_{\text{PCAC,I}}(g_0^2, aM, c_{\text{sw}}, c_A, P_{\text{SF}}) + O(a^z)\ . \qquad (9.106)$$

The idea for the determination of the improvement coefficients c_{sw} and c_A at given bare parameters $\{g_0^2, aM\}$ is the following. If they assume the correct values, referred to as $c_{\text{sw,I}}$ and $c_{A,\text{I}}$ from now on, this implies $z = 2$ in (9.106) irrespective of the parameters P_{SF}. In other words, the dependence on P_{SF} is a pure lattice artifact, which is required to vanish up to $O(a^2)$ effects. This fact may in turn be used to adjust the improvement coefficients such that both chiral symmetry is restored and the cutoff effects of first order in the lattice spacing disappear in on-shell quantities. The PCAC relation is particularly useful to this end, as it has been shown to suffer from large cutoff effects both perturbatively [46] and non-perturbatively [150], rendering it particularly sensitive to c_{sw}. Hence, one may formulate the following *(principal) improvement condition*

---

[29] Note that the case of primed backward PCAC masses, previously considered in (5.54), is included here as part of P_{SF}.

to determine $c_{\mathrm{sw,I}}(g_0^2, aM)$ and $c_{A,\mathrm{I}}(g_0^2, aM)$:

$$m^{ij}_{\mathrm{PCAC,I}}(g_0^2, aM, c_{\mathrm{sw,I}}, c_{A,\mathrm{I}}, P^{(1)}_{\mathrm{SF}}) \overset{!}{=} m^{ij}_{\mathrm{PCAC,I}}(\,\dots\,, P^{(2)}_{\mathrm{SF}}) \overset{!}{=} m^{ij}_{\mathrm{PCAC,I}}(\,\dots\,, P^{(3)}_{\mathrm{SF}})\,. \qquad (9.107)$$

Regarding (9.106), this can be understood as one bare PCAC mass to eliminate the numerical constant $\widetilde{Z}^{ij}_{m_{\mathrm{PCAC}}}$, and two more to fix the improvement coefficients $c_{\mathrm{sw,I}}$ and $c_{A,\mathrm{I}}$. It is now obvious that the non-vanishing and non-degeneracy of the quark masses in $N_f = 3+1$ as well as the employed renormalization scheme and potential improvement coefficients of the axial and pseudoscalar current, all encoded in $\widetilde{Z}^{ij}_{m_{\mathrm{PCAC}}}$, play no significant role in the procedure, as mentioned above.

From the many possibilities to probe the quark dynamics by variation of the different parameters embodied in P_{SF}, one needs to select one or two of those. It is beneficial to use x_0 and $\mathcal{O}^{(\prime)ij}$, as this choice requires merely one set of Monte Carlo simulations, at all other parameters in P_{SF} fixed. Hence, using the short notation again, one avails oneself of

$$m^{ij}_{\mathrm{PCAC,I}}(x_0),\; m'^{ij}_{\mathrm{PCAC,I}}(x_0),\; m^{ij}_{\mathrm{PCAC,I}}(y_0),\; m'^{ij}_{\mathrm{PCAC,I}}(y_0) \qquad (9.108)$$

at two different time slices $x_0 \neq y_0$. In order to avoid the need to simultaneously vary c_{sw} and c_A, it is convenient to consider the two principal conditions in (9.107) separately. Moreover, instead of equating PCAC masses at different time slices x_0 and y_0 (which would be necessary for (9.107)), particularly useful improvement conditions [16] are imposed by the equations

$$\text{I)}\quad m^{ij}_{\mathrm{PCAC,I}}(y_0) \overset{!}{=} m'^{ij}_{\mathrm{PCAC,I}}(y_0) \qquad (9.109)$$

$$\text{II)}\quad m^{ij}_{\mathrm{PCAC,I}}(x_0) \overset{!}{=} m'^{ij}_{\mathrm{PCAC,I}}(x_0)\,, \qquad (9.110)$$

i.e. under consideration of all four PCAC masses listed in (9.108). In practice, the first condition is used to determine $c_{A,\mathrm{I}}$. To this end, the PCAC mass is split according to

$$m^{(\prime)ij}_{\mathrm{PCAC,I}}(x_0) = r^{(\prime)ij}(x_0) + c_A\, s^{(\prime)ij}(x_0) \qquad (9.111\mathrm{a})$$

with[30]

$$r^{(\prime)ij}(x_0) = \frac{\frac{1}{2}\,(\partial_0^\star + \partial_0)\, f_A^{(\prime)ij}(x_0)}{2 f_P^{(\prime)ij}(x_0)} \quad \text{and} \quad s^{(\prime)ij}(x_0) = \frac{a\partial_0^*\partial_0 f_P^{(\prime)ij}(x_0)}{2 f_P^{(\prime)ij}(x_0)}\,. \qquad (9.111\mathrm{b})$$

---

[30] Note that the second time derivative $\tilde{\partial}_0^2$ is replaced here by $\partial_0^*\partial_0$, cf. app. A.3.

The improvement coefficient $c_{A,\text{I}} \equiv c_{A,\text{I}}(y_0)$ that leads to the fulfillment of (9.109) is thus determined at the time slice y_0 by

$$\text{I}') \quad c_{A,\text{I}}(y_0) = -\frac{r'^{ij}(y_0) - r^{ij}(y_0)}{s'^{ij}(y_0) - s^{ij}(y_0)} \; . \tag{9.112}$$

The subsequent application of $c_{A,\text{I}}$ in (9.110) is expressed by the use of *effective quark masses* (cf. (9.111a))

$$M^{(\prime)ij}(x_0, y_0) = r^{(\prime)ij}(x_0) + c_{A,\text{I}}(y_0)\, s^{(\prime)ij}(x_0) \; , \tag{9.113}$$

in terms of which the *improvement condition* for $c_{\text{sw,I}}$ becomes

$$\text{II}') \quad \Delta M^{ij}(x_0, y_0) := M^{ij}(x_0, y_0) - M'^{ij}(x_0, y_0) \overset{!}{=} 0 \; . \tag{9.114}$$

Other improvement conditions are in principle possible, leading to different results covered by the inevitable $O(a)$ ambiguity of the improvement coefficients[31]. However, the improvement condition (9.114) is characterized by its proximity to perturbation theory, i.e. the tree-level value of $\Delta M^{ij}(x_0, y_0)$ is very small [151] and the results for $c_{\text{sw,I}}$ stay rather close to their perturbative counterpart. This issue is further elucidated in app. I.3, and is the reason why (9.114) has found successful application in previous determinations of c_{sw}.

The two distinct valence quark flavors i and j that enter ΔM^{ij} still have to be chosen. In our case of $N_f = 3+1$, these can have the mass either of the light quark, $m^{\star}_{q,l}$, or the charm quark, $m^{\star}_{q,c}$. Since we aim to improve primarily *light* quark physics, i.e. correlation function where the charm quark enters only as a sea quark, we opt for two light quarks. The *improvement condition* in its final form thus reads

$$\Delta M^{ud}(x_0, y_0) = M^{ud}(x_0, y_0) - M'^{ud}(x_0, y_0) \overset{!}{=} 0 \; . \tag{9.115a}$$

The time slice arguments x_0 and y_0 also need to be specified. They should be far away from the boundaries, but also from each other, so that

$$x_0 = \frac{3}{4}T \quad \text{and} \quad y_0 = \frac{1}{4}T \tag{9.115b}$$

is a reasonable choice. Note that unlike in the case of previous determinations of c_{sw}, cf. (F.3), we don't use the tree-level value of ΔM^{ud} on the right hand side of (9.115a).

---

[31] The effective quark mass $M^{ij}(x_0, y_0)$, for instance, depends on the definition of $c_{A,\text{I}}$ only at order $O(a^2)$.

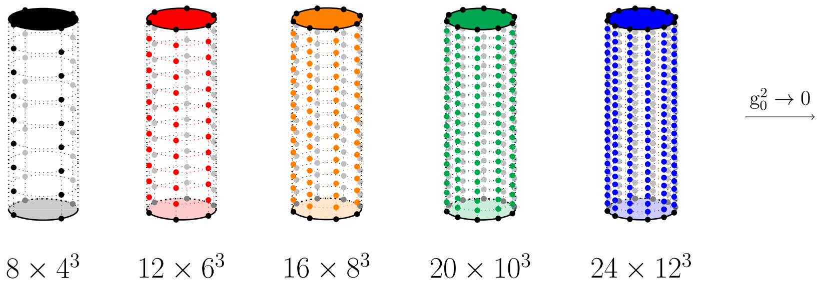


Figure 9.8: c_{sw} is determined at $T = T^\star \sim 0.8$ fm, but—in contrast to the LCP simulations—with $L = T/2$. Moreover, non-vanishing boundary fields (9.119) are used, indicated by the colored sides. Compare to fig. 9.1.

Moreover, it can easily be shown that the arguments x_0 and y_0 are interchangeable. In particular, the improvement conditions (9.112) and (9.114) imply

$$c_{A,\text{I}}(x_0) = c_{A,\text{I}}(y_0) \tag{9.116a}$$

$$\text{and} \qquad \Delta M^{ij}(y_0, x_0) = 0 \,. \tag{9.116b}$$

Finally, to complete the definition of ΔM^{ud} and the improvement condition, we need to fix the remaining Schrödinger functional parameters

$$\{T, L, C_k, C_k', \theta\} \in P_{\text{SF}} \,, \tag{9.117}$$

cf. (9.104). As described in sec. 9.1, a rather large volume (cf. (9.27)) is needed for the line of constant physics. In contrast, the improvement condition is better imposed in a smaller volume, as ΔM^{ud} is cheaper to compute and more sensitive to c_{sw} in that case (see e.g. [22, 46]). Since the bare parameters determined by the LCP have to stay the same though, in regard to the setup of those simulations, we replace

$$L \to L/2 \tag{9.118}$$

while keeping the lattice spacing constant, i.e. we change $L/a \to (L/2)/a$. However, as the size of boundary effects ($O(a^2)$ after improvement) in the measurement of ΔM^{ud} decreases with the distance of the measurement points $x_0 = 3T/4$ and $y_0 = T/4$ from the boundary, it is advantageous to keep T large. Therefore, while decreasing L, we leave $T = T^\star$ unaltered. The SF setup of the improvement simulations is illustrated in fig. 9.8. In addition, the computation of ΔM^{ud} requires non-vanishing boundary gauge fields, because otherwise, with a zero SF background field, $F_{\mu\nu}$ vanishes to tree-level order in perturbation theory, such that the correlation functions entering the improvement condition become insensitive to $c_{\rm sw}$ (which multiplies $F_{\mu\nu}$ in the improved action) to that order, see [46]. On the other hand, the background field in lattice units ($a^2F_{\mu\nu}$) should also not be too strong to avoid large lattice effects. A good compromise between the two criteria, which leads to a non-vanishing but small background field (5.27), is given by the boundary conditions [16, 46]

$$(\phi_1, \phi_2, \phi_3) = \frac{1}{6}\,(-\pi, 0, \pi) \tag{9.119a}$$

$$(\phi_1', \phi_2', \phi_3') = \frac{1}{6}\,(-5\pi, 2\pi, 3\pi)\ , \tag{9.119b}$$

cf. (5.7) and (5.25). These have proven to be useful in previous determinations of $c_{\rm sw}$, and will also be employed in this work. Concerning the SF boundary improvement coefficients c_t and $\widetilde{c}_t$ as well as the boundary fermion angle, we stick to the perturbative expressions as well as $\theta = 0$ as employed in the tuning simulations, see sec. 9.3.

9.5 Setup of the numerical simulations

In this section, we provide information on the numerical simulations performed to *tune*[32] the bare parameters to the LCP (cf. sec. 9.3) and to impose the *improvement*[33] condition (cf. sec. 9.4). For convenience, a condensed overview on the respective lattice setups, gathered from the discussion given in previous chapters, can be found in tab. 9.8.

General procedure and employed software

As has been explained in detail in sec. 4.1, the gauge field and fermion integrations are treated separately, see (4.12). The production of representative gauge field configurations is executed under use of the `openQCD` package[34] [77, 98] and the contained implementation of Schrödinger functional boundary conditions. The gradient flow coupling $\Phi_1 = \overline{g}^2_{\rm GF}$ is the only pure gauge field observable relevant for our purposes (tuning). As such, its evaluation does not involve the integration of fermion fields, and it can be measured "on the fly" by use of the built-in adaptive integrator (see e.g. [105]). In contrast, the quantities Φ_2, Φ_3 (tuning) and $m^{ud}_{\rm PCAC}, M^{ud}, \Delta M^{ud}$ (improvement) de-

---

[32] We will refer to the respective simulations as "tuning" or "improvement" simulations in the following.

[33] See footnote 32.

[34] Version 1.2.

	$N_f = 2$ (LCP)	$N_f = 3+1$ (LCP)	$N_f = 3+1$ (IMPR)	Ref.
S_G	S_G^{W}	S_G^{LW}	S_G^{LW}	(9.33), (9.1)
S_{flow}	S_G^{W}	S_G^{W}	S_G^{W}	(9.18)
T	$T^\star \sim 0.8$ fm	$T^\star \sim 0.8$ fm	$T^\star \sim 0.8$ fm	(9.27)
L	T	T	$T/2$	(9.32), (9.62), (9.118)
ϕ, ϕ'	0	0	$\neq 0$	(9.14), (9.119)
θ	0.5	0	0	(9.15)
c_{sw}	NP	—	—	(3.114)
c_t	2-loop	1-loop	1-loop	(5.37a), (5.42a)
$\widetilde{c}_t$	1-loop	tree-level	tree-level	(5.37b), (5.42b)
c_A	NP	1-loop	implicitly NP	(3.116), (3.115)

Table 9.8: Comparison of the lattice setups (illustrated in fig. 9.1, fig. 9.8) and improvement coefficients used in the $N_f = 2$ and $N_f = 3+1$ simulations. In the case of the latter, we differentiate between the tuning (LCP) and the improvement (IMPR) runs. The last column refers to the respective equations where details can be found. NP stands for "non-perturbative".

pend on the Schrödinger functional correlation functions introduced in sec. 5.3. The fermion integration for the determination of those leads to propagators (cf. (4.9) and [152]), whose computation is performed under use of the respective gauge field ensembles and with the help of the `sfcf` package[35] [152]. The ensemble mean values and error estimates are provided for all observables by the `UWerr` package [88] and its implementation of the gamma method (cf. sec. 4.10). In this context, $\Phi_1 = \overline{g}^2_{\mathrm{GF}}$ is a *primary observable*, while the other quantities are treated as *derived observables*.

In the following sec. 9.5.1, we present details on the algorithmic setup of our simulations. Afterwards, in sec. 9.5.2, we describe the way these are implemented in practice.

9.5.1 Algorithmic setup

Although our simulations are of the type $N_f = 3+1$ with respect to the masses of the dynamical quarks, only two of the degenerate quarks (the up and down) may be treated collectively in MC simulations, as has been discussed in sec. 4.4. The strange and charm quark, in contrast, are implemented solely in terms of the RHMC and

---

[35] I am grateful to Christian Wittemeier, Jochen Heitger and Hubert Simma for providing me with the code.

the associated reweighting (sec. 4.6). Hence, from the algorithmic point of view, we conduct $N_f = 2+1+1$ simulations. The parameters that govern the performance of the Markov chain Monte Carlo algorithm need to be chosen with care, in order for the following requirements to be fulfilled. First, of course, it should be ensured that

#1. the computation of a MC step is fast and effective.

Moreover, as the computations will be rather expensive in $N_f = 2+1+1$ anyway,

#2. the expectation value

$$P_{acc} = \left\langle \min\left(1, e^{-\Delta H}\right)\right\rangle \,, \tag{9.120}$$

of the acceptance probability (cf. (4.53)) should be rather large. We aim for

$$P_{acc} \in [0.90, 0.99] \,. \tag{9.121}$$

However, an eye also has to be kept on feasibility, especially as the tuning procedure schedules the simulation at many different points in bare parameter space (cf. sec. 9.3). Hence,

#3. the amount of algorithmic parameters should be moderate to keep the associated effort to adjust them small.

In the following, we will discuss the algorithmic choices one by one, ordered by their affiliation to the different areas of a MC simulation covered in sec. 4.

RHMC (sec. 4.6)

For the correctness of the RHMC algorithm, the spectra $[r_{f,a}^2, r_{f,b}^2]$ of the operators $Q_f Q_f^\dagger$ with $f \in \{s, c\}$ need to be accurately estimated, the practical implementation of which will be discussed in sec. 9.5.2. Given that this is the case, the maximal relative deviation δ of the rational approximation R_f, cf. (4.66), decreases with the number n of poles, while the numerical effort increases, cf. (4.75). We thus opt for the *minimal n that fulfills the condition (4.69)*. In practice, for the charm quark, $n = 6$ is virtually always sufficient, while the strange quark requires either $n = 6$ or $n = 8$ (as we will see in tab. 9.12 and tab. 9.13).

T/a	8	12	16	20	24
μ	1.2	1.0	1.0	0.5	0.5

Table 9.9: Hasenbusch frequency splitting masses μ (cf. (4.97)) employed for the various lattice sizes T/a.

Solvers (sec. 4.7)

We use the *conjugate gradient* (CG) for the degenerate flavors $f \in \{u, d\}$, and the *multi-shift conjugate gradient* (MSCG) for the rationally approximated pseudofermion actions of the flavors $f \in \{s, c\}$. The reason why we refrain from the more sophisticated solvers described in sec. 4.7, is mainly that the computational gain turns out to be small compared to the increased effort to tune the involved parameters. With regard to the *deflated* solvers, this is arguably due to the use of SF boundary conditions and the induced mass gap (cf. (5.23)), as well as the fact that the up and down masses in our simulations are substantially far away from the chiral limit, close to which deflation is most useful. The use of *SAP* solvers is not suitable for our lattices[36].

Hasenbusch frequency splitting (sec. 4.8)

In order to empirically optimize the efficiency of the pseudofermion integration, several test simulations with different versions of the splitting under use of various μ^2 were performed. With all other parameters of the simulations held fixed, the acceptance as well as the MC time per trajectory were compared. Although the multiple splitting of the up and down quark fermion determinant was also considered, it turned out to be sufficient to use two factors as in (4.97). For simplicity, one and the same μ^2 is employed for all simulations at a certain T/a. The results can be found in tab. 9.9.

Twisted mass reweighting (sec. 4.8)

The use of twisted mass reweighting is not necessary as the quark masses differ significantly from zero (cf. (9.11)).

Multiple time scale integration (sec. 4.9)

The integration of the gluonic and the different fermionic forces is performed in the framework of multiple time scale integration. For the sake of simplicity, i.e. with

---

[36] This is due to the fact that technically, in the `openQCD` code, an additional time slice is used for the implementation of the Schrödinger functional.

T/a	8		12		16		20		24	
L/a	8	4	12	6	16	8	20	10	24	12
$L_{\rm local}$	4	4	6	6	8	8	10	10	(6,4,4)	6
$\prod_i n_i$	8	1	8	1	64	8	8	1	144	8

Table 9.10: Chosen degrees of parallelization for the different lattice sizes.

regard to requirement #3 above, a rather simple scheme with only *two integration levels* is applied. While the gluonic force is integrated on a fine level with step size ϵ_0, all the fermionic forces are integrated uniformly on a coarser level with step size ϵ_1. Actually, since in both cases an *OMF integrator of 4th order* is applied throughout, the fixed choice $N_0 = 1$ already leads to an integration of the gauge field with an effective step size that is five times smaller than the one of the fermion fields (see the discussion in sec. 4.9). The only parameter then left to rule the integration scheme is N_1 (cf. (4.103)).

Parallelization

Our simulations are carried out on lattices of the size $(T/a) \times (L/a)^3$ with (cf. tab. 9.8)

$$T/a \in \{8, 12, 16, 20, 24\} \tag{9.122a}$$

$$\text{and} \qquad L/a \in \{T/a, T/2a\}\ . \tag{9.122b}$$

They can be parallelized in the spatial directions[37], if there exists a divisor $L_{\rm local}/a \geq 4$ of L/a with $L_{\rm local} = L/n$ and $n > 1$. In all but one case we choose an isotropic parallelization, with the respective highest possible degree of parallelization, n^3. Only for $T/a = L/a = 24$, different $L_{\rm local}$ for the different spatial directions are used. Our selection is listed in tab. 9.10.

9.5.2 Practical implementation

Adjustment of algorithmic parameters

According to the discussion in the previous section, a simulation at some point $\vec{x}$ in bare parameter space, irrespective of whether it is a tuning or improvement run, requires the adjustment of the following parameters:

---

[37] Parallelization in the temporal direction is not possible due to the technical complication of a formal additional time slice used in the `openQCD` code.

- Ranges $[r_{f,a}, r_{f,b}]$ of the rational approximations with $f \in \{s, c\}$
- Number of poles n of the rational approximations
- Number of steps N_1 that t'_m is divided into for the integration of the forces F.

To do so, before the actual simulation, we create an ensemble which is relatively small but sufficient to determine the acceptance probability P_{acc} (cf. (9.120)) as well as the extrema of the spectra of $Q_s Q_s^\dagger$ and $Q_c Q_c^\dagger$ (cf. (4.61)). The latter is done under use of the `ms2` program, which is part of the `openQCD` package. It also provides the number of poles n needed to fulfill the condition (4.69). With these informations at hand, the number of poles and the ranges are adapted (the latter under use of a safety margin), while N_1 is adjusted by hand to increase or decrease the acceptance such that it is expected to lie in the aspired range (9.121) according to requirement #2. This way, the parameters are iteratively optimized until the results stabilize. In the process, each run is started from the last configuration of the previous one, which serves the purpose of a pre-thermalization. Afterwards, the actual simulation is started.

Reweighting

The RHMC reweighting factors $W_{f,1}$ (4.68) with $f \in \{s, c\}$ are computed under use of the program `ms1` of the `openQCD` package. The total reweighting factor $W = W_{s,1} \cdot W_{c,1}$ is implemented for each primary observable according to

$$\langle O \rangle = \frac{\langle OW \rangle_{\text{mod}}}{\langle W \rangle_{\text{mod}}} \,. \tag{4.64}$$

A derived observable, like for instance the (unimproved) PCAC mass m^{ud}_{PCAC}, is then obtained by

$$\begin{aligned} m^{ud}_{\text{PCAC}}(x_0) &\overset{(5.51)}{=} \frac{\tilde{\partial}_0 f_A^{ud}(x_0)}{2 f_P^{ud}(x_0)} = \frac{\tilde{\partial}_0 \left\langle A_0^{ud}(x_0)\; \mathcal{O}^{du} \right\rangle}{2 \left\langle P^{ud}(x_0)\; \mathcal{O}^{du} \right\rangle} \\ &= \frac{\left\langle A_0^{ud}(x_0+a)\; \mathcal{O}^{du} \right\rangle - \left\langle A_0^{ud}(x_0-a)\; \mathcal{O}^{du} \right\rangle}{4 \left\langle P^{ud}(x_0)\; \mathcal{O}^{du} \right\rangle} \\ &\overset{(4.64)}{=} \frac{\left\langle A_0^{ud}(x_0+a)\; \mathcal{O}^{du}\; W \right\rangle_{\text{mod}} - \left\langle A_0^{ud}(x_0-a)\; \mathcal{O}^{du}\; W \right\rangle_{\text{mod}}}{4 \left\langle P^{ud}(x_0)\; \mathcal{O}^{du}\; W \right\rangle_{\text{mod}}} \,. \end{aligned} \tag{9.123}$$

Note that the factor $\langle W \rangle_{\text{mod}}$ drops out in ratios of reweighted observables.

Quantity		LCP	IMPR	Ref.	Results
Algorithmic	$e^{-\Delta H}$	yes	yes	(4.53)	
	P_{acc}	yes	yes	(9.120)	
Basic	E_P	yes	yes	(I.7)	app. I.4
	Q_{top}	yes	yes	(6.31)	app. I.4
	$\bar{g}^2_{\mathrm{GF}}$, $\bar{g}^2_{\mathrm{GF},0}$	yes	yes[38]	(6.7), (7.4)	sec. 9.6.5
	SFCF	yes	yes	(5.46)-(5.49)	app. I.3
LCP	Φ_1, Φ_2, Φ_3	yes	no	(9.16), (9.23)	sec. 9.6.1
Masses	m^{ij}_{PCAC}	yes	yes	(9.111a)	sec. 9.6.5
	R_{PCAC}	no	yes	(9.138)	sec. 9.6.5
	c_A	no	yes	(9.112)	app. I.4
	M^{ud}	no	yes	(9.113)	sec. 9.6.2
	ΔM^{ud}	no	yes	(9.115)	sec. 9.6.2

Table 9.11: Main observables obtained from the Monte Carlo simulations in the framework of the tuning (LCP) and the imposition of the improvement condition (IMPR). The definitions of the respective quantities are referred to in the second last column, while the last column directs to the sections where the corresponding results (usually for all T/a and c_{sw}) are discussed.

Analyzed quantities

Several algorithmic and basic quantities are computed by default, while further specific observables are only evaluated either for the tuning or the improvement simulations. An overview is given in tab. 9.11. The parameter S_τ (cf. (4.115)) that enters the error analysis with `UWerr` is adjusted individually for each simulation and observable, and usually lies in the range of $S_\tau \in [2, 10]$. As an example, detailed results of the simulations at[39] $T^\star/a = 24$ and $c_{\mathrm{sw}} = 1.8$ are shown in app. I.1.

Thermalization and replica

In the case of the *tuning* runs, a single simulation with $n_{\mathrm{rep}} = 1$ is sufficient and convenient. This is because the question whether a promising run fulfills the tuning criteria is often decided only as more statistics is accumulated (cf. (9.79b)). A single run offers more flexibility with regard to continuation runs and facilitates the choice of N_{th}.

---

[38] Since the normalization for the gradient flow coupling is not known for non-vanishing boundary conditions (9.119), $t^2 E(t)$ is considered here instead.

[39] We remind the reader that the notational use of $T^\star/a$ indicates simulations at the line of constant physics, while T/a denotes the lattice resolution in general, see footnote 5.

In contrast, for the *improvement* runs, significantly larger ensembles need to be generated to obtain sufficiently small errors, in which case replica are essential. Our procedure involves two separate runs. First, a single *thermalization run* with $n_{\mathrm{rep}} = 1$, where a certain number $N_{\mathrm{th}}^{(\mathrm{th})}$ of configurations is discarded after visual inspection of the histories of the observables listed in tab. 9.11, and in consideration of their autocorrelations. Second, from the remaining ensemble, $n_{\mathrm{rep}} = 8$ configurations of maximal distance to each other in Monte Carlo time are selected. These provide the starting points for the $n_{\mathrm{rep}} = 8$ *replica runs*, from which the main results are derived. If the distance between the start configurations is again large compared to the largest autocorrelation, these can be considered independent, and the replica runs do in principle not need to be thermalized. However, this is checked again, see app. I.2, and depending on whether there is an indication of an incomplete thermalization, a certain amount of configurations $N_{\mathrm{th}}^{(\mathrm{rep})}$ is again discarded.

Finally, note that in all cases, a new simulation of the type $n_{\mathrm{rep}} = 1$ at certain parameters T/a, L/a, c_{sw}, and $\vec{x}$ is preferably started from a configuration of a run with the same T/a, L/a and similar c_{sw} and $\vec{x}$.

9.6 Results

An overview on the created ensembles at the line of constant physics is given in tab. 9.12 and tab. 9.13. The main results of the *tuning*[40] to the LCP (sec. 9.1-9.3) and the imposition of the *improvement*[41] condition (sec. 9.4) are presented in sec. 9.6.1 and sec. 9.6.2, respectively. The outcome for $c_{\mathrm{sw}}(g_0^2)$ for the various lattice sizes is combinedly described by a Padé fit in sec. 9.6.3. In sec. 9.6.4, the results of the aforementioned sections are revisited under consideration of errors due to the unavoidable mismatch of the line of constant physics. This is followed by a check of cutoff effects of a renormalization-independent ratio of PCAC masses and the gradient flow coupling $\overline{g}^2_{\mathrm{GF}}$ in sec. 9.6.5. Finally, in app. I, we investigate the stability of the results (app. I.2), the time dependence of the SFCF (app. I.3), as well as additional results for E_P, Q_{top}, and c_A (app. I.4).

$T^\star/a$	c_{sw}	N_{cf}	P_{acc}	N_1	n_s: $[r_{s,a}, r_{s,b}]$	n_c: $[r_{c,a}, r_{c,b}]$
8	1.9	3500	94.1(1)	7	6: $[0.10, 7.55]$	6: $[0.81, 7.99]$
	2.0	5000	92.9(1)	7	6: $[0.09, 7.93]$	6: $[0.83, 8.34]$
	2.1	3000	91.3(2)	7	6: $[0.08, 7.76]$	6: $[0.69, 8.29]$
	2.2	3000	95.5(1)	8	6: $[0.09, 8.28]$	6: $[0.77, 8.52]$
	2.3	5500	94.0(1)	8	6: $[0.08, 8.28]$	6: $[0.72, 8.52]$
	2.4	5000	92.2(2)	8	6: $[0.09, 8.28]$	6: $[0.74, 8.52]$

$T^\star/a$	c_{sw}	N_{cf}	P_{acc}	N_1	n_s: $[r_{s,a}, r_{s,b}]$	n_c: $[r_{c,a}, r_{c,b}]$
12	1.8	1400	96.2(2)	9	6: $[0.07, 7.26]$	6: $[0.46, 7.46]$
	1.9	1400	94.1(3)	9	6: $[0.07, 7.60]$	6: $[0.43, 7.61]$
	2.0	2100	92.2(3)	9	6: $[0.06, 7.88]$	6: $[0.40, 8.08]$
	2.1	1300	89.0(5)	9	6: $[0.06, 7.88]$	6: $[0.40, 8.08]$
	2.2	2000	87.0(4)	9	6: $[0.08, 7.88]$	6: $[0.40, 8.08]$

$T^\star/a$	c_{sw}	N_{cf}	P_{acc}	N_1	n_s: $[r_{s,a}, r_{s,b}]$	n_c: $[r_{c,a}, r_{c,b}]$
16	1.5	2300	89.2(5)	9	6: $[0.06, 6.90]$	6: $[0.32, 7.00]$
	1.6	2500	87.1(4)	9	6: $[0.05, 6.90]$	6: $[0.32, 7.00]$
	1.7	1400	84.8(5)	9	6: $[0.05, 7.40]$	6: $[0.33, 7.57]$
	1.8	1400	87.7(5)	10	8: $[0.05, 7.24]$	6: $[0.31, 7.32]$
	1.9	1500	84.9(5)	10	8: $[0.04, 7.48]$	6: $[0.29, 7.58]$
	2.0	1800	87.9(5)	11	8: $[0.05, 8.03]$	6: $[0.30, 7.74]$

$T^\star/a$	c_{sw}	N_{cf}	P_{acc}	N_1	n_s: $[r_{s,a}, r_{s,b}]$	n_c: $[r_{c,a}, r_{c,b}]$
20	1.5	1300	91.7(3)	8	8: $[0.05, 6.74]$	6: $[0.26, 6.85]$
	1.6	1200	96.2(2)	9	8: $[0.04, 6.90]$	6: $[0.26, 7.04]$
	1.7	1300	95.4(2)	9	8: $[0.05, 7.06]$	6: $[0.25, 7.16]$
	1.8	1500	93.9(2)	9	8: $[0.04, 7.06]$	6: $[0.25, 7.16]$
	1.9	1500	92.4(3)	9	8: $[0.04, 7.06]$	6: $[0.24, 7.16]$

$T^\star/a$	c_{sw}	N_{cf}	P_{acc}	N_1	n_s: $[r_{s,a}, r_{s,b}]$	n_c: $[r_{c,a}, r_{c,b}]$
24	1.5	1500	95.7(2)	9	8: $[0.04, 6.94]$	6: $[0.19, 7.02]$
	1.6	2000	94.1(2)	9	8: $[0.03, 6.94]$	6: $[0.19, 7.02]$
	1.7	2000	93.2(3)	9	8: $[0.03, 7.00]$	6: $[0.19, 7.10]$
	1.8	2000	96.5(1)	10	8: $[0.03, 7.00]$	6: $[0.19, 7.10]$

Table 9.12: Overview on the *tuning* ensembles at the line of constant physics. $N_{\mathrm{cf}} = n_{\mathrm{rep}} \times N_{\mathrm{cf}}/\mathrm{replicum}$ gives the number of configurations, with the distance t'_m between two consecutive configurations being 2 MDU in all cases. P_{acc} is the acceptance (9.121), N_1 the integration step size (4.103), and n_f: $[r_{f,a}, r_{f,b}]$ denotes the number of poles and the ranges of the rational approximations (sec. 4.6) that were used in the simulations for the flavors $f \in \{s, c\}$. These may include a safety margin with respect to the actually measured spectra. The bare parameters that correspond to the ensembles can be found in tab. 9.14 and tab. 9.15.

$T^\star/a$	$c_{\rm sw}$	$N_{\rm cf}$	P_{acc}	N_1	n_s: $[r_{s,a}, r_{s,b}]$	n_c: $[r_{c,a}, r_{c,b}]$
8	1.9	8×4000	96.13(3)	6	6: [0.22, 7.51]	6: [0.94, 7.87]
	2.0		95.55(3)	6	6: [0.10, 8.10]	6: [0.80, 8.40]
	2.1		94.81(4)	6	6: [0.08, 8.15]	6: [0.69, 8.51]
	2.2		94.04(5)	6	6: [0.10, 8.28]	6: [0.77, 8.52]
	2.3		93.11(5)	6	6: [0.11, 8.56]	6: [0.81, 8.98]
	2.4		91.90(6)	6	6: [0.12, 8.80]	6: [0.81, 9.00]

$T^\star/a$	$c_{\rm sw}$	$N_{\rm cf}$	P_{acc}	N_1	n_s: $[r_{s,a}, r_{s,b}]$	n_c: $[r_{c,a}, r_{c,b}]$
12	1.8	8×4000	91.69(6)	6	6: [0.07, 7.45]	6: [0.50, 7.70]
	1.9		98.77(1)	8	6: [0.07, 7.60]	6: [0.43, 7.70]
	2.0		99.41(1)	9	6: [0.06, 7.88]	6: [0.40, 8.08]
	2.1		95.03(4)	7	8: [0.06, 8.33]	6: [0.44, 8.52]
	2.2		94.09(5)	7	6: [0.08, 8.64]	6: [0.40, 8.87]

$T^\star/a$	$c_{\rm sw}$	$N_{\rm cf}$	P_{acc}	N_1	n_s: $[r_{s,a}, r_{s,b}]$	n_c: $[r_{c,a}, r_{c,b}]$
16	1.5	8×4000	98.90(1)	8	6: [0.06, 6.90]	6: [0.32, 7.00]
	1.6		99.52(1)	9	6: [0.06, 6.90]	6: [0.32, 7.00]
	1.7		98.40(2)	8	8: [0.06, 7.42]	6: [0.31, 7.59]
	1.8		98.03(2)	8	8: [0.06, 7.49]	6: [0.31, 7.64]
	1.9		99.10(1)	9	8: [0.04, 7.54]	6: [0.29, 7.66]
	2.0		99.48(1)	10	8: [0.05, 8.03]	6: [0.30, 7.74]

$T^\star/a$	$c_{\rm sw}$	$N_{\rm cf}$	P_{acc}	N_1	n_s: $[r_{s,a}, r_{s,b}]$	n_c: $[r_{c,a}, r_{c,b}]$
20	1.5	8×3000	97.19(3)	8	8: [0.05, 6.74]	6: [0.26, 6.85]
	1.6		96.51(3)	8	8: [0.06, 7.10]	6: [0.28, 7.20]
	1.7		98.42(1)	9	8: [0.05, 7.20]	6: [0.24, 7.30]
	1.8		98.08(2)	9	8: [0.05, 7.10]	6: [0.24, 7.20]
	1.9		93.59(6)	8	8: [0.05, 7.50]	6: [0.24, 7.60]

$T^\star/a$	$c_{\rm sw}$	$N_{\rm cf}$	P_{acc}	N_1	n_s: $[r_{s,a}, r_{s,b}]$	n_c: $[r_{c,a}, r_{c,b}]$
24	1.5	8×4000	98.61(1)	9	6: [0.05, 7.00]	6: [0.22, 7.10]
	1.6		98.26(1)	9	8: [0.05, 7.00]	6: [0.20, 7.10]
	1.7		94.06(4)	8	8: [0.05, 7.15]	6: [0.20, 7.20]
	1.8		97.40(2)	9	8: [0.05, 7.50]	6: [0.20, 7.60]

Table 9.13: Overview on the *improvement* ensembles at the line of constant physics. $N_{\rm cf} = n_{\rm rep} \times N_{\rm cf}/\text{replicum}$ gives the number of configurations, with the distance t'_m between two consecutive configurations being 2 MDU in all cases. P_{acc} is the acceptance (9.121), N_1 the integration step size (4.103), and n_f: $[r_{f,a}, r_{f,b}]$ denotes the number of poles and the ranges of the rational approximations (sec. 4.6) that were used in the simulations for the flavors $f \in \{s, c\}$. These may include a safety margin with respect to the actually measured spectra. The bare parameters that correspond to the ensembles can be found in tab. 9.14 and tab. 9.15.

c_{sw}	$g^2_{0,\mathrm{L}}$	$\kappa_{l,\mathrm{L}}$	$\kappa_{c,\mathrm{L}}$	Φ_1	Φ_2	Φ_3
1.9	1.7848	0.13741	0.12033	7.17(4)	0.59(3)	5.95(1)
2.0	1.8067	0.13645	0.11960	7.17(3)	0.56(2)	5.99(1)
2.1	1.8340	0.13560	0.11900	7.36(5)	0.61(3)	6.02(1)
2.2	1.8534	0.13450	0.11820	7.17(5)	0.60(3)	6.03(1)
2.3	1.8825	0.13369	0.11770	7.36(3)	0.62(2)	5.97(1)
2.4	1.9069	0.13278	0.11750	7.31(4)	0.57(3)	5.93(1)

$(T^\star/a = 8)$

c_{sw}	$g^2_{0,\mathrm{L}}$	$\kappa_{l,\mathrm{L}}$	$\kappa_{c,\mathrm{L}}$	Φ_1	Φ_2	Φ_3
1.8	1.7334	0.13742	0.12776	7.13(5)	0.58(3)	5.93(2)
1.9	1.7600	0.13656	0.12714	7.26(4)	0.60(3)	5.92(2)
2.0	1.7876	0.13573	0.12653	7.35(5)	0.62(3)	5.99(3)
2.1	1.8152	0.13486	0.12591	7.29(6)	0.60(4)	5.93(3)
2.2	1.8423	0.13401	0.12529	7.42(7)	0.60(4)	5.91(3)

$(T^\star/a = 12)$

c_{sw}	$g^2_{0,\mathrm{L}}$	$\kappa_{l,\mathrm{L}}$	$\kappa_{c,\mathrm{L}}$	Φ_1	Φ_2	Φ_3
1.5	1.6266	0.13918	0.13224	7.38(9)	0.56(2)	5.93(3)
1.6	1.6482	0.13825	0.13139	7.42(10)	0.61(2)	6.06(7)
1.7	1.6741	0.13741	0.13087	7.29(8)	0.59(2)	5.84(5)
1.8	1.7036	0.13661	0.13018	7.32(7)	0.57(2)	5.90(5)
1.9	1.7331	0.13582	0.12949	7.44(9)	0.59(2)	5.94(3)
2.0	1.7626	0.13503	0.12880	7.46(7)	0.59(2)	6.03(4)

$(T^\star/a = 16)$

c_{sw}	$g^2_{0,\mathrm{L}}$	$\kappa_{l,\mathrm{L}}$	$\kappa_{c,\mathrm{L}}$	Φ_1	Φ_2	Φ_3
1.5	1.5868	0.13822	0.13287	7.20(7)	0.61(2)	5.91(8)
1.6	1.6153	0.13749	0.13240	7.30(10)	0.57(2)	5.91(5)
1.7	1.6434	0.13669	0.13178	7.29(7)	0.60(2)	5.85(5)
1.8	1.6714	0.13588	0.13126	7.49(5)	0.63(2)	5.81(6)
1.9	1.6993	0.13504	0.13046	7.36(9)	0.62(2)	5.81(8)

$(T^\star/a = 20)$

c_{sw}	$g^2_{0,\mathrm{L}}$	$\kappa_{l,\mathrm{L}}$	$\kappa_{c,\mathrm{L}}$	Φ_1	Φ_2	Φ_3
1.5	1.5527	0.13751	0.13325	7.19(6)	0.62(2)	6.02(8)
1.6	1.5783	0.13672	0.13253	7.23(11)	0.59(1)	6.02(6)
1.7	1.6065	0.13595	0.13202	7.24(9)	0.62(2)	6.00(5)
1.8	1.6344	0.13515	0.13132	7.11(7)	0.56(1)	5.87(6)

$(T^\star/a = 24)$

Table 9.14: Results of the tuning to the line of constant physics. The bare parameters are labelled with a subscript L.

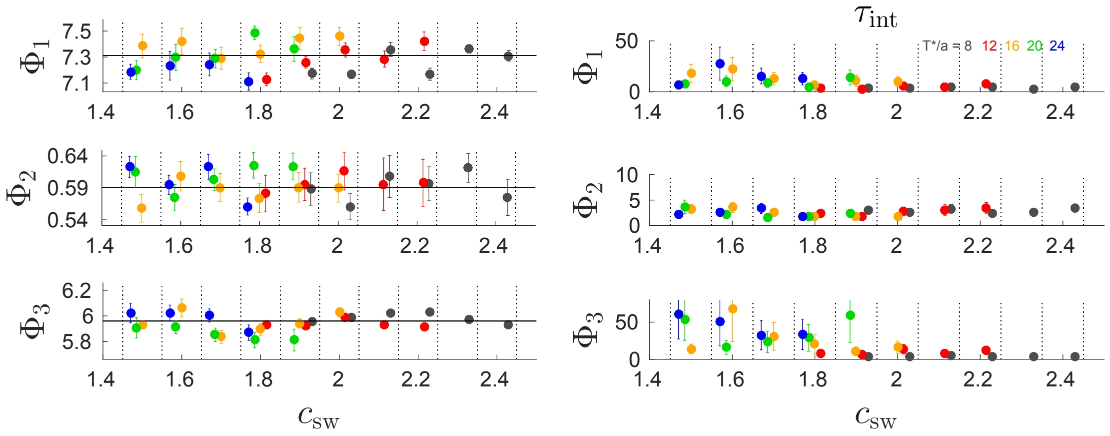


Figure 9.9: Accuracy of the tuning to the line of constant physics. The observables Φ_i (see also tab. 9.14) and their autocorrelations are shown in the *left* and *right panel*, respectively. The points are shifted horizontally depending on $T^\star/a$, but in all cases correspond to the respective central values of the intervals separated by the dotted vertical lines.

9.6.1 Tuning to the line of constant physics

The tuning results, i.e. the bare parameters $\vec{x}_{\mathrm{L}} = (g^2_{0,\mathrm{L}}, \kappa_{l,\mathrm{L}}, \kappa_{c,\mathrm{L}})$ at the line of constant physics and their associated values Φ_i are given in tab. 9.14. The achieved accuracy of the tuning is illustrated in the left panel of fig. 9.9, where the ranges of Φ_i covered by the plots correspond to the allowed deviations from the LCP, respectively. The fulfillment of the tuning criteria is thus reflected by the fact that all data points with their respective error bars fit into the plot range. The dependence of the LCP bare parameters on c_{sw}, depicted in fig. 9.10, may well be described by linear fits whose slope parameters are listed in the same figure.

Note that these fits[42] were employed to estimate initial guesses $\vec{x}^{\mathrm{initial}}$ for the tuning procedure at a new c_{sw} and given T/a, cf. step S0 in the iterative tuning approach in sec. 9.3. Fig. 9.11 shows an example of a second type of interpolation[43] of the LCP bare parameters, in $T^\star/a$ at given c_{sw}, which has been made use of as well.

---

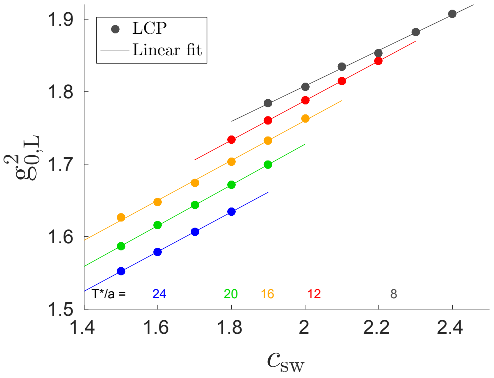
[40] See footnote 32.

[41] See footnote 32.

[42] Actually, precursors of those fits, with preliminary data.

[43] See footnote 42.

$T^\star/a$	$\frac{d}{dc_{\mathrm{sw}}}\, g^2_{0,\mathrm{L}}$	$\frac{d}{dc_{\mathrm{sw}}}\, \kappa_{l,\mathrm{L}}$	$\frac{d}{dc_{\mathrm{sw}}}\, \kappa_{c,\mathrm{L}}$
8	0.2449	−0.00929	−0.00590
12	0.2730	−0.00852	−0.00617
16	0.2755	−0.00825	−0.00674
20	0.2811	−0.00797	−0.00596
24	0.2733	−0.00785	−0.00630



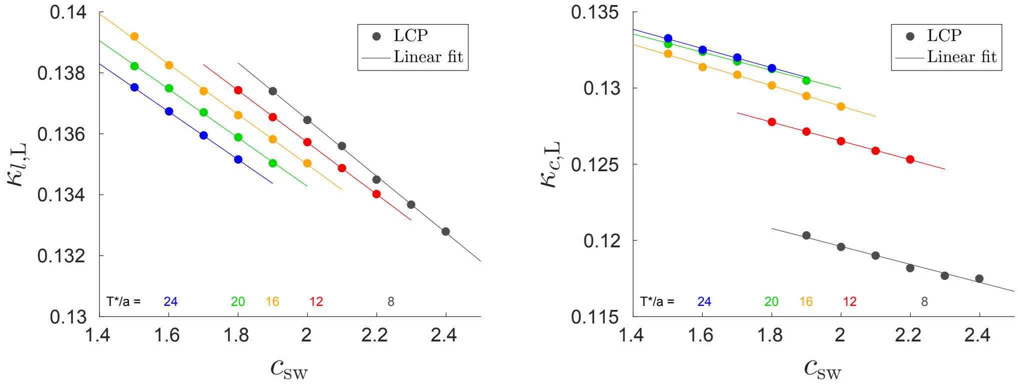


Figure 9.10: Results for the bare parameters tuned to the line of constant physics. *Top left:* Linear slope parameters of the bare parameters at the LCP as a function of c_{sw}. *Top right:* Bare coupling $g^2_{0,\mathrm{L}}$ at the LCP. *Bottom left:* Light hopping parameter $\kappa_{l,\mathrm{L}}$ at the LCP. *Bottom right:* Charm hopping parameter $\kappa_{c,\mathrm{L}}$ at the LCP.

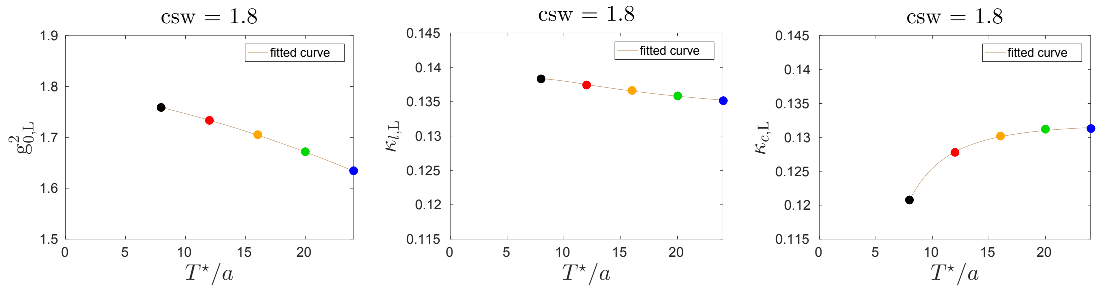


Figure 9.11: Dependences of $g^2_{0,\mathrm{L}}, \kappa_{l,\mathrm{L}}, \kappa_{c,\mathrm{L}}$ on $T^\star/a$ and the respective polynomial interpolations to for $c_{\mathrm{sw}} = 1.8$.

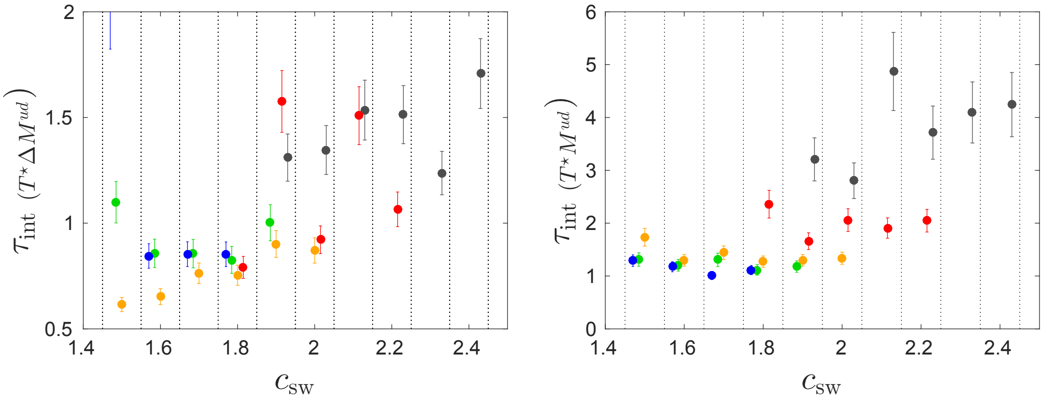


Figure 9.12: Autocorrelations of $T^\star \Delta M^{ud}(\frac{3}{4}T^\star, \frac{1}{4}T^\star)$ *(left)* and $T^\star M^{ud}(T^\star/2, T^\star/4)$ *(right)* in the improvement simulations. The points are shifted horizontally, cf. fig. 9.9.

9.6.2 Improvement condition: Individual and global linear fits

The results for $T^\star \Delta M^{ud}$, in terms of which the improvement condition (9.115) is formulated along the LCP, as well as additional quantities of interest are shown in tab. 9.15, while the observed autocorrelations of $T^\star \Delta M^{ud}$ are illustrated in fig. 9.12.

Two different kinds of fits for $T^\star \Delta M^{ud}$ as a function of $c_{\rm sw}$ and $T^\star/a$ are performed. First, the *individual* fits,

$$T^\star \Delta M^{ud}(c_{\rm sw}, T^\star/a) = s^{(T^\star/a)} \quad \cdot \left[c_{\rm sw} - c_{\rm sw,I}^{(T^\star/a)}\right] \ , \tag{9.124a}$$

which consider each $T^\star/a$ independently by a fit linear in $c_{\rm sw}$, with a total of 10 fit parameters. Second, the *global* fit, where only a single slope parameter s appears as one of 6 fit parameters:

$$T^\star \Delta M^{ud}(c_{\rm sw}, T^\star/a) = s \cdot \left(\frac{a}{T^\star}\right) \cdot \left[c_{\rm sw} - c_{\rm sw,I}^{(T^\star/a)}\right] \ . \tag{9.124b}$$

The respective fits are depicted in fig. 9.13.

For each lattice size $T^\star/a$, the fit parameter $c_{\rm sw,I} \equiv c_{\rm sw,I}^{(T^\star/a)}$ directly provides the improvement coefficient at which the improvement condition $T^\star \Delta M^{ud} = 0$ is fulfilled[44]. Subsequent interpolations of the bare parameters to the respective $c_{\rm sw,I}$, under use of the linear fits shown in fig. 9.10, lead to the corresponding improvement point bare

---

[44] Note that this is not the case for (otherwise equivalent) linear fits of e.g. the form $T^\star \Delta M^{ud}(c_{\rm sw}, T^\star/a) = a^{(T^\star/a)} + b^{(T^\star/a)} \cdot c_{\rm sw}$.

$(T^\star/a = 8)$

c_{sw}	$g^2_{0,\mathrm{L}}$	$\kappa_{l,\mathrm{L}}$	$\kappa_{c,\mathrm{L}}$	$t^2E(t)$ $[10^{-2}]$	$T^\star m^{ud}_{\mathrm{PCAC}}$	$T^\star M^{ud}$	$T^\star \Delta M^{ud}$
1.9	1.7848	0.13741	0.12033	3.831(4)	0.504(7)	0.684(25)	0.047(7)
2.0	1.8067	0.13645	0.11960	3.851(4)	0.491(8)	0.621(23)	0.028(7)
2.1	1.8340	0.13560	0.11900	3.895(5)	0.509(8)	0.707(32)	0.021(8)
2.2	1.8534	0.13450	0.11820	3.898(5)	0.516(8)	0.643(28)	-0.014(8)
2.3	1.8825	0.13369	0.11770	3.945(6)	0.521(9)	0.724(30)	-0.013(8)
2.4	1.9069	0.13278	0.11750	3.963(6)	0.521(10)	0.723(31)	-0.016(9)

$(T^\star/a = 12)$

c_{sw}	$g^2_{0,\mathrm{L}}$	$\kappa_{l,\mathrm{L}}$	$\kappa_{c,\mathrm{L}}$	$t^2E(t)$ $[10^{-2}]$	$T^\star m^{ud}_{\mathrm{PCAC}}$	$T^\star M^{ud}$	$T^\star \Delta M^{ud}$
1.8	1.7334	0.13742	0.12776	7.474(8)	0.247(4)	0.311(11)	0.032(4)
1.9	1.7600	0.13656	0.12714	7.576(8)	0.242(4)	0.306(9)	0.012(7)
2.0	1.7876	0.13573	0.12653	7.699(9)	0.237(4)	0.300(11)	0.020(5)
2.1	1.8152	0.13486	0.12591	7.837(10)	0.243(5)	0.300(10)	-0.004(7)
2.2	1.8423	0.13401	0.12529	7.898(10)	0.223(4)	0.267(11)	-0.011(6)

$(T^\star/a = 16)$

c_{sw}	$g^2_{0,\mathrm{L}}$	$\kappa_{l,\mathrm{L}}$	$\kappa_{c,\mathrm{L}}$	$t^2E(t)$ $[10^{-2}]$	$T^\star m^{ud}_{\mathrm{PCAC}}$	$T^\star M^{ud}$	$T^\star \Delta M^{ud}$
1.5	1.6266	0.13918	0.13224	8.187(9)	0.160(3)	0.192(6)	0.026(3)
1.6	1.6482	0.13825	0.13139	8.213(9)	0.159(2)	0.164(5)	0.018(4)
1.7	1.6741	0.13741	0.13087	8.315(9)	0.156(2)	0.139(6)	0.014(3)
1.8	1.7036	0.13661	0.13018	8.466(9)	0.167(2)	0.144(6)	0.004(3)
1.9	1.7331	0.13582	0.12949	8.607(10)	0.165(3)	0.136(6)	-0.008(3)
2.0	1.7626	0.13503	0.12880	8.735(11)	0.154(2)	0.113(6)	-0.015(4)

$(T^\star/a = 20)$

c_{sw}	$g^2_{0,\mathrm{L}}$	$\kappa_{l,\mathrm{L}}$	$\kappa_{c,\mathrm{L}}$	$t^2E(t)$ $[10^{-2}]$	$T^\star m^{ud}_{\mathrm{PCAC}}$	$T^\star M^{ud}$	$T^\star \Delta M^{ud}$
1.5	1.5868	0.13822	0.13287	8.365(13)	0.150(2)	0.153(5)	0.018(4)
1.6	1.6153	0.13749	0.13240	8.461(14)	0.131(2)	0.135(5)	0.006(3)
1.7	1.6434	0.13669	0.13178	8.569(12)	0.139(2)	0.117(5)	-0.000(3)
1.8	1.6714	0.13588	0.13126	8.610(13)	0.132(3)	0.097(5)	-0.003(3)
1.9	1.6993	0.13504	0.13046	8.703(16)	0.140(2)	0.093(5)	-0.007(5)

$(T^\star/a = 24)$

c_{sw}	$g^2_{0,\mathrm{L}}$	$\kappa_{l,\mathrm{L}}$	$\kappa_{c,\mathrm{L}}$	$t^2E(t)$ $[10^{-2}]$	$T^\star m^{ud}_{\mathrm{PCAC}}$	$T^\star M^{ud}$	$T^\star \Delta M^{ud}$
1.5	1.5527	0.13751	0.13325	8.353(13)	0.129(2)	0.135(4)	0.007(4)
1.6	1.5783	0.13672	0.13253	8.405(13)	0.120(1)	0.109(3)	0.011(2)
1.7	1.6065	0.13595	0.13202	8.463(12)	0.119(1)	0.092(3)	-0.004(2)
1.8	1.6344	0.13515	0.13132	8.519(13)	0.113(2)	0.076(4)	-0.007(2)

Table 9.15: Results of the improvement condition runs along the LCP.

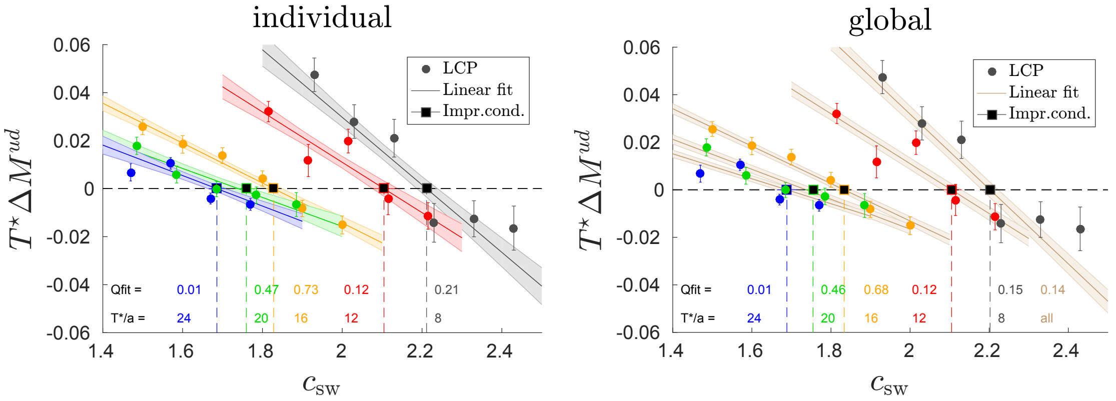


Figure 9.13: Main results for $T^\star \Delta M^{ud}$ of the improvement runs, as listed in tab. 9.15. The *left panel* shows the individual linear fits (9.124a), whereas the global linear fit (9.124b) is displayed in the *right panel*.

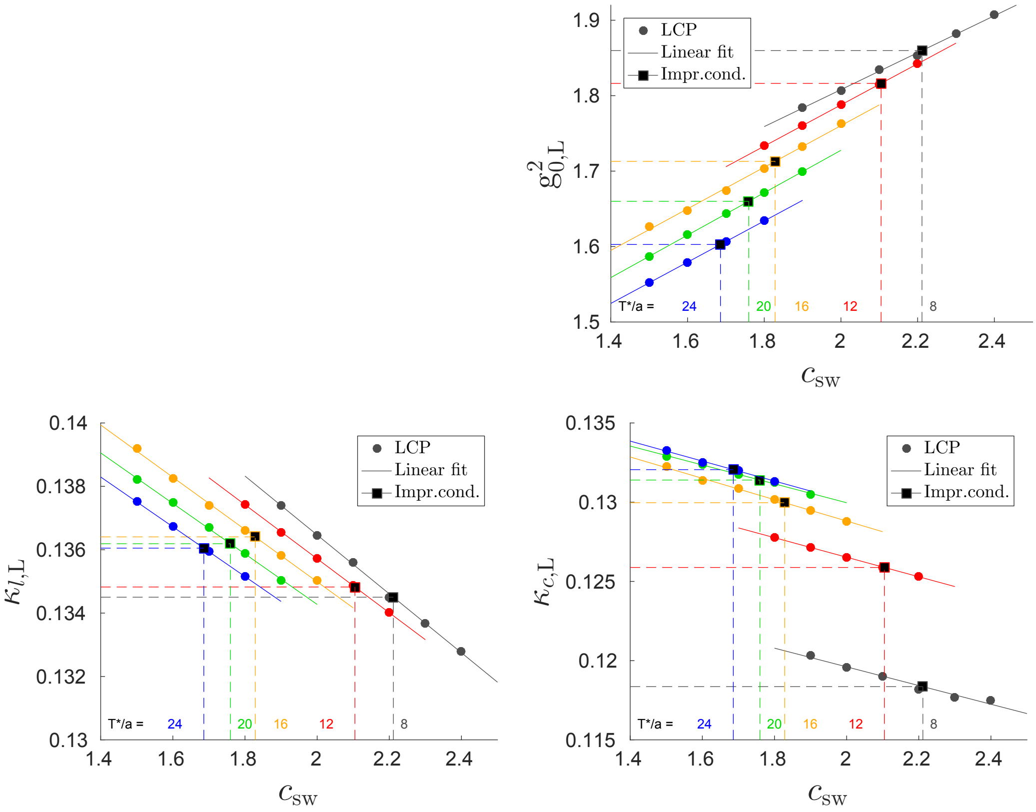


Figure 9.14: Interpolation of the bare parameters $g^2_{0,\rm L}, \kappa_{l,\rm L}, \kappa_{c,\rm L}$ at the LCP to the values $g^2_{0,\rm I}, \kappa_{l,\rm I}, \kappa_{c,\rm I}$ determined by the improvement coefficients $c_{\rm sw,I}$ in the case of the *global* fit. Compare to fig. 9.10

$T^\star/a$	$c_{\mathrm{sw,I}}$	Q_{fit}	$g^2_{0,\mathrm{I}}$	$\kappa_{l,\mathrm{I}}$	$\kappa_{c,\mathrm{I}}$
8	2.211(25)	0.21	1.8597	0.13450	0.11836
12	2.105(30)	0.12	1.8163	0.13482	0.12588
16	1.828(19)	0.73	1.7129	0.13640	0.12997
20	1.760(34)	0.47	1.6600	0.13619	0.13140
24	1.686(20)	0.01	1.6028	0.13605	0.13205

(individual)

$T^\star/a$	$c_{\mathrm{sw,I}}$	Q_{fit}	$g^2_{0,\mathrm{I}}$	$\kappa_{l,\mathrm{I}}$	$\kappa_{c,\mathrm{I}}$
8	2.202(21)	0.15	1.8575	0.13459	0.11841
12	2.105(24)	0.12	1.8163	0.13482	0.12588
16	1.833(19)	0.68	1.7144	0.13636	0.12993
20	1.755(28)	0.46	1.6586	0.13623	0.13143
24	1.688(24)	0.01	1.6035	0.13603	0.13204
		0.14			

(global)

Table 9.16: Results of the respective interpolations to $T^\star \Delta M^{ud} = 0$ for the *individual (top)* and *global (bottom)* fit. While the actual Q_{fit} of the *global* fit is given in the last row of the associated table, the other values of Q_{fit} were formally computed from the χ^2 contribution of the corresponding $T^\star/a$ data and $n_{\mathrm{dof}} = 2$ degrees of freedom for a linear fit. The difference to the Q_{fit} of the corresponding *individual* $T^\star/a$ linear fit may serve as a measure of how much the description of the *individual* $T^\star/a$ data decreases in quality if the *global* fit is used. The covariance matrices of the fit parameters can be found in tab. H.1, see app. H.3.

parameters $g^2_{0,\mathrm{I}}, \kappa_{l,\mathrm{I}}, \kappa_{c,\mathrm{I}}$, via e.g.

$$g^2_{0,\mathrm{I}} = g^2_{0,\mathrm{L}}\big|_{c_{\mathrm{sw}}=0} + \frac{dg^2_{0,\mathrm{L}}}{dc_{\mathrm{sw}}} \cdot c_{\mathrm{sw,I}} \,. \tag{9.125}$$

This is displayed for the case of the *global* fit in fig. 9.14. The overall results of the imposition of the improvement condition are compiled in tab. 9.16.

We find that the *individual* and *global* fits lead to quite similar results. We continue to consider both cases, and will use the more restrictive *global* fit in the end. It is particularly valuable with regard to $T^\star/a = 24$, where data is available to a lesser extent, see also app. I.2.

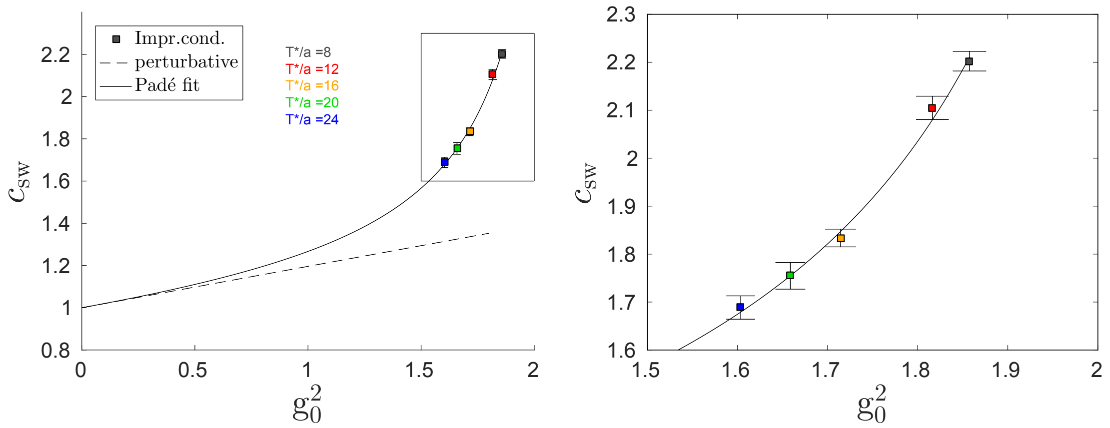


Figure 9.15: *Left:* Padé fit (cf. (9.129)) of the results from the imposition of the improvement condition, which were obtained under use of the *global* fit (cf. right panel of fig. 9.13). *Right:* Magnification of the rectangle area shown on the left.

9.6.3 Padé fit of $c_{\mathrm{sw}}(g_0^2)$

The resulting combinations[45] $c_{\mathrm{sw,I}}(g_{0,\mathrm{I}}^2)$ of both the *individual* and *global* linear fits (cf. tab. 9.16) are now to be interpolated. Apart from describing the data well, the fit function should incorporate the perturbative behavior of $c_{\mathrm{sw,I}}$, given in (3.113a) and (3.113c). To this end, we employ a Padé fit of the form

$$c_{\mathrm{sw,I}}(g_{0,\mathrm{I}}^2) = \frac{1 + a g_{0,\mathrm{I}}^2 + b g_{0,\mathrm{I}}^4 + c g_{0,\mathrm{I}}^6}{1 + (a - 0.196) g_{0,\mathrm{I}}^2} \; . \tag{9.126}$$

The *global* fit results lead to

$$a = -0.254(10) \; , \quad b = -0.055(6) \; , \quad c = +0.004(5) \; , \tag{9.127}$$

with again very similar results for the *individual* fits. As the coefficient c is compatible with zero, it is sufficient to use a Padé fit of second order, which gives

$$a = \; -0.257(7) \; , \quad b = -0.050(4) \; . \tag{9.128}$$

---

[45] We stick to the *non-standard* notation with the subscripts I for the moment, for clarity and consistency with the previous discussion.

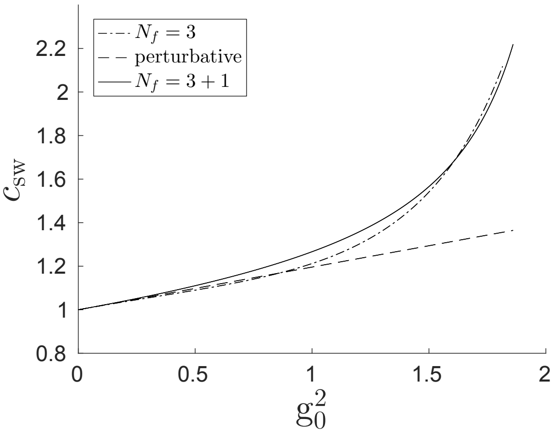


Figure 9.16: Comparison of our result (9.129) in $N_f = 3+1$ (solid line) to the non-perturbative determination of $c_{\mathrm{sw}}(g_0^2)$ in $N_f = 3$ [22] (dash-dot line), cf. tab. F.1 in app. F. The dashed line represents the perturbative 1-loop formula for the Lüscher–Weisz gauge action, see (3.113c).

The fit quality amounts to $Q_{\mathrm{fit}} = 0.59$, and the covariance matrix of the fit parameters is listed in tab. H.1, see app. H.3. We hence quote the main result of our work[46],

$$c_{\mathrm{sw}}(g_0^2) = \frac{1 - 0.257 g_0^2 - 0.050 g_0^4}{1 - 0.453 g_0^2} \qquad \text{for } \ 0 \leq g_0^2 \leq 1.86\ . \tag{9.129}$$

It is illustrated in fig. 9.15, and compared to the previously obtained $N_f = 3$ result for $c_{\mathrm{sw}}(g_0^2)$ in conjunction with the Lüscher–Weisz gauge action in fig. 9.16.

9.6.4 Errors from deviations from the LCP

So far, we have neglected the fact that the line of constant physics was hit only up to a certain precision, defined by the tuning criteria (9.79). We will now semi-quantitatively discuss how this mismatch affects the previously obtained results on $c_{\mathrm{sw}}(g_0^2)$.

The fact that the bare parameters $\vec{x}_{\mathrm{L}} = (g_{0,\mathrm{L}}^2, \kappa_{l,\mathrm{L}}, \kappa_{c,\mathrm{L}})$ only approximately correspond to the LCP is expressed in the occurrence of errors[47] $\tilde{\Delta}\vec{x}_{\mathrm{L}} = (\tilde{\Delta} g_{0,\mathrm{L}}^2, \tilde{\Delta}\kappa_{l,\mathrm{L}}, \tilde{\Delta}\kappa_{c,\mathrm{L}})$, that need to be propagated to errors $\tilde{\Delta} c_{\mathrm{sw,I}}$ in addition to the statistical ones, $\Delta c_{\mathrm{sw,I}}$, which stem from the imposition of the improvement condition. We restrict the discussion to $\tilde{\Delta} g_{0,\mathrm{L}}^2$. In principle, it would be desirable to derive it from the first hand principles described in sec. 9.3, i.e. as the error of $\vec{x}_{\mathrm{L}} = \vec{x}^{\mathrm{min}}$ in the last iteration of the tuning procedure. However, not only is this out of reach, as we have discussed, but the order of magnitude of $\tilde{\Delta} g_{0,\mathrm{L}}^2$ can also be estimated from the following considerations. The dependence of $g_{0,\mathrm{L}}^2$ on c_{sw} is obviously quite well compatible with a linear behavior, see fig. 9.10. If we consider the deviations from the fit function statistical, and

---

[46] We now switch to *standard* notation, $c_{\mathrm{sw,I}} \to c_{\mathrm{sw}}$ and $g_{0,\mathrm{I}}^2 \to g_0^2$.

[47] We denote the errors by $\tilde{\Delta}$ to differentiate them from previously occurring statistical errors, like $\Delta c_{\mathrm{sw,I}}$.

$T^\star/a$	$g^2_{0,\mathrm{I}}$	$\tilde{\Delta} g^2_{0,\mathrm{L}}$	$\tilde{\Delta} g^2_{0,\mathrm{I}}$
8	1.8575	0.0023	0.0010
12	1.8163	0.0003	0.0002
16	1.7144	0.0032	0.0014
20	1.6586	0.0002	0.0001
24	1.6035	0.0009	0.0005

Table 9.17: Estimates for uniform errors $\tilde{\Delta} g^2_{0,\mathrm{L}}$ from setting $\chi^2/n_{\mathrm{dof}} = 1$ with regard to the linear fits shown in fig. 9.10. The errors are transferred to $g^2_{0,\mathrm{I}}$, which is obtained by interpolation to $c_{\mathrm{sw,I}}$ (cf. (9.132)). While the displayed central values $g^2_{0,\mathrm{I}}$ correspond to the *global* case (cf. tab. 9.16), the errors $\tilde{\Delta} g^2_{0,\mathrm{I}}$ are the same for both fits.

assume $\chi^2/n_{\mathrm{dof}} = 1$ (cf. app. H.2), this determines the scale of uniform errors $\tilde{\Delta} g^2_{0,\mathrm{L}}$. The results of this estimation can be found in tab. 9.17. In the following, for some fixed $T^\star/a$, we label the data at the different investigated values of c_{sw} in ascending order by a subscript (j) with $j = 1, \ldots, N_{c_{\mathrm{sw}}}$. For instance, at $T^\star/a = 16$, we have $N_{c_{\mathrm{sw}}} = 6$, with $j = 4$ corresponding to $c_{\mathrm{sw}} = 1.8$, i.e. $g^2_{0,\mathrm{L},(4)} = 1.7036$ and $T^\star \Delta M^{ud}_{(4)} = 0.004$, cf. tab. 9.15. The whole set of bare couplings at the LCP is then denoted by

$$\{g^2_{0,\mathrm{L}}\} = \{g^2_{0,\mathrm{L},(1)}, \ldots, g^2_{0,\mathrm{L},(j)}, \ldots, g^2_{0,\mathrm{L},(N_{c_{\mathrm{sw}}})}\} \,. \tag{9.130}$$

The errors $\tilde{\Delta} g^2_{0,\mathrm{L}}$ now propagate to the results $c_{\mathrm{sw}}(g_0^2)$ in a twofold way. First, the imposition of the improvement condition, (9.124), depends on the bare couplings $\{g^2_{0,\mathrm{L}}\}$ at the various c_{sw} at which $T^\star \Delta M^{ud}$ is determined,

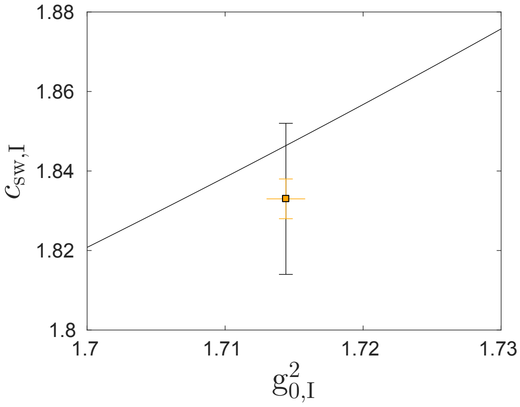
$$c_{\mathrm{sw,I}} \equiv c_{\mathrm{sw,I}}\left(\{g^2_{0,\mathrm{L}}\}\right) \,. \tag{9.131}$$

Second, the interpolation of $g^2_{0,\mathrm{L}}$ to $g^2_{0,\mathrm{I}}$, see (9.125) and fig. 9.14, depends on $\{g^2_{0,\mathrm{L}}\}$ via both the linear fit parameters, and again $c_{\mathrm{sw,I}}$:

$$g^2_{0,\mathrm{I}} = g^2_{0,\mathrm{L}}\big|_{c_{\mathrm{sw}}=0}\left(\{g^2_{0,\mathrm{L}}\}\right) + \frac{dg^2_{0,\mathrm{L}}}{dc_{\mathrm{sw}}}\left(\{g^2_{0,\mathrm{L}}\}\right) \cdot c_{\mathrm{sw,I}}\left(\{g^2_{0,\mathrm{L}}\}\right) \,. \tag{9.132}$$

Consequently, the mismatch of the coupling, $\tilde{\Delta} g^2_{0,\mathrm{L}}$, expresses itself in additional errors $\tilde{\Delta} c_{\mathrm{sw,I}}$ and $\tilde{\Delta} g^2_{0,\mathrm{I}}$, depicted in fig. 9.17. There exist correlations between the two due to (9.132), but it is presumably sufficient to consider them independently to estimate the size of their effects on the total error of $c_{\mathrm{sw,I}}$, which will be done in the following for the example of $T^\star/a = 16$.

Figure 9.17: Illustration of the errors on $c_{\mathrm{sw,I}}$ as a function of $g^2_{0,\mathrm{I}}$ using the example of $T^\star/a = 16$. The central point and the statistical error $\Delta c_{\mathrm{sw,I}}$ shown in black, as well as the fit function are taken over from fig. 9.15. The orange error bars represent the errors due to the mismatch of the LCP, $\tilde{\Delta} g^2_{0,\mathrm{I}}$ (cf. tab. 9.17) and $\tilde{\Delta} c_{\mathrm{sw,I}}$ (cf. (9.135)).



- $\tilde{\Delta} c_{\mathrm{sw,I}}$

To estimate $\tilde{\Delta} c_{\mathrm{sw,I}}$, we investigate the dependence (9.131) by evaluating $T^\star \Delta M^{ud}_{(j)}$ again at slight variations of the elements $g^2_{0,\mathrm{L},(j)}$ of $\{g^2_{0,\mathrm{L}}\}$, which after the usual interpolation to $c_{\mathrm{sw,I}}$ give information on the derivatives in

$$\tilde{\Delta} c_{\mathrm{sw,I}} = \sqrt{\sum_j \left(\frac{\partial c_{\mathrm{sw,I}}}{\partial g^2_{0,\mathrm{L},(j)}} \cdot \tilde{\Delta} g^2_{0,\mathrm{L}} \right)^2} \,, \tag{9.133}$$

where correlations are neglected. At $T^\star/a = 16$ and $c_{\mathrm{sw}} = 1.8$, for instance, we find

$$\frac{\partial c_{\mathrm{sw,I}}}{\partial \left(T^\star \Delta M^{ud}_{(4)} \right)} \sim 3 \qquad \text{and} \qquad \frac{\partial \left(T^\star \Delta M^{ud}_{(4)} \right)}{\partial g^2_{0,\mathrm{L},(4)}} \sim -0.5 \,, \tag{9.134}$$

where the first quantity was computed for the case of the *individual* fit, and also for the other $j \in \{1, \ldots, N_{c_{\mathrm{sw}}}\}$. The evaluation of the second quantity needs much more effort since additional Monte Carlo simulations at modified $g^2_{0,\mathrm{L},(j)}$ are required. Assuming, somewhat simplifying, that its order of magnitude is the same for all j (or c_{sw}), together with (9.133) and $\tilde{\Delta} g^2_{0,\mathrm{I}} = 0.0014$ (see tab. 9.17), this leads to a rough estimate of

$$\tilde{\Delta} c_{\mathrm{sw,I}} \sim 0.005 \,. \tag{9.135}$$

- $\tilde{\Delta} g^2_{0,\mathrm{I}}$

The errors $\tilde{\Delta} g^2_{0,\mathrm{I}}$ are obtained via (9.132) from propagation of the errors of the fit parameters, $\tilde{\Delta} g^2_{0,\mathrm{L}}\big|_{c_{\mathrm{sw}}=0}$ and $\tilde{\Delta} \left(dg^2_{0,\mathrm{L}} / dc_{\mathrm{sw}} \right)$, but neither $\Delta c_{\mathrm{sw,I}}$ nor $\tilde{\Delta} c_{\mathrm{sw,I}}$, and are listed in tab. 9.17. They may subsequently be transferred to additional contributions to the total error of $c_{\mathrm{sw,I}}$ under use of the Padé fit (9.129), according to (H.15). Irrespective of the order of the Padé fit and the use of the global or individual fit results, one obtains

$$\left|\frac{\partial c_{\mathrm{sw,I}}}{\partial g_{0,\mathrm{I}}^2}\right|_{g_{0,\mathrm{I}}^2=1.7144} \cdot \tilde{\Delta} g_{0,\mathrm{I}}^2 \sim 0.003 \,, \tag{9.136}$$

which hence, like $\tilde{\Delta} c_{\mathrm{sw,I}}$ (9.135), turns out to be much smaller than the statistical error (cf. tab. 9.16),

$$\Delta c_{\mathrm{sw,I}} \sim 0.019 \,. \tag{9.137}$$

Conclusion

In conclusion, we find that under the given assumptions, the achieved precision of the tuning to the LCP renders both the associated errors $\tilde{\Delta} c_{\mathrm{sw,I}}$ and $\tilde{\Delta} g_{0,\mathrm{I}}^2$ dominated by the statistical errors $\Delta c_{\mathrm{sw,I}}$ that emerge from the imposition of the improvement condition. Finally, we note that the errors $\tilde{\Delta} \kappa_{l,\mathrm{L}}$ and $\tilde{\Delta} \kappa_{c,\mathrm{L}}$ from the mismatch of the hopping parameters contribute to $\tilde{\Delta} c_{\mathrm{sw,I}}$ as well. This could be investigated in a similar way to (9.133, 9.134, 9.135), but is assumed to be insignificant in comparison with (9.137), too.

9.6.5 Cutoff effects and cross checks

In this last section, we investigate two additional quantities, the PCAC mass ratio R_{PCAC} for the tuning and improvement simulations, as well as the gradient flow coupling $\overline{g}_{\mathrm{GF}}^2$ (without normalization) for the improvement simulations. In particular, we aim to

P0. check that the continuum results are in accordance with expectations

P1. verify the absence of $O(a)$ cutoff effects and thus our results for c_{sw}

P2. investigate the size of remaining $O(a^2)$ cutoff effects.

PCAC mass ratio

If all the improvement coefficients assume proper values, defined by appropriate improvement conditions (like (9.115) for c_{sw}), the cutoff effects of the renormalized masses

$$m_{\mathrm{R}}^{ij} = \widetilde{Z}_{m_{\mathrm{PCAC}}}^{ij}(g_0^2, aM, a\mu) \cdot \; m_{\mathrm{PCAC,I}}^{ij}(g_0^2, aM, c_{\mathrm{sw}}, c_A, P_{\mathrm{SF}}) + O(a^z) \tag{9.106}$$

are of second order in the lattice spacing, $z = 2$. It is instructive to verify this behavior by considering the PCAC mass ratio

$$R_{\mathrm{PCAC}} := \frac{2m^{uc}_{\mathrm{PCAC,I}}(T^\star/2) - m^{ud}_{\mathrm{PCAC,I}}(T^\star/2)}{m^{c\tilde{c}}_{\mathrm{PCAC,I}}(T^\star/2)} = 1 + a\,(b_A - b_P)\,\frac{m_{q,c}}{2} + O(a^2)\,. \quad (9.138)$$

The expression on the right hand side follows from using (9.106), after which the flavor-independent part of the renormalization constants $\left(\widetilde{Z}^{ij}_{m_{\mathrm{PCAC}}}\right)^{-1}$, i.e. the ratio of $\widetilde{Z}_P$ and $\widetilde{Z}_A$ (cf. (8.46, 8.47, 9.105)), drops out. Note that the masses of the up and down quark were neglected on the $O(a)$ level, $am_{q,u}, am_{q,d} \ll am_{q,c}$ and $m^u_{\mathrm{R}}, m^d_{\mathrm{R}} \ll m^c_{\mathrm{R}}$. Furthermore, the time slice $x_0 \in P_{\mathrm{SF}}$, at which the unrenormalized PCAC masses are taken, was specifically chosen as $x_0 = T^\star/2$ here. Consequently, the deviation of R_{PCAC} from 1 is a pure lattice artifact that may be studied on the lattice under use of PCAC masses of different flavors. Given the estimate $am_{q,c} \lesssim 0.5$ (cf. (8.38b)) and the available results for $b_A - b_P$ in theories with different N_f (cf. sec. 3.9), the size of the $O(a)$ term can be assumed to be rather small. If this is the case, a continuum extrapolation of R_{PCAC} linear in $(a/T^\star)^2$ should be applicable.

This will be carried out in the following for both the tuning and the improvement simulations. Although the latter are the very same simulations which determined[48] $c_{\mathrm{sw,I}}$ (and $c_{A,\mathrm{I}}$) in the first place, the investigation is useful also with regard to P1 as the improvement condition involves the PCAC masses at time slices $x_0 = 3T^\star/4$ and $y_0 = T^\star/4$ (cf. (9.115b)), whereas the PCAC mass ratio R_{PCAC} depends on the PCAC mass in the center of the SF, $x_0 = T^\star/2$ (cf. (9.138)). Even more importantly, R_{PCAC} involves the charm as a valence quark, in contrast to the PCAC masses which enter the improvement condition. The tuning simulations, in contrast, are truly independent of the determination of the improvement coefficients, and in addition employ a SF with different boundary conditions and larger spatial extent, as has been explained in sec. 9.4. Each case thus provides an independent check.

We begin with the evaluation of R_{PCAC} for all combinations of $\{T^\star/a, c_{\mathrm{sw}}\}$, the results being listed in tab. 9.18 and tab. 9.19 for the tuning and improvement runs, respectively. For each $T^\star/a$, an interpolation to the respective $c_{\mathrm{sw,I}}$, displayed in the left panel of fig. 9.18, leads to the PCAC mass ratios $R_{\mathrm{PCAC,I}}$ at the improvement points, to be found in tab. 9.20. These are finally extrapolated to the continuum, as depicted in the right panel of fig. 9.18.

---

[48] We use the notation $c_{\mathrm{sw,I}}$ again for the coefficient that fulfills the improvement condition, to distinguish it from values of c_{sw} at which the simulations took place, cf. footnotes 45 and 46.

$(T^\star/a = 8)$

$c_{\rm sw}$	$g_{0,\rm L}^2$	$\kappa_{l,\rm L}$	$\kappa_{c,\rm L}$	$T^\star m_{\rm PCAC}^{ud}$	$T^\star m_{\rm PCAC}^{uc}$	$T^\star m_{\rm PCAC}^{c\bar{c}}$	$R_{\rm PCAC}$
1.9	1.7848	0.13741	0.12033	0.112(5)	2.020(5)	4.713(7)	0.8333(2)
2.0	1.8067	0.13645	0.11960	0.105(4)	2.009(4)	4.713(5)	0.8302(2)
2.1	1.8340	0.13560	0.11900	0.112(6)	2.002(6)	4.709(7)	0.8265(3)
2.2	1.8534	0.13450	0.11820	0.109(5)	1.991(4)	4.698(6)	0.8243(2)
2.3	1.8825	0.13369	0.11770	0.111(4)	1.933(4)	4.550(5)	0.8255(2)
2.4	1.9069	0.13278	0.11750	0.101(5)	1.891(5)	4.464(6)	0.8248(2)

$(T^\star/a = 12)$

$c_{\rm sw}$	$g_{0,\rm L}^2$	$\kappa_{l,\rm L}$	$\kappa_{c,\rm L}$	$T^\star m_{\rm PCAC}^{ud}$	$T^\star m_{\rm PCAC}^{uc}$	$T^\star m_{\rm PCAC}^{c\bar{c}}$	$R_{\rm PCAC}$
1.8	1.7334	0.13742	0.12776	0.101(6)	1.711(5)	3.620(5)	0.9171(3)
1.9	1.7600	0.13656	0.12714	0.103(4)	1.685(3)	3.572(4)	0.9147(3)
2.0	1.7876	0.13573	0.12653	0.101(5)	1.662(4)	3.526(5)	0.9139(2)
2.1	1.8152	0.13486	0.12591	0.098(6)	1.630(5)	3.464(6)	0.9129(4)
2.2	1.8423	0.13401	0.12529	0.096(6)	1.595(4)	3.396(4)	0.9112(3)

$(T^\star/a = 16)$

$c_{\rm sw}$	$g_{0,\rm L}^2$	$\kappa_{l,\rm L}$	$\kappa_{c,\rm L}$	$T^\star m_{\rm PCAC}^{ud}$	$T^\star m_{\rm PCAC}^{uc}$	$T^\star m_{\rm PCAC}^{c\bar{c}}$	$R_{\rm PCAC}$
1.5	1.6266	0.13918	0.13224	0.094(3)	1.612(3)	3.283(3)	0.9533(2)
1.6	1.6482	0.13825	0.13139	0.098(3)	1.620(2)	3.300(2)	0.9522(2)
1.7	1.6741	0.13741	0.13087	0.098(3)	1.564(3)	3.181(3)	0.9529(2)
1.8	1.7036	0.13661	0.13018	0.095(3)	1.553(3)	3.166(3)	0.9514(2)
1.9	1.7331	0.13582	0.12949	0.094(3)	1.541(3)	3.145(3)	0.9496(2)
2.0	1.7626	0.13503	0.12880	0.091(3)	1.524(3)	3.120(3)	0.9480(2)

$(T^\star/a = 20)$

$c_{\rm sw}$	$g_{0,\rm L}^2$	$\kappa_{l,\rm L}$	$\kappa_{c,\rm L}$	$T^\star m_{\rm PCAC}^{ud}$	$T^\star m_{\rm PCAC}^{uc}$	$T^\star m_{\rm PCAC}^{c\bar{c}}$	$R_{\rm PCAC}$
1.5	1.5868	0.13822	0.13287	0.106(3)	1.599(3)	3.193(3)	0.9681(2)
1.6	1.6153	0.13749	0.13240	0.094(3)	1.532(3)	3.065(3)	0.9693(2)
1.7	1.6434	0.13669	0.13178	0.096(2)	1.499(2)	2.995(2)	0.9691(1)
1.8	1.6714	0.13588	0.13126	0.095(2)	1.430(2)	2.851(2)	0.9698(2)
1.9	1.6993	0.13504	0.13046	0.095(2)	1.431(2)	2.856(2)	0.9687(2)

$(T^\star/a = 24)$

$c_{\rm sw}$	$g_{0,\rm L}^2$	$\kappa_{l,\rm L}$	$\kappa_{c,\rm L}$	$T^\star m_{\rm PCAC}^{ud}$	$T^\star m_{\rm PCAC}^{uc}$	$T^\star m_{\rm PCAC}^{c\bar{c}}$	$R_{\rm PCAC}$
1.5	1.5527	0.13751	0.13325	0.099(2)	1.552(2)	3.072(2)	0.9781(1)
1.6	1.5783	0.13672	0.13253	0.096(2)	1.541(2)	3.056(2)	0.9771(1)
1.7	1.6065	0.13595	0.13202	0.093(2)	1.464(2)	2.897(2)	0.9785(2)
1.8	1.6344	0.13515	0.13132	0.087(1)	1.438(1)	2.852(1)	0.9780(1)

Table 9.18: Results for the PCAC masses of different quark flavors and the mass ratio $R_{\rm PCAC}$ (cf. (9.138)) for the *tuning* runs.

$(T^\star/a = 8)$

$c_{\rm sw}$	$g^2_{0,\rm L}$	$\kappa_{l,\rm L}$	$\kappa_{c,\rm L}$	$T^\star m^{ud}_{\rm PCAC}$	$T^\star m^{uc}_{\rm PCAC}$	$T^\star m^{c\tilde{c}}_{\rm PCAC}$	$R_{\rm PCAC}$
1.9	1.7848	0.13741	0.12033	0.504(7)	3.119(8)	6.218(8)	0.9221(4)
2.0	1.8067	0.13645	0.11960	0.491(8)	3.100(9)	6.211(10)	0.9191(3)
2.1	1.8340	0.13560	0.11900	0.509(8)	3.105(9)	6.217(10)	0.9169(3)
2.2	1.8534	0.13450	0.11820	0.516(8)	3.101(8)	6.217(9)	0.9146(4)
2.3	1.8825	0.13369	0.11770	0.521(9)	3.073(9)	6.166(11)	0.9122(3)
2.4	1.9069	0.13278	0.11750	0.521(10)	2.972(10)	5.940(11)	0.9129(4)

$(T^\star/a = 12)$

$c_{\rm sw}$	$g^2_{0,\rm L}$	$\kappa_{l,\rm L}$	$\kappa_{c,\rm L}$	$T^\star m^{ud}_{\rm PCAC}$	$T^\star m^{uc}_{\rm PCAC}$	$T^\star m^{c\tilde{c}}_{\rm PCAC}$	$R_{\rm PCAC}$
1.8	1.7334	0.13742	0.12776	0.247(4)	2.095(4)	4.144(5)	0.9514(3)
1.9	1.7600	0.13656	0.12714	0.242(4)	2.059(4)	4.078(5)	0.9505(3)
2.0	1.7876	0.13573	0.12653	0.237(4)	2.026(4)	4.020(4)	0.9489(3)
2.1	1.8152	0.13486	0.12591	0.243(5)	1.993(5)	3.950(5)	0.9473(3)
2.2	1.8423	0.13401	0.12529	0.223(4)	1.941(4)	3.865(5)	0.9465(3)

$(T^\star/a = 16)$

$c_{\rm sw}$	$g^2_{0,\rm L}$	$\kappa_{l,\rm L}$	$\kappa_{c,\rm L}$	$T^\star m^{ud}_{\rm PCAC}$	$T^\star m^{uc}_{\rm PCAC}$	$T^\star m^{c\tilde{c}}_{\rm PCAC}$	$R_{\rm PCAC}$
1.5	1.6266	0.13918	0.13224	0.160(3)	1.799(3)	3.538(3)	0.9717(2)
1.6	1.6482	0.13825	0.13139	0.159(2)	1.799(3)	3.546(3)	0.9701(2)
1.7	1.6741	0.13741	0.13087	0.156(2)	1.740(3)	3.422(3)	0.9710(2)
1.8	1.7036	0.13661	0.13018	0.167(2)	1.739(3)	3.414(3)	0.9696(2)
1.9	1.7331	0.13582	0.12949	0.165(3)	1.727(3)	3.395(3)	0.9686(2)
2.0	1.7626	0.13503	0.12880	0.154(2)	1.698(3)	3.353(3)	0.9668(2)

$(T^\star/a = 20)$

$c_{\rm sw}$	$g^2_{0,\rm L}$	$\kappa_{l,\rm L}$	$\kappa_{c,\rm L}$	$T^\star m^{ud}_{\rm PCAC}$	$T^\star m^{uc}_{\rm PCAC}$	$T^\star m^{c\tilde{c}}_{\rm PCAC}$	$R_{\rm PCAC}$
1.5	1.5868	0.13822	0.13287	0.150(2)	1.721(2)	3.359(3)	0.9804(1)
1.6	1.6153	0.13749	0.13240	0.131(2)	1.640(2)	3.213(3)	0.9803(2)
1.7	1.6434	0.13669	0.13178	0.139(2)	1.613(2)	3.149(3)	0.9802(2)
1.8	1.6714	0.13588	0.13126	0.132(3)	1.530(3)	2.987(3)	0.9807(2)
1.9	1.6993	0.13504	0.13046	0.140(2)	1.538(3)	2.998(3)	0.9794(3)

$(T^\star/a = 24)$

$c_{\rm sw}$	$g^2_{0,\rm L}$	$\kappa_{l,\rm L}$	$\kappa_{c,\rm L}$	$T^\star m^{ud}_{\rm PCAC}$	$T^\star m^{uc}_{\rm PCAC}$	$T^\star m^{c\tilde{c}}_{\rm PCAC}$	$R_{\rm PCAC}$
1.5	1.5527	0.13751	0.13325	0.129(2)	1.632(2)	3.179(2)	0.9858(1)
1.6	1.5783	0.13672	0.13253	0.120(1)	1.613(2)	3.154(2)	0.9850(2)
1.7	1.6065	0.13595	0.13202	0.119(1)	1.536(1)	2.995(2)	0.9859(1)
1.8	1.6344	0.13515	0.13132	0.113(2)	1.510(2)	2.948(2)	0.9856(1)

Table 9.19: Results for the PCAC masses of different quark flavors and the mass ratio $R_{\rm PCAC}$ (cf. (9.138)) for the *improvement* runs.

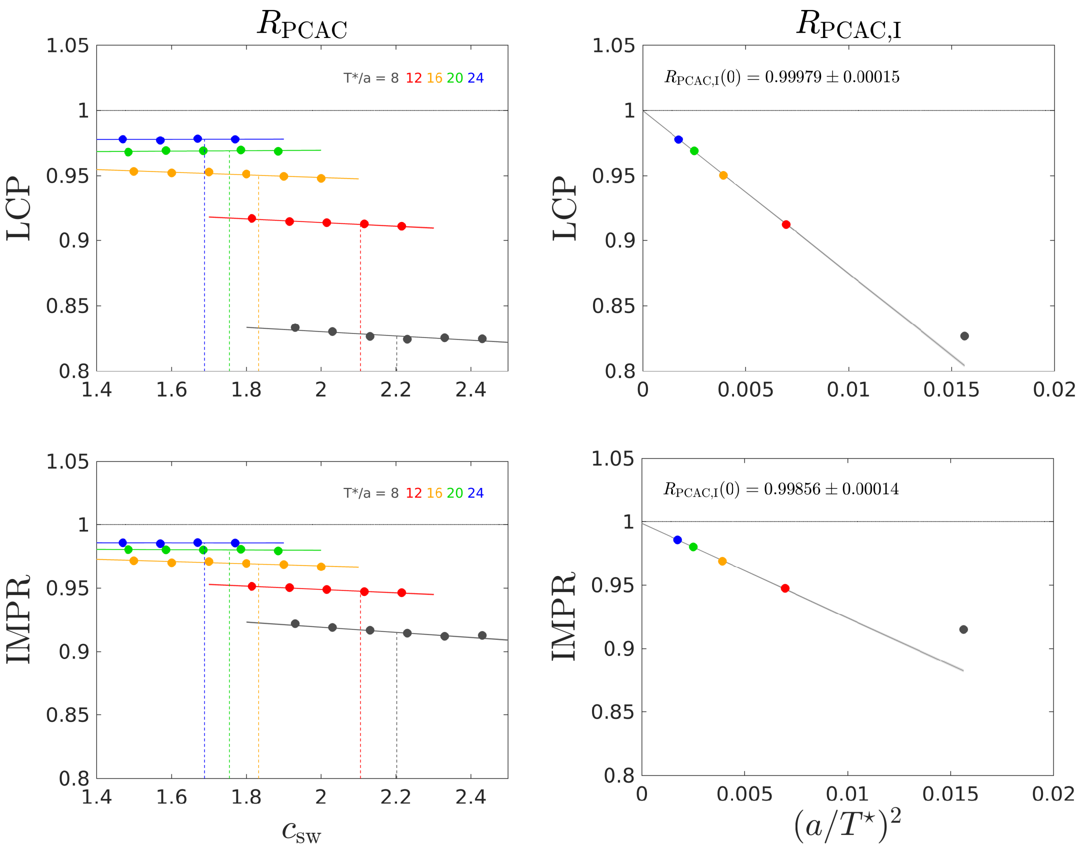


Figure 9.18: *Left:* PCAC mass ratio R_{PCAC} (9.138) for the different combinations of $\{T^{\star}/a, c_{\mathrm{sw}}\}$, and the respective linear interpolations (solid lines) for each $T^{\star}/a$. The vertical, dashed lines indicate the improvement points $c_{\mathrm{sw,I}}$ determined by the *global* fit, at which the $O(a)$ improved PCAC mass ratios $R_{\mathrm{PCAC,I}}$ are determined. *Right:* Improved PCAC mass ratios $R_{\mathrm{PCAC,I}}$ for the different lattice sizes $T^{\star}/a$ and their extrapolation to the continuum under neglect of $T^{\star}/a = 8$.

global		LCP		IMPR	
$T^\star/a$	$c_{\mathrm{sw,I}}$	Q_{fit}	$R_{\mathrm{PCAC,I}}$	Q_{fit}	$R_{\mathrm{PCAC,I}}$
8	2.202(21)	0.00	0.82679(9)(33)	0.00	0.91517(15)(33)
12	2.105(24)	0.06	0.91248(18)(34)	0.51	0.94759(18)(34)
16	1.833(19)	0.00	0.95035(9)(19)	0.00	0.96894(9)(19)
20	1.755(28)	0.00	0.96906(8)(5)	0.01	0.98013(11)(5)
24	1.688(24)	0.00	0.97784(7)(1)	0.00	0.98564(7)(1)
Q_{fit}			0.00		0.00
∞			0.99979(15)		0.99856(14)

Table 9.20: Results for the mass ratio $R_{\mathrm{PCAC,I}}$ (9.138), interpolated to the improvement points $c_{\mathrm{sw,I}}$ under use of the data shown in tab. 9.18 (tuning, LCP) and tab. 9.19 (improvement, IMPR). The first error on $R_{\mathrm{PCAC,I}}$ follows from the associated linear fit (cf. left panel of fig. 9.18), whose fit quality is given in the previous column, respectively, while the second error is due to the uncertainty of $c_{\mathrm{sw,I}}$, determined by the global fits. The combined errors are taken into account for the continuum extrapolations (cf. right panel of fig. 9.18), for which the fit quality and the continuum values are given in the last two rows. Note that the point at $T^\star/a = 8$ is not taken into account here.

In the case of the *global* fits, the procedure yields

$$\begin{aligned} & R_{\mathrm{PCAC,I}}(a/T^\star \to 0) = 0.99979(15) \qquad \text{for the tuning runs} \\ \text{and} \quad & R_{\mathrm{PCAC,I}}(a/T^\star \to 0) = 0.99856(14) \qquad \text{for the measurement runs.} \end{aligned} \tag{9.139}$$

The results for the *individual* fits are almost identical. With regard to P0, we thus find approximate agreement of the continuum limit of $R_{\mathrm{PCAC,I}}$ with the expected value 1 (cf. (9.138)). The deviation is presumably due to the fact that the errors from the LCP have not been taken into account. These are probably non-negligible especially in view of the given large statistical precision of R_{PCAC}, and propagate to the (first) error of $R_{\mathrm{PCAC,I}}$. This would also explain why the fit quality of R_{PCAC} as a linear function of c_{sw} is quite bad in many cases, see tab. 9.20. Furthermore, the investigation reveals a dominance of $O(a^2)$ cutoff effects. However, the dependence of R_{PCAC} on c_{sw} is negligible (at large $T^\star/a$) or weak (at small $T^\star/a$). Consequently, the continuum limit $R_{\mathrm{PCAC,I}}$ is almost insensitive to c_{sw}, and may thus not serve as a cross check of our results for $c_{\mathrm{sw,I}}$ (cf. P1). Finally, referring to P2, we observe a moderate size of $O(a^2)$ cutoff effects, which in view of the involvement of the charm quark might not be

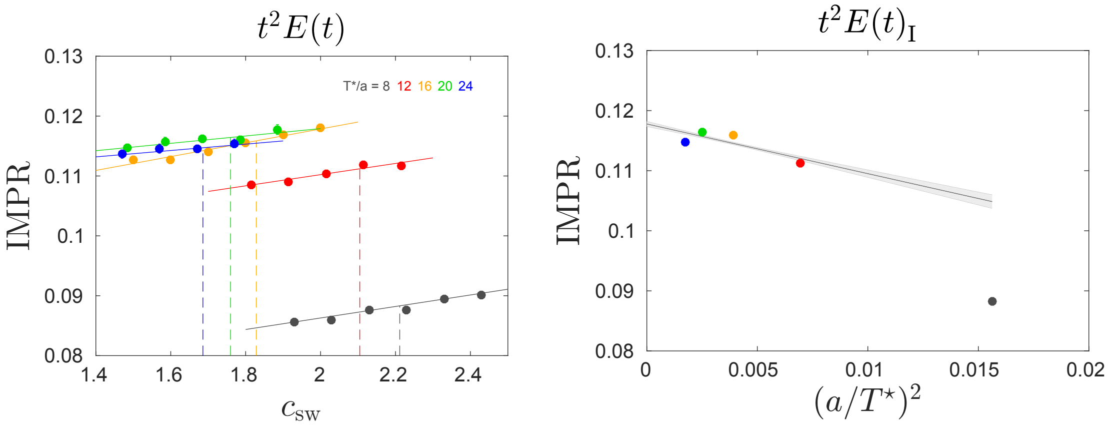


Figure 9.19: *Left:* $t^2E(t)$ for the different combinations of $\{T^\star/a, c_{\mathrm{sw}}\}$, and the respective linear interpolations (solid lines) for each $T^\star/a$. The vertical, dashed lines indicate the improvement points $c_{\mathrm{sw,I}}$ determined by the *global* fit, at which the $O(a)$ improved $t^2E(t)$ are determined. *Right:* Improved $t^2E(t)_{\mathrm{I}}$ for the different lattice sizes $T^\star/a$ and their extrapolation to the continuum under neglect of $T^\star/a = 8$.

self-evident but is reassuring.

Gradient flow coupling

For the tuning to the LCP, the gradient flow coupling $\Phi_1 = \bar{g}^2_{\mathrm{GF}}$ was fixed to $\Phi_1^\star$. There was no need to distinguish between $T^\star$ and L as they were the same, i.e. we had $t = c^2L^2/8 = c^2T^{\star 2}/8$ (cf. (6.7)). In the case of the improvement simulations, where the spatial volume is reduced by a factor two with respect to the tuning simulations (cf. (9.118)), we retain the scaling of the flow time by specifying $t = c^2T^{\star 2}/8$. However, for the boundary conditions of the improvement simulations (cf. (9.119)), the normalization of the gradient flow coupling is not known. Hence, we only display the results for $t^2E(t)$ in the left panel of fig. 9.19, in combination with the smoothing fraction $c = 0.25$ up to which measurements were taken. Interpolation to $c_{\mathrm{sw,I}}$ for each $T^\star/a$ leads to the data whose extrapolation to the continuum is shown in the right panel of the same figure.

As for R_{PCAC}, we assume that the error due to imprecise tuning to the LCP is not negligible in view of the statistical precision of our data. With regard to P1 and P2, we thus interpret the modest cutoff effects to be in accordance with $(a/T^\star)^2$ scaling, except for the case of $T^\star/a = 8$. The latter, however, may be understood by noting that the ratio of the lattice spacing and the smoothing range, $\epsilon = a/\sqrt{8t} = a/(cT^\star) = 1/2$, becomes large, see [149] for details. We find that $t^2E(t)$ also only rather weakly depends on c_{sw}, which in combination with the limited quality of the data does not allow for a precision test of our results for c_{sw}.

Finally, we remark that as the topological charge Q_{top} at $c = 0.3$ does not leave the trivial sector (cf. app. I.4), the gradient flow coupling $\overline{g}^2_{\mathrm{GF}}$ associated with $t^2E(t)$ basically equals the corresponding modified coupling $\overline{g}^2_{\mathrm{GF},0}$.

10 Summary

In the first part of this work, the gradient flow coupling was investigated in pure $SU(3)$ Yang–Mills theory in a volume of approximately $V \sim (0.8\ \mathrm{fm})^4$. A significant correlation between the coupling and the topological charge was found, which increases in strength with the smoothing fraction c. Simulations suffering from a bad sampling of topological sectors and critical slowing down would lead to a biased determination of $\overline{g}^2_{\mathrm{GF}}$ in the continuum. In contrast, the determination of an alternative definition for the coupling, $\overline{g}^2_{\mathrm{GF},0}$, which takes into account only the trivial topological sector, is not affected by the bad topology sampling in the studied volume.

The second, major part of this thesis was devoted to $O(a)$ improvement of the Wilson fermion action in the $N_f = 4^*$ theory of four non-degenerate, massive fermions. Due to the large cutoff effects that go along with the charm quark mass, as well as the numerous improvement terms that arise in the case of non-degeneracy, the previously expedient approach of a mass-independent renormalization and improvement scheme had to be abandoned. We argued that a mass-dependent scheme is more suitable to accommodate the dynamical charm quark effects, and presented an outline of the associated renormalization patterns of the bare parameters and bilinears. Within this framework, the determination of the Sheikholeslami–Wohlert coefficient c_{sw} as the first step of the Symanzik improvement program in $N_f = 4^*$ was performed. The improvement condition was formulated in the Schrödinger functional with $N_f = 3 + 1$ massive, dynamical quarks, which in comparison with $N_f = 4^*$ represents a vast simplification at the bearable cost of a rather insignificant introduction of small overall $O(a^2)$ cutoff effects. It was imposed along a line of constant physics, which employs the gradient flow coupling and effective meson masses, set up to implicitly fix the volume of the Schrödinger functional as well as the light and charm quark masses. The relation between these two sets of quantities was established under use of the $N_f = 2$ theory. The lattice extent was held constant at a value of $T^\star \sim 0.8$ fm, chosen large enough to ensure a good signal, while small enough to avoid the involvement of non-trivial topological sectors and the associated problems of critical slowing down and topology freezing that had been considered earlier. Both the use of massive Wilson fermions and the fixed volume approach represent an innovation with respect to previous determinations of c_{sw} with a smaller number of dynamical quark flavors.

Our numerical effort involved the tree-level improved Lüscher–Weisz gauge action. The sensitivity of the quantity $T^{\star}\Delta M^{ud}$ (that is required to vanish by the improvement condition) on c_{sw} was found to be satisfying even for the finest lattices, and the results for the different lattice sizes could reasonably be described by a restrictive, global fit. The main result of this work is expressed in the interpolating formula

$$c_{\mathrm{sw}}(g_0^2) = \frac{1 - 0.257 g_0^2 - 0.050 g_0^4}{1 - 0.453 g_0^2} \qquad \text{for } \ 0 \leq g_0^2 \leq 1.86\,. \tag{9.129}$$

It builds the basis for further steps in the Symanzik improvement program, such as the determination of c_A, and may be employed in future large volume simulations under inclusion of dynamical charm quark effects, to extract physical observables in the continuum limit with severely reduced errors.

Appendix

A Notation

A.1 Groups and group transformations

The group $SU(N)$

$SU(N)$ is the group of complex $N \times N$ matrices which are unitary,

$$\Omega\Omega^\dagger = \Omega^\dagger\Omega = \mathbb{1} \, , \tag{A.1}$$

and special, $\det\Omega = 1$. As a compact *Lie group*, its elements can be parametrized by

$$\Omega = \exp\{i\alpha^a T^a\} \, , \tag{A.2}$$

with $\alpha^a \in \mathbb{R}$ and $a \in \{1, \dots, N^2 - 1\}$. Here, T^a are hermitean, traceless $N \times N$ matrices, the generators of the group, which obey the *Lie Algebra*

$$[T^a, T^b] = i f^{abc} T^c \tag{A.3}$$

with totally antisymmetric $SU(N)$ structure constants. For $N = 2$, these are equal to the Levi–Civita tensor components ϵ^{abc}, while for $N = 3$ the non-vanishing components are:

$$f^{123} = 1, \tag{A.4}$$

$$f^{147} = f^{246} = f^{345} = f^{516} = f^{257} = f^{637} = \frac{1}{2}, \tag{A.5}$$

$$f^{458} = f^{678} = \frac{\sqrt{3}}{2} \, . \tag{A.6}$$

The anticommutator of the generators reads

$$\left\{T^a, T^b\right\} = d^{abc} T^c + \frac{1}{N}\delta^{ab}\mathbb{1}_N \, , \tag{A.7}$$

where the coefficients d^{abc} are totally symmetric.

The generators are usually chosen as $T^a = \lambda^a/2$, where λ^a are the 3 Pauli matrices for $N = 2$, and the 8 *Gell-Mann matrices* for $N = 3$:

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; \lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{A.8}$$

$$\lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \; \lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \tag{A.9}$$

$$\lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \; \lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \; \lambda^8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{A.10}$$

In particular, this implies the normalizations

$$T_R = 1/2 \tag{A.11}$$

$$\text{and} \qquad C_F = \frac{N^2-1}{2N} = \begin{cases} 3/4 & \text{for } N = 2 \\ 4/3 & \text{for } N = 3 \end{cases} \tag{A.12}$$

for the trace condition[1] and the Casimir operator,

$$\mathrm{Tr}[T^a \cdot T^b] = T_R \cdot \delta^{ab} \tag{A.13}$$

$$\text{and} \qquad T^a T^a = C_F \cdot \mathbb{1}_N\,, \tag{A.14}$$

respectively.

Instead of the hermitean generators T^a, one may also employ *anti-hermitean generators* $\tilde{T}^a$, related by $\tilde{T}^a = iT^a$. Those obey

$$[\tilde{T}^a, \tilde{T}^b] = f^{abc}\tilde{T}^c \tag{A.15}$$

$$\text{and} \qquad \mathrm{Tr}[\tilde{T}^a \cdot \tilde{T}^b] = -T_R \cdot \delta^{ab}\,, \tag{A.16}$$

as follows immediately from (A.3) and (A.13).

---

[1] Note that in contrast to the main text, where tr and Tr denote the trace in color and flavor space, respectively, the symbol Tr stands for the trace in general here.

Group transformations

In QCD, two particular group transformations are of fundamental importance. On the one hand, these are the *local gauge transformations*

$$\Omega(x) = \exp(i\alpha^a(x)T^a) \in SU(N_c)_c \tag{A.17}$$

with $N_c = 3$. They act in color space—indicated by the subscript c—for instance on the color components of a spinor $\psi = (\psi_{\text{red}}, \psi_{\text{green}}, \psi_{\text{blue}})^\top$. On the other hand, there are several *global flavor transformations*, which are marked with a subscript f. These are

$$\begin{aligned}
U_L \;\; &= \exp\{i\alpha_L^a T^a\} \in SU(N_f)_{fL} && \text{(A.18a)}\\
U_R \;\; &= \exp\{i\alpha_R^a T^a\} \in SU(N_f)_{fR} && \text{(A.18b)}\\
U_L^{(1)} &= \exp\{i\alpha_L\} \quad\; \in U(1)_{fL} && \text{(A.18c)}\\
U_R^{(1)} &= \exp\{i\alpha_R\} \quad\; \in U(1)_{fR}\ , && \text{(A.18d)}
\end{aligned}$$

which act only on the left- or right-handed components (cf. (2.21)), respectively:

$$\psi'_{L/R} = U_{L/R}\psi_{L/R} \qquad \text{and} \qquad \overline{\psi}'_{L/R} = \overline{\psi}_{L/R}U^\dagger_{L/R}\ . \tag{A.19}$$

In $N_f = 4$, we have $\psi_L = (\psi_{uL}, \psi_{dL}, \psi_{sL}, \psi_{cL})^\top$ and $\psi_R = (\psi_{uR}, \psi_{dR}, \psi_{sR}, \psi_{cR})^\top$. Important combined left- and right-handed transformations are the vector and axial transformations, which are obtained from (A.18) by taking

$$\alpha_V = \alpha_L = \alpha_R \qquad \text{and} \qquad \alpha_A = -\alpha_L = \alpha_R\ , \tag{A.20}$$

respectively. Acting on $\psi = (\psi_u, \psi_d, \psi_s, \psi_c)^\top$ in the case of $N_f = 4$, they amount to

$$\begin{aligned}
V \;\;\; &= \exp\{i\alpha_V^a T^a\} \quad\; \in SU(N_f)_{fV} && \text{(A.21a)}\\
A \;\;\; &= \exp\{i\alpha_A^a \gamma_5 T^a\} \in {''SU(N_f)''_{fA}} && \text{(A.21b)}\\
V^{(1)} &= \exp\{i\alpha_V\} \qquad\;\; \in U(1)_{fV} && \text{(A.21c)}\\
A^{(1)} &= \exp\{i\alpha_A\gamma_5\} \quad\; \in U(1)_{fA}\ , && \text{(A.21d)}
\end{aligned}$$

with

$$\begin{aligned} \psi' &= V\psi \\ \psi' &= A\psi \end{aligned} \qquad \text{and} \qquad \begin{aligned} \overline{\psi}' &= \overline{\psi}V^\dagger \\ \overline{\psi}' &= \overline{\psi}A \ . \end{aligned} \tag{A.22}$$

As indicated above, the transformations denoted by V, $V^{(1)}$ and $A^{(1)}$ form subgroups of the chiral group (2.23), respectively. This is not the case for the axial, flavor mixing transformations A. These form a coset of $SU(N_f)_{fV}$, which is reflected by the use of the notation $''SU(N_f)''_{fA}$.

A.2 Gamma matrices

In *Minkowski space* with the metric $g_{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)$, the gamma matrices γ_μ obey the anti-commutation relation

$$\{\gamma_\mu, \gamma_\nu\} = 2g_{\mu\nu}\mathbb{1}_4 \ . \tag{A.23}$$

For the matrix $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$, the following equations hold:

$$\{\gamma_5, \gamma_\mu\} = 0 \qquad \text{and} \qquad \gamma_5^2 = 1 \ . \tag{A.24}$$

In *Euclidean space*, the gamma matrices are related to those in Minkowski space by

$$\gamma_0^E = \gamma_0 \qquad \text{and} \qquad \gamma_j^E = i\gamma_j \ , \quad j = 1,2,3 \ . \tag{A.25}$$

The anti-commutation relation (A.23) becomes

$$\{\gamma_\mu^E, \gamma_\nu^E\} = 2\delta_{\mu\nu}\mathbb{1}_4 \ , \tag{A.26}$$

while (A.24) remains valid also for $\gamma_5^E = \gamma_0^E\gamma_1^E\gamma_2^E\gamma_3^E$.

A.3 Discrete derivatives

We consider a smooth function $f : \mathbb{R}^4 \to \mathbb{R}$, whose derivatives in the continuum with respect to the μ-direction are denoted by f', f'', and so forth. On a lattice with lattice spacing a, discretizations of these derivatives may involve $f(x + na\hat{\mu})$ with $n \in \mathbb{Z}$. Straight forward discretizations of the *first derivative* f' are obtained by the *forward* and *backward* lattice derivatives,

$$\partial_\mu f(x) = \frac{1}{a}\left[f(x+a\hat{\mu}) - f(x)\right] \tag{A.27}$$

$$\partial^*_\mu f(x) = \frac{1}{a}\left[f(x) - f(x-a\hat{\mu})\right] \ , \tag{A.28}$$

respectively. They introduce discretization errors of the form $O(a)$, as can readily be seen by using the Taylor expansion

$$f(x \pm a\hat{\mu}) = f(x) \pm af'(x) + \frac{a^2}{2}f''(x) + O(a^3) \tag{A.29}$$

in the above expressions. The situation is improved in the case of the *symmetric* lattice derivative,

$$\tilde{\partial}_\mu f(x) = \frac{1}{2}\left(\partial^*_\mu + \partial_\mu\right) f(x) = \frac{1}{2a}\left[f(x+a\hat{\mu}) - f(x-a\hat{\mu})\right] \ , \tag{A.30}$$

which equals f' up to $O(a^2)$. Concerning the discretization of the *second derivative* f'', the same order $O(a^2)$ of discretization effects is obtained by the use of

$$\tilde{\partial}^2_\mu f(x_0) = \frac{1}{(2a)^2}\left[f(x_0+2a\hat{\mu}) \ - 2f(x_0) + f(x_0-2a\hat{\mu})\ \right] \tag{A.31}$$

$$\text{and} \qquad \partial^*_\mu\partial_\mu f(x_0) = \frac{1}{a^2} \quad \left[f(x_0+a\hat{\mu}) \quad - 2f(x_0) + f(x_0-a\hat{\mu}) \quad \right] \ . \tag{A.32}$$

Obviously, the latter corresponds to the former with the replacement $2a \to a$. It is thus advantageous to use $\partial^*_\mu\partial_\mu f$, as its leading discretization errors are reduced by a factor four. Furthermore, it may be evaluated closer to the edges of a finite lattice.

B The axial current

B.1 Representation in the flavor and the generator basis

We consider the axial current with N_f flavors. While concrete expressions are shown for $N_f = 2$ for simplicity, the results will be applied in sec. 8.1.3 for $N_f = 4$. In $\mathbb{C}^{N_f \times N_f}$ regarded as a complex vector space, we define the *flavor basis* as matrices Δ^{ij} with components

$$(\Delta^{ij})_{kl} = \begin{cases} 1 & \text{if } k = i, l = j \\ 0 & \text{otherwise ,} \end{cases} \tag{B.1a}$$

where $i, j \in \{1, \dots, N_f\}$. The axial current in this basis reads

$$A_\mu = \sum_{i,j=1}^{N_f} A_\mu^{ij} \Delta^{ij} \overset{(N_f=2)}{=} \begin{pmatrix} A_\mu^{11} & A_\mu^{12} \\ A_\mu^{21} & A_\mu^{22} \end{pmatrix} , \tag{B.1b}$$

and its components are defined as

$$A_\mu^{ij} = \overline{\psi} \gamma_\mu \gamma_5 \Delta^{ij} \psi = \overline{\psi}^i \gamma_\mu \gamma_5 \psi^j , \qquad i, j \in \{1, \dots, N_f\} . \tag{B.1c}$$

In order to discuss the behavior of the axial current under $SU(N_f)$ flavor transformations and renormalization, it is convenient to switch to the *generator basis*, which consists of the generators of $SU(N_f)$, T^c with $c \in \{1, \dots, N_f^2 - 1\}$, extended by $T^0 = \mathbb{1}/N_f$. In $N_f = 2$, for instance, it reads

$$T^0 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} , \quad T^1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad T^2 = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \quad T^3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{B.2a}$$

In this basis, the axial current becomes

$$A_\mu = \sum_{c=0}^{N_f^2-1} A_\mu^c T^c \overset{(N_f=2)}{=} \frac{1}{2} \begin{pmatrix} A_\mu^0 + A_\mu^3 & A_\mu^1 - iA_\mu^2 \\ A_\mu^1 + iA_\mu^2 & A_\mu^0 - A_\mu^3 \end{pmatrix} , \tag{B.2b}$$

and its components can be written as

$$A_\mu^c = \overline{\psi}\gamma_\mu\gamma_5 T^c \psi , \qquad c \in \{0, \dots, N_f^2 - 1\} . \tag{B.2c}$$

B.2 The improved axial current under flavor transformations

Next, we discuss the behavior of the mass-dependent improvement terms of the axial current, as listed in sec. 8.1.3, under $SU(4)_{fV}$ flavor transformations in the generator basis, while charge conjugation is considered in app. B.3.

Under the generalized, vector spurionic $SU(4)_{fV}$ transformations (cf. (8.2))

$$\psi \overset{V}{\to} V\psi , \qquad \overline{\psi} \overset{V}{\to} \overline{\psi}V^\dagger \qquad \text{and} \qquad M \overset{V}{\to} VMV^\dagger , \tag{B.3}$$

with the notation

$$V^\dagger T^c V = R^{cd}(V) T^d , \tag{8.12}$$

the bilinear combinations as they appear in (8.14a) and (8.14b) behave like the axial current:

$$A_\mu^c \overset{V}{\to} R^{cd}(V) A_\mu^d . \tag{8.13}$$

We show this explicitly for the examples of the ω_0 and ω_1 terms in (8.14a):

$$\begin{aligned} \partial_\mu P^c &= \partial_\mu \left(\overline{\psi}\gamma_5 T^c \psi\right) \\ &\overset{V}{\to} \partial_\mu \left(\overline{\psi}V^\dagger\gamma_5 T^c V\psi\right) \\ &= \partial_\mu \left(R^{cd}(V)\overline{\psi}\gamma_5 T^d \psi\right) \\ &= R^{cd}(V)\, \partial_\mu P^d \end{aligned} \tag{B.4a}$$

and

$$\begin{aligned} \mathrm{Tr}[M] A_\mu^c &\overset{V}{\to} \mathrm{Tr}[VMV^\dagger]\, R^{cd}(V)\, A_\mu^d \\ &= R^{cd}(V)\, \mathrm{Tr}[M] A_\mu^c . \end{aligned} \tag{B.4b}$$

Note that the $SU(4)_{fV}$ matrices V and $V^\dagger$ act in flavor space and therefore commute with the Dirac matrices γ_μ and γ_5.

B.3 The improved axial current under charge conjugation

Under charge conjugation, the axial current in the continuum transforms as

$$A_\mu^c = \overline{\psi}\gamma_\mu\gamma_5 T^c \psi \overset{C}{\to} \overline{\psi}\gamma_\mu\gamma_5 [T^c]^\mathsf{T} \psi \,, \tag{B.5}$$

where the superscript T denotes the transposed matrix. It immediately follows that those components whose generators T^c are diagonal, are invariant under charge conjugation. They are hence referred to as neutral currents, see sec. 8.1.3.

Regarding the improvement terms in $N_f = 4^*$, the ones with the coefficient ω_2 in (8.14a) transform as ($M = M^\mathsf{T}$)

$$\overline{\psi}\gamma_\mu\gamma_5 T^c M \psi \overset{C}{\to} \overline{\psi}\gamma_\mu\gamma_5 M [T^c]^\mathsf{T} \psi \tag{B.6a}$$

$$\text{and} \quad \overline{\psi}\gamma_\mu\gamma_5 M T^c \psi \overset{C}{\to} \overline{\psi}\gamma_\mu\gamma_5 [T^c]^\mathsf{T} M \psi \,. \tag{B.6b}$$

As T^c and M do not commute for non-degenerate masses, the individual terms do not transform like A_μ^c under charge conjugation. However, the sum of both terms does:

$$\overline{\psi}\gamma_\mu\gamma_5 \left\{T^c, M\right\} \psi \overset{C}{\to} \overline{\psi}\gamma_\mu\gamma_5 \left\{[T^c]^\mathsf{T}, M\right\} \psi \,. \tag{B.7a}$$

The other linear combination, the difference of the two original terms, does not transform like A_μ^c:

$$\overline{\psi}\gamma_\mu\gamma_5 \left[T^c, M\right] \psi \overset{C}{\to} -\overline{\psi}\gamma_\mu\gamma_5 \left[[T^c]^\mathsf{T}, M\right] \psi \,. \tag{B.7b}$$

Since the Wilson formulation of Lattice QCD is required to be charge symmetric, $\overline{\psi}\gamma_\mu\gamma_5 \left[T^c, M\right] \psi$ is not included in the Symanzik effective axial current (8.14a).

C Chiral Ward identities

Ward identities may be derived by application of symmetry transformations on the fields in the path integral. These lead to relations between different classical currents, which are reinterpreted as operator relations in the quantum theory. While they hold exact in the continuum, they may be violated by cutoff effects on the lattice, depending on whether the lattice regularization respects or violates the respective symmetry.

We briefly sketch the derivation of the chiral Ward identities in the lattice regularized theory, loosely following [153]. Consider a multilocal operator $O(x_1, \ldots, x_n)$ of quark and gluon fields at the sites $x_1 \neq x_2 \neq \ldots \neq x_n$. Its expectation value (cf. (3.46)),

$$\langle O(x_1, \ldots, x_n) \rangle = \frac{1}{Z[\psi, \overline{\psi}, U]} \int \mathcal{D}[\psi, \overline{\psi}, U] \; e^{-S[\psi, \overline{\psi}, U]} O(x_1, \ldots, x_n) \,, \tag{C.1}$$

is invariant under *local* $SU(N_f)_{fL} \times SU(N_f)_{fR}$ transformations, i.e. (A.18a) and (A.18b) with the replacements $\alpha^a_{L/R} \to \alpha^a_{L/R}(x)$, as these merely constitute a change of integration variables. Hence, now considering the two special kinds of vector ($X = V$) and axial ($X = A$) transformations, one finds

$$\frac{\delta}{\delta \alpha^a_X(x)} \langle O(x_1, \ldots, x_n) \rangle = 0 \,, \tag{C.2}$$

or equivalently

$$\left\langle \frac{\delta O(x_1, \ldots, x_n)}{\delta \alpha^a_X(x)} \right\rangle = \left\langle O(x_1, \ldots, x_n) \, \frac{\delta S^W_F}{\delta \alpha^a_X(x)} \right\rangle \,. \tag{C.3}$$

If the expressions in (C.3) are explicitly written out, the equations for two different flavor indices a may be combined such that the *non-singlet*, so-called point-split currents[1] $\tilde{V}^{ij}_\mu(x)$ and $\tilde{A}^{ij}_\mu(x)$ with $j \neq i$ (cf. (3.75)) emerge for the vector and axial transformation, respectively. For the specific choice $O(x_1, x_2) = \psi_i(x_1)\overline{\psi}_j(x_2)$, it can be shown that the terms that stem from the left hand side of (C.3) yield propagators, present only if $x = x_1$ or $x = x_2$. The remainder from the right hand side of (C.3) contains

---

[1] These involve fields at x and the nearest neighbor points, see [153].

terms which combine either the point-split vector current $\tilde{V}_\mu^{ij}(x)$ with the scalar density $S^{ij}(x)$, or the point-split axial current with the pseudoscalar density $P^{ij}(x)$. In the latter case of the axial transformation, the variation of the Wilson term yields an additional $O(a)$ contribution with an operator O_{W5}^{ij} of dimension five:

$$\begin{aligned}
X = V: \quad & -\delta(x_2 - x)\,\langle \psi_i(x_1)\overline{\psi}_i(x_2)\rangle \quad - \delta(x_1 - x) \quad \left\langle \psi_j(x_1)\overline{\psi}_j(x_2)\right\rangle \\
& = \quad \tilde{\partial}_\mu \left\langle \tilde{V}_\mu^{ij}(x)\, O\right\rangle - (m_{0,i} - m_{0,j}) \left\langle S^{ij}(x)\, O\right\rangle \; . \qquad \text{(C.4)} \\
X = A: \quad & -\delta(x_2 - x)\,\langle \psi_i(x_1)\overline{\psi}_i(x_2)\rangle\,\gamma_5 - \delta(x_1 - x)\gamma_5 \left\langle \psi_j(x_1)\overline{\psi}_j(x_2)\right\rangle \\
& = \quad \tilde{\partial}_\mu \left\langle \tilde{A}_\mu^{ij}(x)\, O\right\rangle - (m_{0,i} + m_{0,j}) \left\langle P^{ij}(x)\, O\right\rangle \\
& \quad - a \left\langle O_{W5}^{ij}\, O\right\rangle \; . \qquad \text{(C.5)}
\end{aligned}$$

Although of higher order in the lattice spacing, the operator O_{W5}^{ij} may under renormalization mix with operators of lower order, namely those already present on the right hand side of (C.5) [154]:

$$\left[O_{W5}^{ij}\right]_R = Z_5 \left[O_{W5}^{ij} + \frac{1}{a}\left\{(Z_{\tilde{A}} - 1)\,\tilde{\partial}_\mu \tilde{A}_\mu^{ij}(x) + \left(\overline{m}_i + \overline{m}_j\right) P^{ij}\right\}\right] \; . \qquad \text{(C.6)}$$

Hence, by using (C.6) in (C.5) to replace O_{W5}^{ij}, additional terms arise at leading order, and (C.5) may be cast into the form

$$\begin{aligned}
X = A: \quad & -\delta(x_2 - x)\,\langle \psi_i(x_1)\overline{\psi}_i(x_2)\rangle\,\gamma_5 - \delta(x_1 - x)\gamma_5 \left\langle \psi_j(x_1)\overline{\psi}_j(x_2)\right\rangle \\
& = Z_{\tilde{A}}\, \tilde{\partial}_\mu \left\langle \tilde{A}_\mu^{ij}(x)\, O\right\rangle - (\overline{m}^i_{\text{PCAC}} + \overline{m}^j_{\text{PCAC}}) \left\langle P^{ij}(x)\, O\right\rangle \\
& \quad + O(a) \; . \qquad \text{(C.7)}
\end{aligned}$$

The *bare PCAC mass*[2] $\overline{m}^i_{\text{PCAC}}$ is related to the ordinary bare mass $m_{0,i}$ by

$$\overline{m}^i_{\text{PCAC}} = m_{0,i} - \overline{m}_i \; , \qquad \text{(C.8)}$$

where the mixing coefficients $\overline{m}_i$ are regular functions of the bare parameters of mass-dimension one, $\overline{m}_i = \frac{w(g_0^2, a\mathcal{M})}{a}$ [154].

Several important conclusions may be drawn from (C.4) and (C.7), and the observation that the propagators on the left hand sides of these equations are finite and

---

[2] Note that the definition here differs by cutoff effects from m^i_{PCAC}, the quantity employed in the main text (cf. e.g. (3.81)).

renormalized after multiplication with the field renormalization constants. The *first implication* concerns the renormalization of the (point-split) vector and axial current. In the special case of degenerate bare masses $m_{0,i}$ and $m_{0,j}$ $(X = V)$ or vanishing PCAC masses $\overline{m}^i_{\mathrm{PCAC}}$ and $\overline{m}^j_{\mathrm{PCAC}}$ $(X = A)$, this embodies that the respective correlation functions with $\tilde{V}^{ij}_\mu$ or $\tilde{A}^{ij}_\mu$ have to be finite and renormalized as well. Hence, while the point-split vector current $\tilde{V}^{ij}_\mu$ requires no renormalization, the point-split axial current $\tilde{A}^{ij}_\mu$ is renormalized by a finite and scale-independent factor, i.e. subject to a pure lattice renormalization that vanishes in the continuum (no-renormalization theorem [155, 156]):

$$Z_{\tilde{V}} \qquad = 1 \, , \qquad\qquad Z_{\tilde{A}}(g_0^2) = 1 + O(a) \, . \tag{C.9}$$

Furthermore, it may be argued [154] that the difference between the point-split currents and their respective local counterparts is finite, allowing to replace the former by the latter in (C.4) and (C.7), in conjunction with the finite lattice renormalizations constants (cf. (3.76))

$$Z_V(g_0^2) = 1 + O(a) \, , \qquad Z_A(g_0^2) = 1 + O(a) \, . \tag{C.10}$$

A *second implication* of (C.4) and (C.7) in the situation where field renormalization has rendered the left hand sides finite applies to the renormalization of the scalar and pseudoscalar density. The fact that the total derivative of the point-split vector or axial current vanishes after integration over x implies relations between the renormalization constants that appear on the respective finite right hand sides, namely

$$Z_m(g_0^2, a\mu) = Z_S(g_0^2, a\mu)^{-1} \tag{C.11a}$$

$$Z_{\overline{m}_{\mathrm{PCAC}}}(g_0^2, a\mu) = Z_P(g_0^2, a\mu)^{-1} \, , \tag{C.11b}$$

up to finite contributions. This behavior is reflected e.g. in (3.107) and (E.15).

A *third and immediate consequence* of (C.4) and (C.7) emerges in the special case of $x \neq x_1, x_2$, which leads to the *chiral lattice Ward identities*. For the vector current $(X = V)$, this is the *PCVC relation* (Partially Conserved Vector Current),

$$Z_V \, \tilde{\partial}_\mu \left\langle V^{ij}_\mu(x) \, O \right\rangle \qquad = (m_{0,i} - m_{0,j}) \left\langle S^{ij}(x) \, O \right\rangle \, . \tag{C.12a}$$

Application of (C.11a) allows to write this equation as[3]

$$\tilde{\partial}_\mu \left\langle (V_{\mathrm{R}})^{ij}_\mu(x)\; O_{\mathrm{R}} \right\rangle = (m^i_{\mathrm{R}} - m^j_{\mathrm{R}}) \cdot \left\langle (S_{\mathrm{R}})^{ij}(x)\; O_{\mathrm{R}} \right\rangle \;. \tag{C.12b}$$

The vector current is hence conserved for *degenerate masses*, $m^i_{\mathrm{R}} = m^j_{\mathrm{R}}$, and the relation is exactly fulfilled as the lattice regularization respects $SU(N_f)_{fV}$ symmetry.

For the axial current ($X = A$), the *PCAC relation* (Partially Conserved Axial Current) reads

$$Z_A\; \tilde{\partial}_\mu \left\langle A^{ij}_\mu(x)\; O \right\rangle \qquad = (\overline{m}^i_{\mathrm{PCAC}} + \overline{m}^j_{\mathrm{PCAC}}) \left\langle P^{ij}(x)\; O \right\rangle \; + O(a) \;. \tag{C.13a}$$

It may be written under use of (C.11b) in terms of renormalized quantities as

$$\tilde{\partial}_\mu \left\langle (A_{\mathrm{R}})^{ij}_\mu(x)\; O_{\mathrm{R}} \right\rangle = (m^i_{\mathrm{R}} + m^j_{\mathrm{R}}) \quad \cdot \quad \left\langle P^{ij}_{\mathrm{R}}(x)\; O_{\mathrm{R}} \right\rangle + O(a) \;. \tag{C.13b}$$

The PCAC relation is thus violated on the lattice, in accordance with the breaking of chiral symmetry by the Wilson term, cf. sec. 3.3.

Note that, although the explicitly given terms in (C.4), (C.5) and (C.7) refer to the special choice of the operator O taken above, the chiral lattice Ward identities (C.12) and (C.13) hold generally for any multilocal operator O with fields at sites different from x. Moreover, we wish to emphasize again that all of the above holds in the case of *non-singlet* currents ($j \neq i$). The singlet case, affected by the chiral ABJ anomaly (cf. sec. 2.2), can be considered separately in terms of an anomalous chiral Ward identity, which is however not of importance here.

---

[3] Note that even for non-degenerate masses, the *difference* of two bare masses renormalizes multiplicatively with $Z_m(g_0^2, a\mu)$. This can be readily seen from (3.66).

D Spurionic chiral symmetry

We have seen in the main text that chiral symmetry is broken in several ways. Beside the spontaneous breaking that goes along with the non-vanishing expectation value of the quark condensate (sec. 2.2), this happens explicitly on the level of the Lagrangian in a twofold way. First, by non-vanishing and even more so by non-degenerate quark masses (sec. 2.2), and second, by the Wilson regularization on the lattice (see sec. 3.3). In order to disentangle these two sources of explicit symmetry breaking, it is instructive to consider *spurionic chiral transformations*, under which the bare mass matrix (3.13) is assigned the behavior

$$\mathcal{M} \to \mathcal{M}' = V \ \mathcal{M} V^{\dagger} \ , \qquad V \in SU(N_f)_{fV} \tag{D.1}$$

$$\mathcal{M} \to \mathcal{M}' = A^{\dagger} \mathcal{M} A^{\dagger} \ , \qquad A \in {}^{''}SU(N_f)^{''}_{fA} \ . \tag{D.2}$$

Evidently, in combination with the transformations of the fields, (A.22), the mass term $\overline{\psi}\mathcal{M}\psi$ in the action then obeys a generalized, *spurionic chiral symmetry*, irrespective of the nature of the mass matrix $\mathcal{M}$. This allows to consider the effects of the Wilson term in isolation. Concretely, the imposition of (spurionic) vector symmetry (D.1) is relevant for the construction of the Symanzik effective theory in the case of non-degenerate, massive quarks, sec. 8.1, while the breaking of (spurionic) axial symmetry (D.2) has important consequences for the quark mass renormalization in the Wilson formulation, which will be discussed in app. E.

In preparation for the latter, we now study the behavior of the quark masses under spurionic chiral symmetry transformations for the specific case of $N_f = 4^*$ in more detail, following the lines of [153]. It is useful to decompose the bare mass matrix (3.13) into an $SU(4)_{fV}$ flavor singlet and non-singlet part [133]:

$$\mathcal{M} = m_{\text{av}} \cdot \mathbb{1}_4 + \sum_{c=1}^{15} \hat{m}^c \lambda^c \ . \tag{D.3}$$

This is a generalization of $\mathcal{M}$, from which the physical case can be retrieved by restricting the generalized Gell-Mann matrices in the sum to the diagonal ones, $c = 3, 8, 15$.

Using $\mathrm{Tr}[\lambda^a] = 0$ and the trace property (A.13), one finds

$$m_{\mathrm{av}} = \frac{1}{4}\mathrm{Tr}[\mathcal{M}] \tag{D.4}$$

$$\hat{m}^a = \frac{1}{2}\mathrm{Tr}[\mathcal{M}\lambda^a] \ . \tag{D.5}$$

Under infinitesimal spurionic vector transformations (D.1), the masses transform as

$$m_{\mathrm{av}} \to m_{\mathrm{av}} \tag{D.6}$$

$$\hat{m}^a \to \hat{m}^a - f^{abc}\alpha^b\hat{m}^c \ , \tag{D.7}$$

where the explicit form (A.21a) of $V \in SU(4)_{fV}$ with (A.20) and the commutator (A.3) of the Gell-Mann matrices were used. Under infinitesimal spurionic axial transformations (D.2) with $A \in {}''SU(4)''_{fA}$ as given in (A.21b), one finds

$$m_{\mathrm{av}} \to m_{\mathrm{av}} - \frac{2i}{4}\alpha^c\hat{m}^c \tag{D.8}$$

$$\hat{m}^a \to \hat{m}^a - d^{abc}\alpha^b\hat{m}^c - i\alpha^a m_{\mathrm{av}} \ , \tag{D.9}$$

Hence, under spurionic vector transformations, the average mass is a singlet, whereas the non-singlet mass components form an adjoint multiplet. In contrast, under spurionic axial transformations, the flavor singlet and non-singlet mass components transform into each other.

E Quark mass renormalization

The observed behavior of the generalized mass $\mathcal{M}$ under spurionic vector transformations, (D.6) and (D.7), and spurionic axial transformations, (D.8) and (D.9), has important consequences for the renormalization of quark masses in the Wilson regularization. The vector transformations leave room for different multiplicative renormalizations of the singlet mass m_{av} and the non-singlet mass components $\hat{m}^c$. Moreover, the invariance of m_{av} allows this term to mix with the identity $\mathbb{1}$, which makes it subject to an additive renormalization that manifests itself in the appearance of the *critical mass* m_{crit}:

$$\hat{m}^c_R \quad = Z_m(g_0^2, a\mu) \;\; \hat{m}^c \tag{E.1}$$

$$m_{R,\text{av}} = Z_{m^0}(g_0^2, a\mu) \;\; [m_{\text{av}} - m_{\text{crit}}] \;\; , \tag{E.2}$$

This rather complicated renormalization pattern would become obsolete if spurionic axial symmetry was respected, as the mixing of all the mass components under these transformations would enforce a uniform, multiplicative renormalization. However, since this spurionic axial symmetry is explicitly violated by the Wilson formulation of Lattice QCD, there is no constraint which prevents the masses from renormalizing according to (E.1) and (E.2).

E.1 Unimproved $N_f = 4^*$ theory

We consider the physical case again, where the bare mass matrix (3.13) is split into (cf. (D.3))

$$\mathcal{M} = m_{\text{av}} \cdot \mathbb{1}_4 + \sum_{c=3,8,15} \hat{m}^c \lambda^c \;, \tag{E.3}$$

i.e. the sum in the non-singlet part now only involves diagonal generators λ^c. The individual masses accordingly read

$$m_{0,i} = m_{\mathrm{av}} + \sum_{c=3,8,15} \hat{m}^c \lambda^c_{ii} \,. \tag{E.4}$$

Taking into account that the two parts renormalize differently, the renormalized quark mass of flavor i may be defined in terms of the matrix

$$M = \mathrm{diag}(m_{q,u}, m_{q,d}, m_{q,s}, m_{q,c}) \tag{3.64}$$

of bare subtracted quark masses

$$m_{q,i} = m_{0,i} - m_{\mathrm{crit}} \tag{3.65}$$

as follows [133]:

$$\begin{aligned}
m^i_{\mathrm{R}} &= m_{R,\mathrm{av}} + \sum_c \hat{m}^c_R \lambda^c_{ii} \\
&\overset{(E.1,E.2)}{=} Z_{m^0}(g_0^2, a\mu)\,[m_{\mathrm{av}} - m_{\mathrm{crit}}] + Z_m(g_0^2, a\mu) \sum_c \hat{m}^c \lambda^c_{ii} \\
&\overset{(E.4,3.65)}{=} Z_{m^0}(g_0^2, a\mu)\,[m_{\mathrm{av}} - m_{\mathrm{crit}}] + Z_m(g_0^2, a\mu)\left((m_{q,i} + m_{\mathrm{crit}}) - m_{\mathrm{av}}\right) \\
&= Z_m(g_0^2, a\mu)\left[\frac{Z_{m^0}(g_0^2, a\mu)}{Z_m(g_0^2, a\mu)}\,[m_{\mathrm{av}} - m_{\mathrm{crit}}] + \left(m_{q,i} - [m_{\mathrm{av}} - m_{\mathrm{crit}}]\right)\right] \\
&= Z_m(g_0^2, a\mu)\left[m_{q,i} + \left(\frac{Z_{m^0}(g_0^2, a\mu)}{Z_m(g_0^2, a\mu)} - 1\right)[m_{\mathrm{av}} - m_{\mathrm{crit}}]\right] \\
&\overset{(E.6)}{=} Z_m(g_0^2, a\mu)\left[m_{q,i} + \left(\frac{Z_{m^0}(g_0^2, a\mu)}{Z_m(g_0^2, a\mu)} - 1\right)\mathrm{Tr}[M]/N_f\right] \,,
\end{aligned} \tag{E.5}$$

where in the last step

$$m_{\mathrm{av}} - m_{\mathrm{crit}} = \left(\sum_i m_{0,i}\right)/N_f - m_{\mathrm{crit}} = \mathrm{Tr}[M]/N_f \,. \tag{E.6}$$

was used. The ratio of the flavor singlet and non-singlet renormalization constant is usually abbreviated by

$$r_m(g_0^2) = \frac{Z_{m^0}(g_0^2, a\mu)}{Z_m(g_0^2, a\mu)} \,, \tag{3.63}$$

such that (E.5) becomes

$$m_{\mathrm{R}}^i = Z_m(g_0^2, a\mu)\left[m_{q,i} + \left(r_m(g_0^2) - 1\right)\mathrm{Tr}\left[M\right]/N_f\right]\,. \tag{3.66}$$

E.2 Improved $N_f = 4^*$ theory

For the $O(a)$ improved expression of m_{R}^i, one has to replace $g_0^2 \to \widetilde{g}_0^2$ and $m_{q,i} \to \widetilde{m}_{q,i}$ (cf. (8.20, 8.22)) in (3.66):

$$m_{\mathrm{R}}^i = Z_m(\widetilde{g}_0^2, a\mu)\left[\widetilde{m}_{q,i} + \left(r_m(\widetilde{g}_0^2) - 1\right)\mathrm{Tr}\left[\widetilde{M}\right]/N_f\right]\,, \tag{E.7}$$

where (cf. (3.64)) $\widetilde{M} = \mathrm{diag}(\widetilde{m}_{q,u}, \widetilde{m}_{q,d}, \widetilde{m}_{q,s}, \widetilde{m}_{q,c})$. In that case, after a little algebra,

$$\begin{aligned}
\widetilde{m}_{q,i} \quad + &\quad \left(r_m(\widetilde{g}_0^2) - 1\right)\mathrm{Tr}\left[\widetilde{M}\right]/N_f \\
= &\left[m_{q,i} + ac_1 m_{q,i}^2 + ac_2\mathrm{Tr}[M]m_{q,i} + ac_3\mathrm{Tr}[M^2] + ac_4(\mathrm{Tr}[M])^2\right] + \frac{r_m(\widetilde{g}_0^2) - 1}{N_f}\times \\
&\quad \left[\mathrm{Tr}[M] + ac_1\mathrm{Tr}[M^2] + ac_2(\mathrm{Tr}[M])^2 + ac_3N_f\mathrm{Tr}[M^2] + ac_4N_f(\mathrm{Tr}[M])^2\right] \\
= m_{q,i} &\;+ \left(r_m(\widetilde{g}_0^2) - 1\right)\frac{\mathrm{Tr}[M]}{N_f} + a\left[c_1 m_{q,i}^2 + c_2\mathrm{Tr}[M]m_{q,i}\right. \\
&\;+ \left(c_3 + \frac{r_m(g_0^2) - 1}{N_f}\left(c_1 + N_f c_3\right)\right)\mathrm{Tr}[M^2] \\
&\;+ \left.\left(c_4 + \frac{r_m(g_0^2) - 1}{N_f}\left(c_2 + N_f c_4\right)\right)(\mathrm{Tr}[M])^2\right] \\
= m_{q,i} &\;+ \left(r_m(\widetilde{g}_0^2) - 1\right)\frac{\mathrm{Tr}[M]}{N_f} + a\left[c_1 m_{q,i}^2 + c_2\mathrm{Tr}[M]m_{q,i}\right. \\
&\;+ \left(r_m(g_0^2)\left(c_1 + N_f c_3\right) - c_1\right)\frac{\mathrm{Tr}[M^2]}{N_f} \\
&\;+ \left.\left(r_m(g_0^2)\left(c_2 + N_f c_4\right) - c_2\right)\frac{(\mathrm{Tr}[M])^2}{N_f}\right]\,,
\end{aligned} \tag{E.8}$$

one finds

$$m_{\mathrm{R}}^{i} = Z_m(\widetilde{g}_0^2, a\mu)\Big[m_{q,i} + \big(r_m(\widetilde{g}_0^2) - 1\big)\frac{\mathrm{Tr}[M]}{N_f} + a\cdot\Big\{b_m(g_0^2)m_{q,i}^2 + \overline{b}_m(g_0^2)\mathrm{Tr}[M]m_{q,i} + \Big(r_m(g_0^2)d_m(g_0^2) - b_m(g_0^2)\Big)\frac{\mathrm{Tr}[M^2]}{N_f} + \Big(r_m(g_0^2)\overline{d}_m(g_0^2) - \overline{b}_m(g_0^2)\Big)\frac{(\mathrm{Tr}[M])^2}{N_f}\Big\}\Big] . \tag{8.25}$$

The new coefficients are related to the old ones by

$$b_m = c_1 , \qquad d_m = c_1 + N_f c_3 , \tag{8.23a}$$

$$\overline{b}_m = c_2 , \qquad \overline{d}_m = c_2 + N_f c_4 . \tag{8.23b}$$

While the renormalized quark masses m_{R}^{i} on the lattice depend on the bare (improved) coupling $\widetilde{g}_0^2$ and the renormalization scale μ, their continuum extrapolated counterparts

$$\overline{m}_i(\mu) = \lim_{a\to 0} m_{\mathrm{R}}^{i}(\widetilde{g}_0^2, a\mu) \tag{E.10}$$

are related to the scale-independent continuum RGI masses M_i via the flavor-independent conversion factor (cf. (2.36))

$$h(\mu) = M/\overline{m}(\mu) , \tag{E.11}$$

which in particular cancels the μ-dependence of $\overline{m}_i$,

$$M_i = h(\mu)\cdot\overline{m}_i(\mu) . \tag{E.12}$$

E.3 Improved $N_f = 4$ and massless $N_f = 2$ theories

We now consider the renormalization of the quark mass, (8.25), for two special cases. On the one hand, the theory with $N_f = 4$ *degenerate* masses, cf. (3.84), and on the other hand, the *massless* $N_f = 2$ case, where in addition to the dynamical, massless up and down quarks, one treats the strange and charm quark in the quenched approximation, cf. tab. 9.1.

In $N_f = 4$, after use of $\mathrm{Tr}[M] = N_f m_q$, $\mathrm{Tr}[M^2] = N_f m_q^2$ and $(\mathrm{Tr}[M])^2 = N_f^2 m_q^2$, the b_m and $\overline{b}_m$ terms vanish, while the $\overline{d}_m$ term may be absorbed in the one with d_m. Hence, (8.25) reduces to

$$m_{\mathrm{R}} = Z_{m^0}(\tilde{g}_0^2, a\mu)\left[1 + ad_m(g_0^2)m_q\right] m_q \ . \tag{E.13}$$

In the case of $N_f = 2$, one has to differentiate in (8.25) between the effects of dynamical and valence quarks. The former are embodied in the appearance of the flavor trace of the bare subtracted mass matrix, which vanishes by definition, $\mathrm{Tr}[M] = 0$. Thus, the renormalization prescription reads

$$m_{\mathrm{R}}^i = Z_m(g_0^2, a\mu)\left[1 + ab_m(g_0^2)m_{q,i}\right] m_{q,i} \ , \tag{E.14}$$

for a valence quark of flavor i. Note that the use of the improved bare coupling $\tilde{g}_0^2$ as an argument of Z_m is not necessary, since it is equal to g_0^2 for massless sea quarks, see (3.96).

As we have seen in sec. 3.6, a renormalized mass may also be formulated in terms of the PCAC mass. With its improved version in $N_f = 4$ given in (3.107), for $N_f = 2$ we find

$$m_{\mathrm{R}}^i = \frac{Z_A(g_0^2)}{Z_P(g_0^2, a\mu)} \, \frac{1 + ab_A(g_0^2)\ m_{q,i}}{1 + ab_P(g_0^2)\ m_{q,i}} \, m_{\mathrm{PCAC,I}}^{i\tilde{i}}(x) + O(a^2) \ , \tag{E.15}$$

where the tilde notation indicates the involvement of *non-singlet* currents. The two equations (E.14) and (E.15) relate the improved PCAC mass $m_{\mathrm{PCAC,I}}^{i\tilde{i}}$ and the bare subtracted mass $m_{q,i}$ according to

$$m_{\mathrm{PCAC,I}}^{i\tilde{i}} = \frac{Z_m(g_0^2, a\mu)\ Z_P(g_0^2, a\mu)}{Z_A(g_0^2)} \left[1 + \left(b_m(g_0^2) + b_P(g_0^2) - b_A(g_0^2)\right) am_{q,i}\right] m_{q,i} + O(a^2) \ , \tag{E.16}$$

where in $N_f = 2$ the cases $i = s, c$ of the valence strange and charm quark are of particular interest. Finally, we note that the overall renormalization factor

$$Z(g_0^2) = \frac{Z_m(g_0^2, a\mu)\ Z_P(g_0^2, a\mu)}{Z_A(g_0^2)} \ , \tag{E.17}$$

as it appears in (E.16), is finite and independent of the renormalization scale μ, since the divergent parts of Z_m and Z_P cancel each other.

E.4 Improved $N_f = 4^*$ theory in a mass-dependent scheme

In order to perform the transition of the mass renormalization prescription from the mass-independent to the mass-dependent scheme, (8.25) may be reformulated so as to disentangle the sea quark effects (that come with a trace) from the flavor-specific terms (that come with $m_{q,i}$). Adding and subtracting two terms with constants f and $\overline{f}$, we write

$$\begin{aligned} m_{\mathrm{R}}^{i} = Z_m(\widetilde{g}_0^2, a\mu)\Bigg[& \Big(1 + af\, b_m(g_0^2) m_{q,i} + a\overline{f}\, \overline{b}_m(g_0^2)\mathrm{Tr}[M]\Big)\, m_{q,i} + \left(r_m(\widetilde{g}_0^2) - 1\right) \frac{\mathrm{Tr}[M]}{N_f} \\ & + a \cdot \Bigg\{ (1-f)\, b_m(g_0^2) m_{q,i}^2 + \Big(1 - \overline{f}\Big)\, \overline{b}_m(g_0^2)\mathrm{Tr}[M] m_{q,i} \\ & + \Big(r_m(g_0^2) d_m(g_0^2) - b_m(g_0^2)\Big) \frac{\mathrm{Tr}[M^2]}{N_f} + \Big(r_m(g_0^2)\overline{d}_m(g_0^2) - \overline{b}_m(g_0^2)\Big) \frac{(\mathrm{Tr}[M])^2}{N_f} \Bigg\} \Bigg], \end{aligned} \tag{E.18}$$

and factoring out leads to

$$\begin{aligned} m_{\mathrm{R}}^{i} = Z_m(\widetilde{g}_0^2, a\mu) & \Big(1 + af\, b_m(g_0^2) m_{q,i} + a\overline{f}\, \overline{b}_m(g_0^2)\mathrm{Tr}[M]\Big) \Bigg[m_{q,i} + \left(r_m(\widetilde{g}_0^2) - 1\right) \frac{\mathrm{Tr}[M]}{N_f} \\ & + a \cdot \Bigg\{ (1-f)\, b_m(g_0^2) m_{q,i}^2 + \Big(1 - \overline{f} + f\Big)\, \overline{b}_m(g_0^2)\mathrm{Tr}[M] m_{q,i} \\ & + \Big(r_m(g_0^2) d_m(g_0^2) - b_m(g_0^2)\Big) \frac{\mathrm{Tr}[M^2]}{N_f} \\ & + \Big(r_m(g_0^2) \Big(\overline{d}_m(g_0^2) - \overline{f}\, \overline{b}_m(g_0^2)\Big) + \Big(\overline{f} - 1\Big)\, \overline{b}_m(g_0^2)\Big) \frac{(\mathrm{Tr}[M])^2}{N_f} \Bigg\} \Bigg]. \end{aligned} \tag{E.19}$$

In order to make the terms in the second line disappear, $f = 1$ and $\overline{f} = 2$ is chosen, and all the remaining $O(a)$ terms may in analogy to (8.49) be absorbed in the mass-dependent critical mass m_{crit}^{M} (8.48), which is considered in the notation $\widetilde{m}_{\mathrm{crit}}^{M}$. Additional factorization of the term in the first line leads to

$$\begin{aligned} m_{\mathrm{R}}^{i} & = Z_m(\widetilde{g}_0^2, a\mu) \Big(1 + 2a\overline{b}_m(g_0^2)\mathrm{Tr}[M]\Big) \left(1 + ab_m(g_0^2) m_{q,i}\right) \Big[m_{0,i} - \widetilde{m}_{\mathrm{crit}}^{M}(g_0^2, \mathrm{Tr}\,[M]) \Big] \\ & = \widetilde{Z}_m(g_0^2, a\mathrm{Tr}\,[M]\,, a\mu) \left(1 + ab_m(g_0^2) m_{q,i}\right) \Big[m_{0,i} - \widetilde{m}_{\mathrm{crit}}^{M}(g_0^2, \mathrm{Tr}\,[M]) \Big] \end{aligned} \tag{8.50a}$$

where in the last step, the sea quark contributions were absorbed in the renormalization constant $\widetilde{Z}_m(g_0^2, a\mathrm{Tr}\,[M]\,, a\mu)$ in a similar way to (8.44, 8.45).

F Previous non-perturbative determinations of c_{sw} and c_A

F.1 c_{sw}

Previous non-perturbative results of the ALPHA Collaboration on the determination of c_{sw} are available for the Wilson gauge action in $N_f = 0$ [16], $N_f = 2$ [17], $N_f = 4$ [21], and for the Lüscher–Weisz gauge action in $N_f = 3$ [22]. All these works deal with massless quarks and in principle follow the standard procedure discussed in sec. 9.4, with a few crucial differences to be described in this appendix.

Fixed L/a approach

The principal difference to the strategy described in the main text (cf. sec. 9.1 and sec. 9.4) is the fact that, instead of the lattice extent L (cf. (9.7a)), it is the lattice resolution that is fixed to $L/a = 8$, with $T = 2L$ or $T = 2L - a$, see fig. F.1.

This *fixed L/a approach* comes with a few advantages. First, in opposition to the restrictions imposed on g_0^2 by the line of constant physics (LCP), the values of the bare coupling at which c_{sw} is determined may be chosen at one's discretion. Second, in

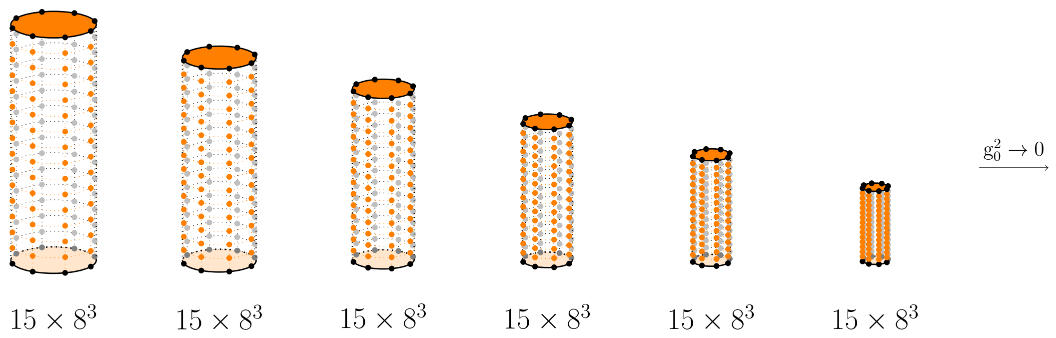


Figure F.1: Fixed L/a and $T = 2L - a$ as used in the determination of c_{sw} in $N_f = 3$ with massless quarks [22], to be compared with the LCP approach illustrated in fig. 9.8 in sec. 9.4.

conjunction with the use of massless quarks[1], it makes the setting of a scale obsolete. It also conveniently avoids the problem of critical slowing down (cf. sec. 6.2.2) for small g_0^2. Finally, as discussed in sec. 9.4, the slope in ΔM^{ud} (cf. 9.115a) as a function of c_{sw} is larger in small volumes, leading to a clearer signal. This also simplifies the determination of c_{sw} at small bare couplings.

However, as a matter of principle, the fixed L/a approach gives rise to potentially large $O(a)$ ambiguities in c_{sw}, see e.g. [92]. Hence, c_{sw} is properly defined only at a LCP, and the fixed L/a approach has to be considered an approximation. Its validity is usually checked by a single simulation at $L/a = 16$, where the weakness of the dependence of c_{sw} on L/a is verified.

Improvement condition

Instead of the LCP (9.7), it is the *improvement point*

$$M^{ud}(T/2, T/4) \overset{!}{=} 0 \,, \tag{F.1}$$

defined in terms of the effective quark mass (9.113), that determines the hopping parameter κ, of which there is only one due to the degeneracy. The *improvement condition* is formulated very similarly to (9.115) in terms of

$$\Delta M^{ud}(3T/4, T/4) = M^{ud}(3T/4, T/4) - M'^{ud}(3T/4, T/4) \,. \tag{F.2}$$

However, instead of requiring this quantity to vanish, one fixes it to its tree-level value

$$\Delta M^{ud}(3T/4, T/4) \overset{!}{=} \Delta M^{ud(0)}(3T/4, T/4) \,, \tag{F.3}$$

to ensure $c_{\mathrm{sw}}(0) = 1$. c_{sw} is only weakly dependent on a possible mismatch in (F.1). Hence, a rough tuning of κ to the improvement point is sufficient.

The work on $N_f = 3$ is special concerning two aspects. First of all, as mentioned before, the Lüscher–Weisz gauge action (3.112) is used instead of the Wilson plaquette gauge action (3.39). Secondly, for technical reasons that is the use of version[2] 1.0 of the `openQCD` code, lattices of temporal size $T = 2L - a$ were employed, such that—since

---

[1] These can be tuned without the knowledge of a scale, cf. (F.1).

[2] We remind the reader that version 1.2 is used in the present work, see sec. 9.5.

Ref.	S_G	N_f	result	range of validity
[16]	S_G^{W}	0	$c_{\mathrm{sw}}(g_0^2) = \frac{1-0.656g_0^2-0.152g_0^4-0.054g_0^6}{1-0.922g_0^2}$	$0 \leq g_0^2 \leq 1$
[17]	S_G^{W}	2	$c_{\mathrm{sw}}(g_0^2) = \frac{1-0.454g_0^2-0.175g_0^4+0.012g_0^6+0.045g_0^8}{1-0.720g_0^2}$	$\beta \geq 5.2$
[22]	S_G^{LW}	3	$c_{\mathrm{sw}}(g_0^2) = \frac{1-0.1921g_0^2-0.1378g_0^4+0.0717g_0^6}{1-0.3881\ g_0^2}$	$\beta \geq 3.3$
[21]	S_G^{W}	4	$c_{\mathrm{sw}}(g_0^2) = \frac{1-0.1372g_0^2-0.1641g_0^4+0.1679g_0^6}{1-0.4031g_0^2}$	$0 \leq g_0^2 \leq 1.2$

Table F.1: Results for $c_{\mathrm{sw}}(g_0^2)$ from previous works.

$T/2$ is not an integer—instead of (F.1),

$$\frac{1}{2}\left(M(L, L/2) + M(L-a, L/2)\right) = 0 \tag{F.4}$$

is used.

Results

The results of the previous determinations of c_{sw} are compiled in tab. F.1 and partly illustrated along with the main result of the present work in fig. 9.16. The formulas incorporate the perturbative behavior of c_{sw}, given by (3.113a) and (3.113c), in a fashion very similar to (9.126).

For completeness, we note that further results on c_{sw} exist, namely in connection with the Wilson gauge action S_G^{W} and $N_f = 3$ [18], the so-called RG-improved gauge action and $N_f = 0, 2, 3$ [19], as well as $N_f = 3$ stout-smeared fermions and the tree-level improved Lüscher–Weisz gauge action [20].

F.2 c_A

Within the ALPHA Collaboration, non-perturbative results on c_A were obtained for massless quarks and the Wilson gauge action in $N_f = 0$ [16], $N_f = 2$ [54], and the Lüscher–Weisz gauge action in $N_f = 3$ [141].

In all of those cases, the associated non-perturbative results for c_{sw} were employed, see app. F.1, but more suitable improvement conditions on which c_A is more sensitive than the one designed for c_{sw} (cf. sec. 9.4) were chosen. In particular, these incorporate vanishing boundary fields (cf. (9.14)). In the case of $N_f = 0$, the improvement condition entails a variation of the fermion angle θ (cf. (5.21)) within the fixed L/a approach (cf. app. F.1). The works on $N_f = 2$ and $N_f = 3$ differ from this, as they both feature a line of constant physics, which fixes the volume of the SF. In addition, based on the work of [157], the improvement condition in those cases is formulated in terms of a variation of the boundary states, designed such that exclusively either the pseudoscalar ground or first excited state contributes. Finally, we note that in the case of $N_f = 3$, the results were projected onto the trivial topological sector, due to topology freezing for small lattice spacings, in a fashion very similar to the one described in sec. 7. We refer to the original publications and (3.116) for further details and numerical results. For completeness, we remark that c_A was also non-perturbatively determined in the case of $N_f = 3$ and the so-called Iwasaki gauge action [158].

G Effective meson masses on the lattice

The quantities

$$\Phi_2 = T \cdot \Gamma^{su} \tag{9.23a}$$

$$\Phi_3 = T \cdot \left(\Gamma^{sc} - \frac{1}{2}\Gamma^{su} \right) , \tag{9.23b}$$

are defined in terms of the effective meson masses

$$\Gamma^{ij} \equiv \Gamma^{ij}_{A} = -\tilde{\partial}_0 \log \left(f^{ij}_{A,\mathrm{I}}(x_0) \right) \big|_{x_0=T/2} , \quad i,j \in \{l,c\} , \tag{9.20}$$

with

$$f^{ij}_{A,\mathrm{I}}(x_0) = f^{ij}_{A}(x_0) + ac_A \tilde{\partial}_0 f^{ij}_{P}(x_0) . \tag{5.53}$$

G.1 Discretizations of Γ

The symmetric derivative $\tilde{\partial}_0$ (cf. (A.30)) in (9.20) can either be applied to the logarithmic function as a whole, or on $f^{ij}_{A,\mathrm{I}}$ after using the chain rule. We label the two resulting discretizations $\Gamma^{ij}_{(1)}$ and $\Gamma^{ij}_{(2)}$, respectively:

$$\Gamma^{ij}_{(1)} = -\frac{1}{2a} \log \left(\frac{f^{ij}_{A,\mathrm{I}}(x_0+a)}{f^{ij}_{A,\mathrm{I}}(x_0-a)} \right) \Bigg|_{x_0=T/2} \tag{9.26}$$

$$\Gamma^{ij}_{(2)} = -\frac{\tilde{\partial}_0 f^{ij}_{A,\mathrm{I}}(x_0)}{f^{ij}_{A,\mathrm{I}}(x_0)} \Bigg|_{x_0=T/2} . \tag{G.1}$$

The explicit expressions can easily be obtained under use of (A.30) and (5.53). They read

$$\Gamma^{ij}_{(1)} = -\frac{1}{2a}\log\left(\frac{f_A^{ij}(\frac{T}{2}+a)+\frac{c_A}{2}\left[f_P^{ij}(\frac{T}{2}+2a)-f_P^{ij}(\frac{T}{2})\right]}{f_A^{ij}(\frac{T}{2}-a)+\frac{c_A}{2}\left[f_P^{ij}(\frac{T}{2})-f_P^{ij}(\frac{T}{2}-2a)\right]}\right)\Bigg|_{x_0=T/2} \tag{G.2}$$

and

$$\Gamma^{ij}_{(2)} = -\frac{1}{2a}\cdot\frac{\left[f_A^{ij}(\frac{T}{2}+a)-f_A^{ij}(\frac{T}{2}-a)\right]+2c_A\left[f_P^{ij}(\frac{T}{2}+a)-2f_P^{ij}(\frac{T}{2})+f_P^{ij}(\frac{T}{2}-a)\right]}{f_A^{ij}(\frac{T}{2})+\frac{c_A}{2}\left[f_P^{ij}(\frac{T}{2}+a)-f_P^{ij}(\frac{T}{2}-a)\right]}\ . \tag{G.3}$$

Note that in (G.3), the alternative discretization $\partial_0^*\partial_0$ (cf. (A.32)) of the second derivative has been used instead of $\tilde{\partial}_0^2$.

The difference between the two quantities $\Gamma^{ij}_{(1)}$ and $\Gamma^{ij}_{(2)}$ is a pure lattice artifact. On the one hand, they both inherit discretization errors from the correlation function $f^{ij}_{A,\mathrm{I}}$ (cf. (9.21))—which themselves stem from the discretized action and the operator $\mathcal{O}^{ji}$ (cf. (5.46)). On the other hand, the respective discretizations of the logarithmic derivative introduce further cutoff effects. The latter can be investigated in isolation by ignoring the former in the approximation of no excited states (cf. (9.21),(9.22)), that is by using

$$f^{ij}_{A,\mathrm{I}}(x_0) \sim e^{-m^{ij}_{\mathrm{PS}}\cdot x_0}\ . \tag{G.4}$$

In doing so, we find

$$\begin{aligned}\Gamma^{ij}_{(1)} &= -\frac{1}{2a}\log\left(\frac{f_A^{ij}(x_0+a)}{f_A^{ij}(x_0-a)}\right)\Bigg|_{x_0=T/2} \overset{(G.4)}{=} -\frac{1}{2a}\log\left(e^{-2am^{ij}_{\mathrm{PS}}}\right)\\ &= m^{ij}_{\mathrm{PS}}\end{aligned} \tag{G.5}$$

and

$$\begin{aligned}\Gamma^{ij}_{(2)} &= -\frac{\tilde{\partial}_0 f_A^{ij}(x_0)}{f(x_0)}\Bigg|_{x_0=T/2} \overset{(G.4)}{=} \frac{1}{a}\sinh\left(am^{ij}_{\mathrm{PS}}\right)\\ &= m^{ij}_{\mathrm{PS}}\left(1+\frac{1}{6}(am^{ij}_{\mathrm{PS}})^2+O(a^4)\right)\ .\end{aligned} \tag{G.6}$$

L_2/a	ens.	N_{cf}	$L_2\Gamma^{s\tilde{s}}_{(2)}$	$L_2\Gamma^{ud}_{(2)}$	$L_2\Gamma^{s\tilde{s}}_{(2),\mathrm{cor}}$	τ_{int}	Q_{fit}
12	C_{12}	500	1.42(7)	0.76(6)	0.66(1)	5(1)	
16	C_{16}	500	0.80(6)	0.11(5)	0.69(1)	7(2)	
20	C_{20}	180	0.51(4)	-0.15(4)	0.66(2)	9(4)	
24	C_{24}	500	0.79(3)	0.19(3)	0.60(1)	17(6)	
32	C_{32}	180	0.48(2)	-0.14(2)	0.62(1)	8(3)	
∞ without $L_2/a = 12$					0.59(1)		0.01
∞ with $L_2/a = 12$					0.61(1)		0.00

Table G.1: Results of the $N_f = 2$ runs to determine $L_2\Gamma^{s\tilde{s}}_{(2),\mathrm{cor}}$ at $z_l^\star = L_2 M_l^\star$ under use of the alternative discretization $\Gamma^{ij}_{(2)}$ (cf. (G.1)). See tab. 9.5 for further explanations and comparison to the results obtained with $\Gamma_{ij} = \Gamma^{ij}_{(1)}$.

Hence, the discretization of the logarithmic derivative as used in $\Gamma^{ij}_{(1)}$ introduces no explicit errors in the absence of excited states, as opposed to the one in $\Gamma^{ij}_{(2)}$. For this reason, we use

$$\Gamma^{ij} \equiv \Gamma^{ij}_{(1)} \tag{G.7}$$

as mentioned in the main text, cf. (9.26). The relative cutoff effects between the two discretizations, apparent in (G.5) and (G.6), will be investigated explicitly in the $N_f = 2$ theory in app. G.2.

G.2 Results for $\Phi_2^\star$ and $\Phi_3^\star$ under use of $\Gamma_{(2)}$

The quantities $\Phi_2^\star$ and $\Phi_3^\star$ have been determined in sec. 9.2.4 and sec. 9.2.5 as the continuum extrapolated values of the specific effective meson masses given in (9.28b) and (9.28c), respectively, under use of the standard discretization $\Gamma^{ij} = \Gamma^{ij}_{(1)}$ (cf. (9.26)). We here provide the results obtained with the alternative discretization $\Gamma^{ij}_{(2)}$ (cf. (G.1)), and check for agreement between the two definitions in the continuum.

The light EMM $\Gamma^{s\tilde{s}}_{(2),\mathrm{cor}}$ and $\Phi^\star_{2,(2)}$

The results for $L_2\Gamma^{s\tilde{s}}_{(2),\mathrm{cor}}$ by use of the discretization $\Gamma^{ij}_{(2)}$ are listed in tab. G.1. The difference from the standard lattice discretization $\Gamma_{ij} = \Gamma^{ij}_{(1)}$ is insignificant. Accordingly,

L_2/a	ens.	N_{cf}	$L_2\Gamma^{sc}_{(2)}$	$L_2\left(\Gamma^{sc}_{(2),\mathrm{cor}} - \frac{1}{2}\Gamma^{(2),\mathrm{cor}}_{s\bar{s}}\right)$	τ_{int}	Q_{fit}	$L_2\Gamma^{sc}_{(2-1),\mathrm{cor}}$
12	C_{12}	500	7.95(8)	6.86(4)	4(1)		0.553(15)
16	C_{16}	500	7.20(10)	6.75(5)	7(2)		0.265(9)
20	C_{20}	180	6.69(8)	6.51(7)	9(4)		0.145(6)
24	C_{24}	500	6.73(6)	6.24(5)	18(7)		0.096(3)
32	C_{32}	180	6.33(5)	6.16(4)	13(5)		0.051(3)
∞ without $L_2/a = 12$				5.94(5)		0.36	
∞ with $L_2/a = 12$				6.10(4)		0.00	

Table G.2: *Left:* Results of the $N_f = 2$ runs to determine $L_2(\Gamma^{sc}_{(2),\mathrm{cor}} - \frac{1}{2}\Gamma^{s\bar{s}}_{(2),\mathrm{cor}})$ at $z_l^\star = L_2 M_l^\star$ and $z_c^\star = L_2 M_c^\star$ under use of the alternative discretization $\Gamma^{ij}_{(2)}$ (cf. (G.1)). See tab. 9.6 for further explanations and comparison to the results obtained with $\Gamma^{ij} = \Gamma^{ij}_{(1)}$. *Right:* Results for $L_2\Gamma^{sc}_{(2-1),\mathrm{cor}}$, which incorporates the relative cutoff effects of the two discretizations, cf. (G.10) and fig. G.1.

the continuum extrapolation leads to

$$\Phi^\star_{2,(2)} = 0.59(1)\ , \tag{G.8}$$

which is exactly the same as $\Phi_2^\star$ (cf. (9.57)).

The charm EMM $\Gamma^{sc}_{(2),\mathrm{cor}}$ and $\Phi^\star_{3,(2)}$

Similarly, the results for $L_2\left(\Gamma^{sc}_{(2),\mathrm{cor}} - \Gamma^{s\bar{s}}_{(2),\mathrm{cor}}\right)$ by use of the discretization $\Gamma^{ij}_{(2)}$ are listed in tab. G.2, and illustrated together with those obtained with $\Gamma_{ij} = \Gamma^{ij}_{(1)}$ in the left panel of fig. G.1.

The continuum extrapolated value

$$\Phi^\star_{3,(2)} = 5.94(5) \tag{G.9}$$

is again well in accordance with $\Phi_3^\star$, see (9.61), as expected. However, in this case, the results on the lattice differ significantly for the two definitions. This comes as no surprise as the respective cutoff effects scale with the masses of the valence quarks, or equivalently, the pseudoscalar mass m^{ij}_{PS} in the approximation of no excited states. This is true also for the difference between $\Gamma^{ij}_{(1)}$ and $\Gamma^{ij}_{(2)}$, which by comparison of (G.5)

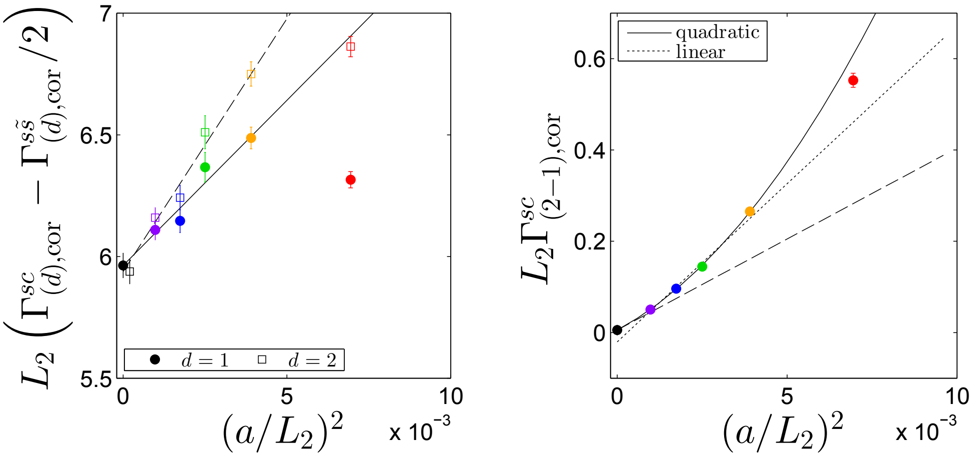


Figure G.1: *Left:* $L_2\left(\Gamma^{sc}_{(d),\mathrm{cor}} - \frac{1}{2}\Gamma^{s\tilde{s}}_{(d),\mathrm{cor}}\right)$ for the two discretizations $d = 1, 2$ of the effective meson masses, as listed in tab. 9.6 and tab. G.2, respectively. The fits, linear in $(a/L_2)^2$, are illustrated by the solid and the dashed line for $d = 1$ and $d = 2$, respectively. The continuum extrapolated value $\Phi^\star_{3,(2)}$ for $d = 2$ (shown as the empty black squared symbol) is slightly shifted to the right for convenience. *Right:* Difference $L_2\Gamma^{sc}_{(2-1),\mathrm{cor}}$ (cf. (G.10) of the two discretizations of the effective meson masses. The solid line illustrates the fit given in (G.11), quadratic in $(a/L_2)^2$, while the dashed line represents the associated tangent with the slope parameter a_1. The dotted line shows a fit linear in $(a/L_2)^2$. The data for $L_2/a = 12$ (red) is not taken into account in any of the above fits.

and (G.6) amounts to

$$L_2\Gamma^{ij}_{(2-1)} \equiv L_2\Gamma^{ij}_{(2)} - L_2\Gamma^{ij}_{(1)} = \frac{1}{6}\left(L_2 m^{ij}_{\mathrm{PS}}\right)^3 \left(\frac{a}{L_2}\right)^2 + O(a^4)\;. \tag{G.10}$$

The relative cutoff effects between the two discretizations are thus quite small in the case of the light quarks that enter $L_2\Gamma^{s\bar{s}}_{(d),\mathrm{cor}}$, but larger if the charm is involved as in the case of the quantity $L_2(\Gamma^{sc}_{(d),\mathrm{cor}} - \frac{1}{2}\Gamma^{s\bar{s}}_{(d),\mathrm{cor}})$. In the right panel of fig. G.1, the relative cutoff effects (G.10) are displayed for $L_2\Gamma^{sc}_{(d)}$. We find that the data may indeed well be described by a fit *quadratic* in $(a/L_2)^2$:

$$L_2\Gamma^{ij}_{(2-1)} = a_0 + a_1(a/L_2)^2 + a_2(a/L_2)^4 \tag{G.11a}$$

with

$$a_0 = 0.005(11)\;, \quad a_1 = 40(11)\;, \quad a_2 = 68(25)\cdot 10^2\;. \tag{G.11b}$$

The slope parameter $a_1 = \frac{1}{6}\left(L_2 m^{ij}_{\mathrm{PS}}\right)^3$ of this fit together with the approximately known scale $L_2 \approx 4\ \mathrm{GeV}^{-1}$ (cf. (9.44a)) allows to estimate the mass

$$m^{sc}_{\mathrm{PS}} \approx 1.55\ \mathrm{GeV} \tag{G.12}$$

of a pseudoscalar meson with a light and a charm valence (anti-)quark in $N_f = 2$. This serves as a cross check, as it is roughly of the order of the masses [26] of the D-mesons $D^0(1865)$, $D^+(1870)$ and $D^+_s(1968)$, whose valence quark content is $c\bar{u}$, $c\bar{d}$ and $c\bar{s}$, respectively. If the relative cutoff effects are modelled by a fit *linear* in $(a/L_2)^2$ (shown as the dotted line in fig. G.1), the same procedure yields an even better value of

$$m^{sc}_{\mathrm{PS}} \approx 1.86\ \mathrm{GeV}\;. \tag{G.13}$$

Note, however, that the quality of the fit in that case is worse as the curvature obviously present in the data is not taken into account, and the continuum extrapolated value is not compatible with zero as for the quadratic fit (cf. a_0 in (G.11b)).

G.3 Mass dependence of the effective meson masses Γ_A, Γ_P in $N_f = 2$

Since we know the relation between the RGI masses and the hopping parameters in $N_f = 2$ (cf. sec. 9.2.2), this theory allows to study the dependence of the quantities

$$L_2 \cdot \Gamma^{s\bar{s}} \quad \text{and} \quad L_2 \cdot \left(\Gamma^{sc} - \frac{1}{2}\Gamma^{s\bar{s}} \right) , \tag{G.14}$$

in terms of which $\Phi_2^\star$ and $\Phi_3^\star$ are defined (cf. (9.28)), on the light and charm quark masses. As explained in sec. 9.1, the definition of the effective meson masses incorporates the axial current,

$$\Gamma^{ij} \equiv \Gamma_A^{ij} = -\tilde{\partial}_0 \log \left(f_{A,\mathrm{I}}^{ij}(x_0) \right) \big|_{x_0=T/2} , \quad i,j \in \{l,c\} , \tag{9.20}$$

and we will compare this to the use of the pseudoscalar density in the similar quantity

$$\Gamma_P^{ij} = -\tilde{\partial}_0 \log \left(f_P^{ij}(x_0) \right) \big|_{x_0=T/2} , \quad i,j \in \{l,c\} . \tag{G.15}$$

To this end, we compute both Γ_A^{ij} and Γ_P^{ij} under use of the $N_f = 2$ ensembles (cf. sec. 9.2.1) at the dimensionless RGI masses $z_i = L_2 M_i$ in the combinations

$$(z_l, z_c) \quad \text{with} \quad z_l < z_c \tag{G.16}$$

and

$$z_l, z_c \in \{0, z_l^\star, 3z_l^\star, 0.5z_c^\star, 0.9z_c^\star, z_c^\star, 1.1z_c^\star\} . \tag{G.17}$$

The results for the respective continuum extrapolated values are given in tab. G.3 and depicted in fig. G.2.

We find that the deviation of Γ_A and Γ_P plays a role only if both the masses of the involved valence quarks are small, i.e. z_l in the case of $\Gamma_{s\bar{s}}$. Once there is a heavy (charm-like) quark involved, the difference becomes insignificant. This is in compliance with the behavior expected from the discussion in sec. 9.1, see (9.24). The dependence of the quantities (G.14) on the dimensionless RGI masses z_l and z_c under use of the pseudoscalar density, $\Gamma \equiv \Gamma_P$, may very well be described by linear functions. An effective description of the data for $\Gamma_A^{s\bar{s}}$ at small z_l is achieved if the deviation from $\Gamma_P^{s\bar{s}}$ is modelled by a third order polynomial fixed to the values $\Gamma_A^{s\bar{s}}(z_l = 0) = 0$

z_l	$\lim_{a\to 0} L_2\Gamma_A^{s\tilde{s}}$	Q_{fit}	$\lim_{a\to 0} L_2\Gamma_P^{s\tilde{s}}$	Q_{fit}
0	0		2.07(12)	0.84
$z_l^\star$	0.59(1)	0.01	2.30(11)	0.73
$3z_l^\star$	1.60(2)	0.04	2.79(10)	0.52
$0.5z_c^\star$	6.17(5)	0.31	6.29(7)	0.04
$0.9z_c^\star$	9.82(5)	0.12	9.82(5)	0.00
$z_c^\star$	10.69(5)	0.07	10.69(5)	0.00

z_l	$\lim_{a\to 0} L_2\left(\Gamma_A^{sc}-\frac{1}{2}\Gamma_A^{s\tilde{s}}\right)$	Q_{fit}	$\lim_{a\to 0} L_2\left(\Gamma_P^{sc}-\frac{1}{2}\Gamma_P^{s\tilde{s}}\right)$	Q_{fit}
0	6.10(5)	0.22	5.28(5)	0.01
$z_l^\star$	5.96(5)	0.30	5.30(4)	0.01
$3z_l^\star$	5.76(4)	0.37	5.31(4)	0.01
$0.5z_c^\star$	5.39(3)	0.13	5.30(3)	0.00
$0.9z_c^\star$	5.35(3)	0.07	5.29(3)	0.00
$z_c^\star$	5.35(3)	0.07	5.29(3)	0.00

z_c	$\lim_{a\to 0} L_2\left(\Gamma_A^{sc}-\frac{1}{2}\Gamma_A^{s\tilde{s}}\right)$	Q_{fit}	$\lim_{a\to 0} L_2\left(\Gamma_P^{sc}-\frac{1}{2}\Gamma_P^{s\tilde{s}}\right)$	Q_{fit}
$z_l^\star$	0.29(1)	0.01	1.10(5)	0.76
$3z_l^\star$	0.83(1)	0.03	1.33(5)	0.53
$0.5z_c^\star$	3.46(4)	0.21	3.07(4)	0.04
$0.9z_c^\star$	5.49(5)	0.30	4.85(4)	0.01
$z_c^\star$	5.96(5)	0.30	5.30(4)	0.01
$1.1z_c^\star$	6.43(5)	0.30	5.73(4)	0.01

Table G.3: Dependence of $\lim_{a\to 0} L_2\Gamma_\phi^{s\tilde{s}}$ on the light quark mass z_l *(top)*, $\lim_{a\to 0} L_2\left(\Gamma_\phi^{sc}-\frac{1}{2}\Gamma_\phi^{s\tilde{s}}\right)$ on the light quark mass z_l *(middle)* and $\lim_{a\to 0} L_2\left(\Gamma_\phi^{sc}-\frac{1}{2}\Gamma_\phi^{s\tilde{s}}\right)$ on the charm quark mass z_c *(bottom)*. All quantities are shown for $\phi = A$ and $\phi = P$ and with the respective other quark mass held fixed to $z_i = z_i^\star$ with $i = l$ or $i = c$. Q_{fit} gives the quality of the continuum extrapolations, which in the case of $\phi = A$ were conducted under use of the corrected quantity $\Gamma_A^{ij} \equiv \Gamma_{\mathrm{cor}}^{ij}$ (cf. (9.55)). Some errors of the results for $\phi = P$ are assumingly underestimated due to the poor fit quality, especially for large masses.

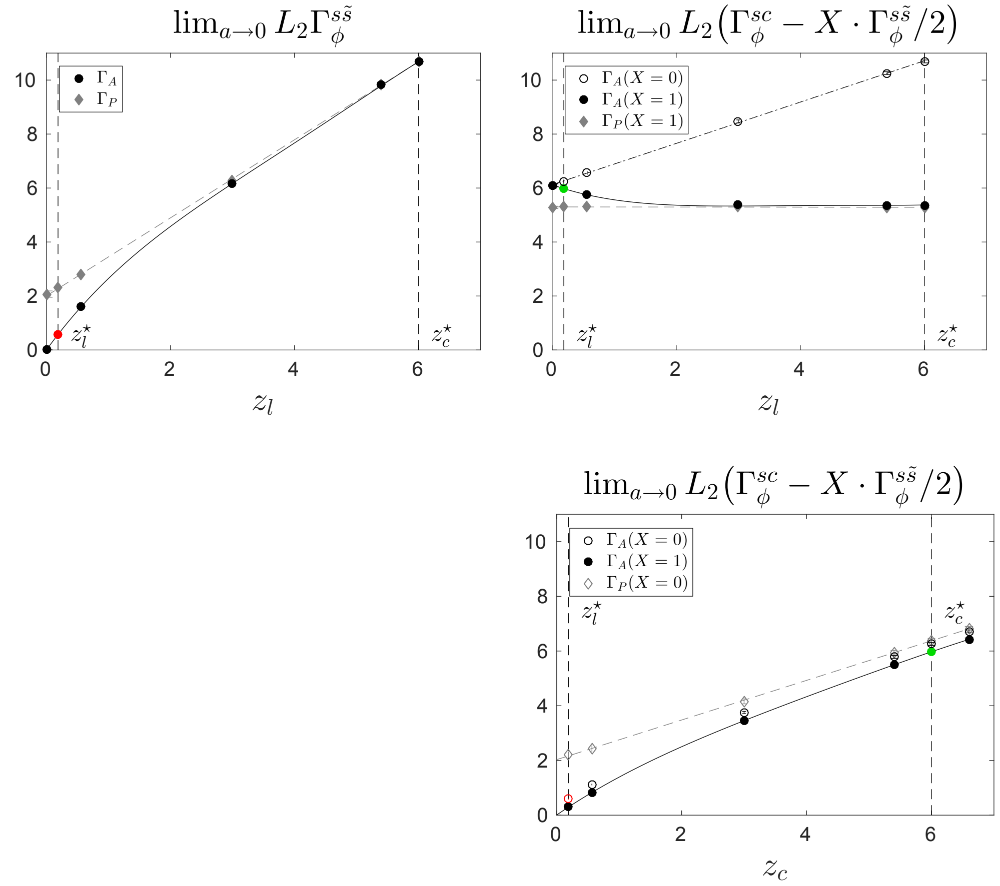


Figure G.2: Continuum extrapolated values of $L_2\Gamma_\phi^{s\tilde{s}}$ *(left)* as well as $L_2(\Gamma_\phi^{sc} - \frac{1}{2}\Gamma_\phi^{s\tilde{s}})$ and $L_2\Gamma_\phi^{sc}$ *(right)* for $\phi = A, P$ as a function of the light mass $z_l = L_2 M_l$ of the strange-like quark flavor s *(top)* and the mass $z_c = L_2 M_c$ of the charm-like quark flavor *(bottom)*. The respective other quark mass is held fixed to $z_c^\star$ *(top)* and $z_l^\star$ *(bottom)*. The points that correspond to $\Phi_2^\star$ and $\Phi_3^\star$ are highlighted in red and green, respectively.

ϕ	A	P
$\frac{\partial}{\partial z_l}\left(\lim_{a\to 0} L_2\Gamma_\phi^{s\tilde{s}}\right)\Big\rvert_{z_l=z_l^\star,z_c=z_c^\star}$	2.95(3)	1.45(1)
$\frac{\partial}{\partial z_l}\left(\lim_{a\to 0} L_2\left(\Gamma_\phi^{sc}-\frac{1}{2}\Gamma_\phi^{s\tilde{s}}\right)\right)\Big\rvert_{z_l=z_l^\star,z_c=z_c^\star}$	-0.71(2)	-0.002(6)
$\frac{\partial}{\partial z_c}\left(\lim_{a\to 0} L_2\left(\Gamma_\phi^{sc}-\frac{1}{2}\Gamma_\phi^{s\tilde{s}}\right)\right)\Big\rvert_{z_l=z_l^\star,z_c=z_c^\star}$	0.78(1)	0.73(1)

Table G.4: Dependence of the quantities $L_2\Gamma_\phi^{s\tilde{s}}$ and $L_2\left(\Gamma_\phi^{sc}-\frac{1}{2}\Gamma_\phi^{s\tilde{s}}\right)$ on the dimensionless RGI masses z_l and z_c at the points $z_l = z_l^\star$, $z_c = z_c^\star$ that define the LCP, under use of the 0^{th} component of the axial current ($\Gamma \equiv \Gamma_A$) and the pseudoscalar density ($\Gamma \equiv \Gamma_P$).

and $\Gamma_A^{s\tilde{s}}(z_l = z_c^\star) = \Gamma_P^{s\tilde{s}}(z_l = z_c^\star)$, see fig. G.2. A very similar approach describes the dependence of Γ_A^{sc} on z_c.

As explained in sec. 9.1, see (9.23), it is advantageous with regard to the tuning in $N_f = 3+1$ (sec. 9.3) for $L_2\Gamma_\phi^{s\tilde{s}}$ to strongly depend on z_l, and $L_2\left(\Gamma_\phi^{sc}-\frac{1}{2}\Gamma_\phi^{s\tilde{s}}\right)$ to strongly (weakly) depend on z_c (z_l). The partial derivatives of these quantities[1] with respect to z_l and z_c at the LCP are shown in tab. G.4 for both the cases $\phi = A$ and $\phi = P$. We find that the wanted dependencies of $L_2\Gamma_\phi^{s\tilde{s}}$ on z_l as well as $L_2\left(\Gamma_\phi^{sc}-\frac{1}{2}\Gamma_\phi^{s\tilde{s}}\right)$ on z_c are stronger for $\phi = A$ than $\phi = P$. The axial current is thus better suited to determine $\Phi_2^\star$ (in $N_f = 2$) and fix Φ_2 to it (in $N_f = 3+1$). However, this comes at the cost that $L_2\left(\Gamma_\phi^{sc}-\frac{1}{2}\Gamma_\phi^{s\tilde{s}}\right)$ does unwantedly depend also on z_l for $\phi = A$, which is not the case for $\phi = P$. The subtraction of half the light EMM $\Gamma_{s\tilde{s}}$ from Γ_{sc} serves its purpose to minimize the dependence on z_l, perfectly for $\phi = P$, while the resulting quantity drops by ~ 12 % from $z_l = 0$ to $z_l = z_c^\star$ in the case of the axial current. This might be a disadvantage with regard to the effort of the tuning of Φ_3 to $\Phi_3^\star$ in $N_f = 3+1$. Nevertheless, we opt for the use of the axial current.

G.4 Γ_A with unimproved axial current correlation function in $N_f = 2$

In this section, the continuum extrapolations of the quantities $L_2\Gamma_{\text{cor}}^{s\tilde{s}}$ and $L_2(\Gamma_{\text{cor}}^{sc} - \frac{1}{2}\Gamma_{\text{cor}}^{s\tilde{s}})$, which lead to $\Phi_2^\star$ and $\Phi_3^\star$, respectively, are considered with the *unimproved* axial current correlation function f_A employed in the EMM Γ, i.e. with $c_A \to 0$. The results

---

[1] Due to the fixed volume, $\partial L_2 = 0$, the derivatives with respect to the RGI Masses M_i may be retrieved under use of $\frac{\partial(L_2\Gamma)}{\partial z_i} = \frac{\partial(L_2\Gamma)}{(\partial L_2)M_i + L_2(\partial M_i)} = \frac{1}{L_2}\frac{\partial(L_2\Gamma)}{\partial M_i}$.

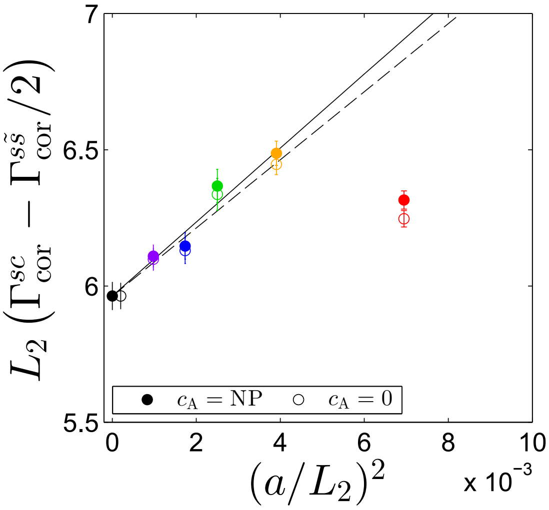
L_2/a	ens.	$L_2\Gamma^{s\tilde{s}}_{\mathrm{cor}}$	Q_{fit}	$L_2\Gamma^{s\tilde{s}}_{0,\mathrm{cor}}$	Q_{fit}
12	C_{12}	0.65(1)		0.63(1)	
16	C_{16}	0.68(1)		0.67(1)	
20	C_{20}	0.66(2)		0.65(2)	
24	C_{24}	0.60(1)		0.60(1)	
32	C_{32}	0.62(1)		0.62(1)	
∞		0.59(1)	0.01	0.59(1)	0.01

L_2/a	ens.	$L_2\left(\Gamma^{sc}_{\mathrm{cor}} - \frac{1}{2}\Gamma^{s\tilde{s}}_{\mathrm{cor}}\right)$	Q_{fit}	$L_2\left(\Gamma^{sc}_{0,\mathrm{cor}} - \frac{1}{2}\Gamma^{s\tilde{s}}_{0,\mathrm{cor}}\right)$	Q_{fit}
12	C_{12}	6.32(3)		6.25(3)	
16	C_{16}	6.49(4)		6.45(4)	
20	C_{20}	6.37(6)		6.34(6)	
24	C_{24}	6.15(5)		6.13(5)	
32	C_{32}	6.11(4)		6.10(4)	
∞		5.96(5)	0.30	5.96(5)	0.31

Table G.5: Additional results of the $N_f = 2$ runs on $L_2\Gamma^{s\tilde{s}}_{\mathrm{cor}}$ *(top)* and $L_2(L_2\Gamma^{sc}_{\mathrm{cor}} - \frac{1}{2}L_2\Gamma^{s\tilde{s}}_{\mathrm{cor}})$ *(bottom)* with the improved and unimproved axial correlation function, $f_{A,\mathrm{I}}$ and f_A, respectively. In the latter case, the EMM are labelled with a subscript 0. Compare to tab. 9.5 and tab. 9.6. The data for $L_2/a = 12$ is excluded from the continuum extrapolations.

are compared to the original ones of sec. 9.2.4 and sec. 9.2.5 in tab. G.5. The effects on the lattice results are rather small in size, and significant only for $L_2(\Gamma^{sc}_{\mathrm{cor}} - \frac{1}{2}\Gamma^{s\tilde{s}}_{\mathrm{cor}})$, illustrated in fig. G.3. The cutoff effects are still in accordance with an $O(a^2)$ behavior. In both cases, the alteration does not affect the continuum extrapolated values, as anticipated.

Figure G.3: $L_2\left(\Gamma^{sc}_{\mathrm{cor}} - \frac{1}{2}\Gamma^{s\tilde{s}}_{\mathrm{cor}}\right)$ as shown in fig. 9.5 (filled symbols), compared to the same quantity with c_A set to zero (empty symbols). The fits, linear in $(a/L_2)^2$, are illustrated by the solid and the dashed line for the non-perturbative c_A and $c_A = 0$, respectively. The continuum extrapolated value in the case of $c_A = 0$ (shown as the empty black squared symbol) is slightly shifted to the right for convenience. The data for $L_2/a = 12$ (red) is not taken into account in the fits.



H Fit methods

The present chapter briefly covers the description of data by fit functions. Although most of the presented methods constitute established textbook knowledge, we include them here for completeness, to set our notation, and to ensure transparency of our results. The presentation is mostly based on [159, 160].

We consider an N-tuple of data $(y_1, \dots y_N)$ with errors $(\sigma_1, \dots \sigma_N)$ at points $(x_1, \dots x_N)$ in D dimensions, $y_n, \sigma_n \in \mathbb{R}$ and $x_n \in \mathbb{R}^D$. These are to be described by a fit function

$$f : \mathbb{R}^D \times \mathbb{R} \to \mathbb{R}\,, \quad (x, c) \mapsto f(x, c) \tag{H.1}$$

with a number of P fit parameters $c = (c_1, \dots, c_P)$.

H.1 Optimal fit parameters by χ^2 minimization

The optimal fit function is defined by the parameters c that minimize the χ^2 function,

$$\chi^2 : \mathbb{R}^P \to \mathbb{R}\,, \quad c \mapsto \chi^2(c) \tag{H.2}$$

with

$$\chi^2(c) = \sum_{n=1}^{N} \frac{(y_n - f(x_n, c))^2}{\sigma_n^2}\,. \tag{H.3}$$

The minimum of χ^2 is given at the point where its gradient vanishes[1], which leads to P equations for the P fit parameters sought for:

$$\frac{\partial \chi^2}{\partial c_p} = -2 \sum_{n=1}^{N} \frac{(y_n - f(x_n, c))}{\sigma_n^2} \frac{\partial f(x_n, c)}{\partial c_p} = 0\,, \qquad p \in \{1, \dots, P\}\,. \tag{H.4}$$

---

[1] χ^2 possesses a global minimum as it is restricted from below. Possibly existing local minima do play a role if the fit parameters are non-linear, see below.

In the case of *linear* fit parameters, where the fit function may be written as

$$f(x,c) = \sum_{q=1}^{P} c_q \, f_q(x) \; , \tag{H.5}$$

the partial derivatives $\frac{\partial f}{\partial c_p}(x_n;c) = f_p(x_n)$ depend on x_n only, such that the equations (H.4) become

$$\sum_{q=1}^{P} \sum_{n=1}^{N} \frac{f_p(x_n)}{\sigma_n^2} \, f_q(x_n) \cdot c_q = \sum_{n=1}^{N} \frac{f_p(x_n)}{\sigma_n^2} \, y_n \; , \qquad p \in \{1, \ldots, P\} \; . \tag{H.6}$$

This linear system can be written in matrix form as

$$Ac = b \; , \tag{H.7a}$$

where

$$A_{pq} = \sum_{n=1}^{N} \frac{f_p(x_n)}{\sigma_n^2} \, f_q(x_n) \; , \quad b_p = \sum_{n=1}^{N} \frac{f_p(x_n)}{\sigma_n^2} \, y_n \; . \tag{H.7b}$$

Note that $A \equiv A(\{x_i\}, \{\sigma_i\})$ and $b \equiv b(\{x_i\}, \{y_i\}, \{\sigma_i\})$. The optimal fit parameters in the linear case (H.5) are thus simply given by $c = A^{-1}b$. In the more general case of non-linear fit parameters, the equations (H.4) are coupled, do not necessarily possess a unique solution and may not be solved analytically. Instead, one employs numerical non-linear optimization techniques (cf. sec. 9.3).

H.2 Error analysis and fit quality

With the optimal fit parameters given, their covariance matrix, which encodes the uncertainties and mutual correlations needs to be determined. Its inverse, to first order in the deviation from the optimal fit parameters, is given by the curvature[2] of χ^2, i.e. one half times the Hesse matrix:

$$\mathrm{cov}^{-1}(c_p, c_q) = \frac{1}{2} \, \frac{\partial \chi^2(c)}{\partial c_p \, \partial c_q} \; . \tag{H.8}$$

---

[2] Note, however, that in the general case of non-linear fit parameters, a P-dimensional paraboloid not necessarily has to be a good approximation of χ^2 around the minimum, in which case better approximations need to be applied to render the method more precisely (e.g. "profile likelihood method").

This corresponds to the notion that the uncertainties are smaller the larger the curvature. For non-linear fit parameters, again, the Hesse matrix can only be computed numerically. In the special case of linear fit parameters, in contrast, the covariance may analytically be obtained from the covariance matrix $\mathrm{cov}(y_n, y_m)$ of the data by linear error propagation:

$$\begin{aligned}
\mathrm{cov}(c_p, c_q) &= \mathrm{cov}((A^{-1})_{pp'}\, b_{p'}, (A^{-1})_{qq'}\, b_{q'}) \\
&= (A^{-1})_{pp'} \sum_{n,m} \frac{f_{p'}(x_n)\, f_{q'}(x_m)}{\sigma_n^2\, \sigma_m^2}\, (A^{-1})_{qq'}\, \underbrace{\mathrm{cov}(y_n, y_m)}_{\sigma_n^2 \delta_{nm}} \\
&= (A^{-1})_{pp'}\, A_{p'q'}\, (A^{-1})_{qq'} \\
&= (A^{-1})_{pq}\ . \qquad \text{(H.9)}
\end{aligned}$$

Note that we have assumed uncorrelated data $\{y_i\}$ here, in the second line. The result is in accordance with (H.8):

$$\frac{1}{2} \frac{\partial \chi^2(c)}{\partial c_p\, \partial c_q} \overset{(H.4)}{=} \frac{1}{2} \frac{\partial}{\partial c_q} \left[-2 \sum_{n=1}^{N} \frac{(y_n - f(x_n; c))}{\sigma_n^2} \frac{\partial f(x_n; c)}{\partial c_p} \right] = \sum_{n=1}^{N} \frac{f_q(x_n)}{\sigma_n^2}\, f_p(x_n) = A_{pq}\ . \qquad \text{(H.10)}$$

Moreover, (H.9) is equivalent to the generic and possibly more familiar expression

$$\mathrm{cov}(c_p, c_q) = \sum_{n,m} \frac{\partial c_p}{\partial y_n} \cdot \mathrm{cov}(y_n, y_m) \cdot \frac{\partial c_q}{\partial y_m}\ , \qquad \text{(H.11)}$$

as can readily be seen by application of (H.7) to the second line of (H.9).

The standard errors of the c_p's are easily retrieved from the diagonal elements of (H.9):

$$\sigma_p = \sqrt{\mathrm{cov}(c_p, c_p)} = \sqrt{(A^{-1})_{pp}}\ . \qquad \text{(H.12)}$$

More generally, the uncertainty $\Delta f \equiv \sigma_f$ of any function f that depends on the fit parameters c, in particular the one of the fit function itself, may be derived by linear propagation of the fit parameter errors under consideration of their correlations, similar to (H.11):

$$(\Delta f)^2\,(x, c) = \sum_{p,q} \frac{\partial f(x, c)}{\partial c_p} \cdot \mathrm{cov}(c_p, c_q) \cdot \frac{\partial f(x, c)}{\partial c_q}\ . \qquad \text{(H.13)}$$

A useful measure which disregards the absolute size of the variances is the correlation

$$\operatorname{cor}(c_p, c_q) = \frac{\operatorname{cov}(c_p, c_q)}{\sqrt{\operatorname{cov}(c_p, c_p) \cdot \operatorname{cov}(c_q, c_q)}} \,. \tag{H.14}$$

Errors $\sigma_{x,i}$ in the variables x_i, if they are independent of those of y_i (i.e. cor $\left(\sigma_{x,i}, \sigma_j\right) = 0 \ \ \forall\ i, j \in \{1, \dots, N\}$), can be projected under use of the partial derivatives of the fit function f with respect to the components of x:

$$\sigma^2_{\text{tot},i} = \sigma_i^2 + \left(f'(x_i) \cdot \sigma_{x,i}\right)^2 \ . \tag{H.15}$$

As the fit function and thus $f'(x_i)$ depends on the deployed errors σ_i, the fitting procedure should then be repeated with $\sigma_{\text{tot},i}$ and possibly iterated until $f'(x_i)$ and therefore $\sigma_{\text{tot},i}$ is stable.

Once the optimal fit parameters are determined, one should evaluate the compatibility of the data with the model chosen to describe it. To this end, one considers the expected distribution of χ^2 for $n \in \{1, \dots, N\}$ independent, random variables $\tilde{y}_n$ with individual, normal probability distributions

$$P_{f(x_n),\sigma_n}(\tilde{y}_n) = \frac{1}{\sqrt{2\pi}\sigma_n}\, e^{-\frac{(\tilde{y}_n - f(x_n))^2}{2\sigma_n^2}} \ , \tag{H.16}$$

whose expectation values and errors are given by the fit function $f(x_n)$ and the errors σ_n of y_n, respectively. The probability pdf(χ^2, N) for a certain $\chi^2 > 0$ to occur in this setting can be obtained by integration over the product of those individual distributions:

$$\operatorname{pdf}(\chi^2, N) = \int d\tilde{y}\, \left[\prod_{n=1}^{N} P_{f(x_n),\sigma_n}(\tilde{y}_n)\right] \delta(\chi^2(\tilde{y}) - \chi^2) = \frac{(\chi^2)^{N/2-1} \cdot e^{-\chi^2/2}}{2^{N/2}\,\Gamma(N/2)} \ . \tag{H.17}$$

The probability for a χ^2 smaller than a value $\chi^2_\star$ to occur is given by the cumulative probability cdf$(\chi^2_\star, N)$, while its counterpart

$$Q_{\text{fit}}(\chi^2_\star, N) = 1 - \underbrace{\int_0^{\chi^2_\star} d\chi^2 \ \operatorname{pdf}(\chi^2, N)}_{\operatorname{cdf}(\chi^2_\star, N)} = \int_{\chi^2_\star}^{\infty} d\chi^2 \ \operatorname{pdf}(\chi^2, N) \tag{H.18}$$

provides the probability for $\chi^2 > \chi^2_\star$.

As the expectation value of χ^2 is monotonically increasing with N, it is instructive

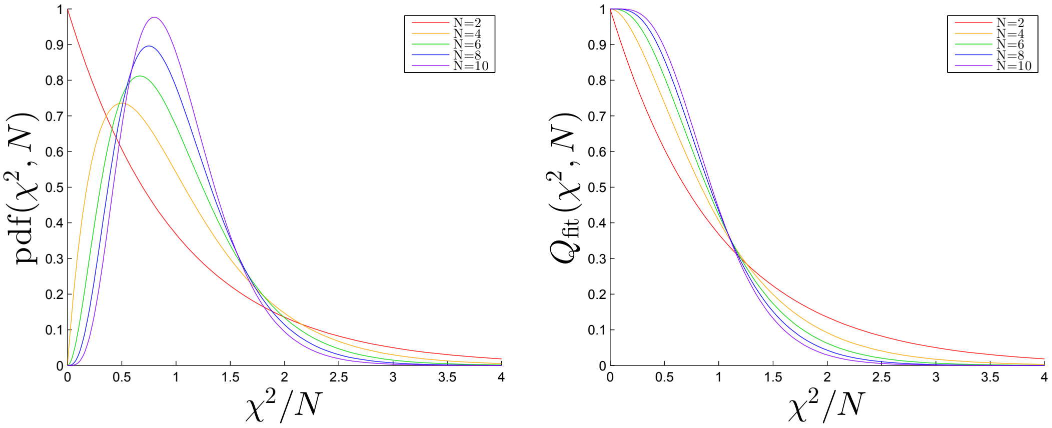


Figure H.1: *Left:* Probability distribution function pdf(χ^2, N) as defined in (H.17). *Right:* Fit quality function $Q_{\text{fit}}(\chi^2, N)$ as defined in (H.18).

to consider the quantity χ^2/N, which for large N approximates a normal distribution whose expectation value and width behave like

$$\langle \chi^2/N \rangle \approx 1 \ , \qquad \sigma_{\chi^2/N} \approx \frac{2}{\sqrt{N}} \ . \tag{H.19}$$

The distributions of χ^2/N and the associated Q_{fit} are depicted for different degrees of freedom N in fig. H.1.

In the case where the data $\{y_n, \sigma_n\}$ determines the fit function f and the associated χ^2_{fit}, the quantity $Q_{\text{fit}}(\chi^2_{\text{fit}}, \cdot)$ provides the probability of statistically distributed data with $\chi^2 > \chi^2_{\text{fit}}$ to occur under the assumption that the fit function represents the correct model. However, the number of degrees of freedom is in that case diminished by the number of fit parameters,

$$n_{\text{dof}} = N - P \ , \tag{H.20}$$

which can be understood in the way that this very number of data points is needed to fix the parameters. Hence,

$$Q_{\text{fit}}(\chi^2_{\text{fit}}, n_{\text{dof}}) = \int_{\chi^2_{\text{fit}}}^{\infty} d\chi^2 \ \text{pdf}(\chi^2, n_{\text{dof}}) \tag{H.21}$$

is used as a measure to estimate the quality of the fit, with a large value indicating a large probability for the fit function to constitute a valid model.

If the data $\{y_i\}$ is given without errors $\{\sigma_i\}$, these are first set to one, so that each data point has the same weight in the determination of the optimal fit parameters. The covariance matrix $\mathrm{cov}(c_p, c_q)$ of the fit parameters may then be estimated by scaling the errors $\{\sigma_i\}$ such that instead of the minimized $\chi^2_{\mathrm{fit}}/n_{\mathrm{dof}}$ one obtains the expectation value, i.e. $\langle \chi^2_{\mathrm{fit}}/n_{\mathrm{dof}} \rangle \overset{!}{=} 1$. However, even if constant σ_i is a valid assumption and the fit model is correct, the procedure leads to biased error estimates due to the presence of statistical fluctuations of $\chi^2_{\mathrm{fit}}/n_{\mathrm{dof}}$ around 1. As the size of these fluctuations follows (H.19), the resulting fit parameter error estimates are more reliable the larger n_{dof}, and should be considered with care especially if n_{dof} is small. In addition, the opportunity to assess the quality of the fit via Q_{fit} (H.21) is not given in the case of missing data errors $\{\sigma_i\}$.

H.3 Details on performed fits

In the present work, the optimal fit parameters and their covariance matrix in the case of all linear or linearizable fit functions were *analytically* calculated as described in sec. H.1. The non-linear Padé fit results (sec. 9.6.3) were conducted with the python package `kafe` [160], which under the surface applies the `TMinuit` class of the `ROOT` package [161]. All results were subjected to an independent cross check under use of the (non-linear) `curve_fit` routine, which is part of the `SciPy` library [162]. In tab. H.1, we list the respective covariance matrices of the fit parameters for completeness.

ref.	fit function	misc	fit parameters	covariance matrix	Q_{fit}
(9.53)	$\overline{g}^2_{\mathrm{GF}}(a^2)$		$\Phi_1^\star = 7.31$ $s_1 = 136.20$	$\begin{pmatrix} 0.0105 & -3.6863 \\ -3.6863 & 1628.2 \end{pmatrix}$	0.99
(9.56)	$L_2\Gamma^{s\bar{s}}_{\mathrm{cor}}(a^2)$		$\Phi_2^\star = 0.59$ $s_2 = 22.82$	$\begin{pmatrix} 0.0001 & -0.0513 \\ -0.0513 & 28.607 \end{pmatrix}$	0.01
(9.60)	$L_2\left(\Gamma^{sc}_{\mathrm{cor}} - \frac{1}{2}\Gamma^{s\bar{s}}_{\mathrm{cor}}\right)(a^2)$		$\Phi_3^\star = 5.96$ $s_3 = 135.67$	$\begin{pmatrix} 0.0025 & -0.8833 \\ -0.8833 & 398.22 \end{pmatrix}$	0.30
(9.124a)	$T^\star \Delta M^{ud}(c_{\mathrm{sw}})$	$T^\star/a = 8$	$c^{(8)}_{\mathrm{sw,I}} = 2.211$ $s^{(8)} = -0.141$	$\begin{pmatrix} 6.3 & 2.1 \\ 2.1 & 3.5 \end{pmatrix} \cdot 10^{-4}$	0.21
(9.124a)	$T^\star \Delta M^{ud}(c_{\mathrm{sw}})$	$T^\star/a = 12$	$c^{(12)}_{\mathrm{sw,I}} = 2.105$ $s^{(12)} = -0.105$	$\begin{pmatrix} 9.2 & 3.2 \\ 3.2 & 2.6 \end{pmatrix} \cdot 10^{-4}$	0.12
(9.124a)	$T^\star \Delta M^{ud}(c_{\mathrm{sw}})$	$T^\star/a = 16$	$c^{(16)}_{\mathrm{sw,I}} = 1.828$ $s^{(16)} = -0.083$	$\begin{pmatrix} 3.5 & 0.7 \\ 0.7 & 0.7 \end{pmatrix} \cdot 10^{-4}$	0.73
(9.124a)	$T^\star \Delta M^{ud}(c_{\mathrm{sw}})$	$T^\star/a = 20$	$c^{(20)}_{\mathrm{sw,I}} = 1.760$ $s^{(20)} = -0.059$	$\begin{pmatrix} 11.6 & 2.2 \\ 2.2 & 1.6 \end{pmatrix} \cdot 10^{-4}$	0.47
(9.124a)	$T^\star \Delta M^{ud}(c_{\mathrm{sw}})$	$T^\star/a = 24$	$c^{(24)}_{\mathrm{sw,I}} = 1.686$ $s^{(24)} = -0.063$	$\begin{pmatrix} 4.1 & 0.3 \\ 0.3 & 1.6 \end{pmatrix} \cdot 10^{-4}$	0.01
(9.124b)	$T^\star \Delta M^{ud}(c_{\mathrm{sw}})$	global	$c^{(24)}_{\mathrm{sw,I}} = 1.688$ $c^{(20)}_{\mathrm{sw,I}} = 1.755$ $c^{(16)}_{\mathrm{sw,I}} = 1.833$ $c^{(12)}_{\mathrm{sw,I}} = 2.105$ $c^{(8)}_{\mathrm{sw,I}} = 2.202$ $s = -1.257$	$\begin{pmatrix} 59.2 & 0.5 & 0.6 & 0.8 & 0.5 & 7.6 \\ 0.5 & 77.3 & 2.9 & 4.0 & 2.4 & 38.7 \\ 0.6 & 2.9 & 34.3 & 5.0 & 3.0 & 48.2 \\ 0.8 & 4.0 & 5.0 & 59.1 & 4.1 & 66.2 \\ 0.5 & 2.4 & 3.0 & 4.1 & 42.2 & 39.3 \\ 7.6 & 38.7 & 48.2 & 66.2 & 39.3 & 638.2 \end{pmatrix}$ $\cdot 10^{-5}$	0.14
(9.126)	$c_{\mathrm{sw,I}}(g^2_{0,\mathrm{I}})$	global 2nd order	$a = -0.257$ $b = -0.050$	$\begin{pmatrix} 6.2 & 3.6 \\ 3.6 & 2.1 \end{pmatrix} \cdot 10^{-5}$	0.59

Table H.1: Fit parameters and their associated covariance matrices for the most important fits employed in this work. The first column of the table refers to the respective explicit forms of the fit functions listed in the second column. The covariance matrices are ordered according to the shown order of the fit parameters, i.e. the element in the first column and row is the variance of the first fit parameter etc.

I Additional results and checks

I.1 Detailed results of the simulations (example)

In this section, we provide detailed results of the numerical simulations described in sec. 9, for the quantities listed in tab. 9.11. We use the runs at $T^\star/a = 24$ and $c_{\mathrm{sw}} = 1.8$ as an example. The results of the *tuning* simulation are listed in tab. I.1. Monte Carlo histories of selected observables (and the impact of different N_{th} for Φ_i) are shown in fig. I.1 and fig. I.2. The corresponding results of the *improvement* simulations can be found in tab. I.2, fig. I.3 and fig. I.4. Furthermore, fig. I.5 gives information on the stability of the results with regard to the thermalization (i.e. the choice of $N_{\mathrm{th}}^{(\mathrm{rep})}$) and replica, see also app. I.2.

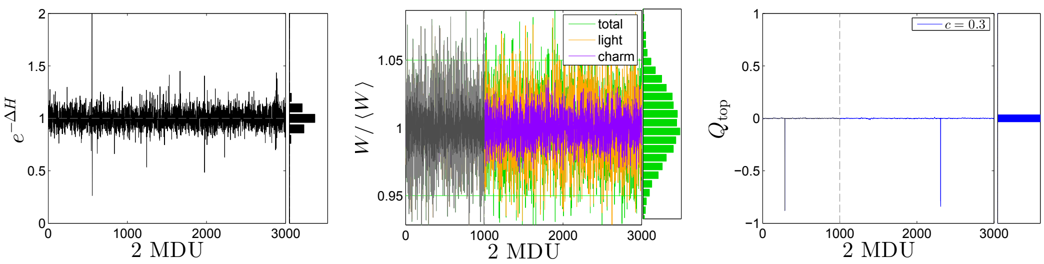


Figure I.1: MC histories and histograms of $e^{-\Delta H}$, normalized RHMC reweighting factors, and the topological charge Q_{top}.

Tuning (LCP)

- Algorithmic & Basic

N_{cf}	3000 (2 MDU)
$N_{\text{th}}^{(\text{th})}$	1000
N_1	10
S_τ	4-10

	Result	$\tau_{\text{int}}[2\text{ MDU}]$
P_{acc}	(96.54 ± 0.12) %	0.53 ± 0.04
E_P	1.79778 ± 0.00003	1.78 ± 0.25
$\exp(-\Delta H)$	0.9994 ± 0.0020	0.49 ± 0.02

- Rational approximation

	n	$r_a - r_b$	spectrum `[ms2]`	$\max(W)$ `[ms1]`
D_s	8	2.60e-02 – 7.00e+00	3.81e-02 – 6.31e+00	8.66e-02
D_c	6	1.90e-01 – 7.10e+00	2.07e-01 – 6.37e+00	3.93e-02

- Topological charge & GF coupling

	$c = 0.3$	
	Result	$\tau_{\text{int}}[2\text{ MDU}]$
Q_{top}	-0.0014 ± 0.0012	1.54 ± 0.31
Q_{top}^2	0.0010 ± 0.0010	1.50 ± 0.30
$\delta^{\epsilon}_{Q_{\text{top}},0}$	(99.8 ± 0.2) %	1.50 ± 0.30
$\bar{g}^2_{\text{GF}}$	7.11 ± 0.07	12.79 ± 6.07
$\bar{g}^2_{\text{GF},0}$	7.11 ± 0.07	12.85 ± 6.10

- Main results

	Result	$D_i[\%]$	$\tau_{\text{int}}[2\text{ MDU}]$
Φ_1	7.11 ± 0.07	-2.8 ± 1.0	12.79 ± 6.07
Φ_2	0.56 ± 0.01	-4.9 ± 2.3	1.77 ± 0.41
Φ_3	5.87 ± 0.06	-1.5 ± 1.0	33.52 ± 20.44

Table I.1: Detailed results for the tuning run at $T^\star/a = 24$, $L/a = 24$, $c_{\text{sw}} = 1.8$ and $g_0^2 = 1.6344$, $\kappa_l = 0.13515$, $\kappa_c = 0.13132$ (cf. tab. 9.14).

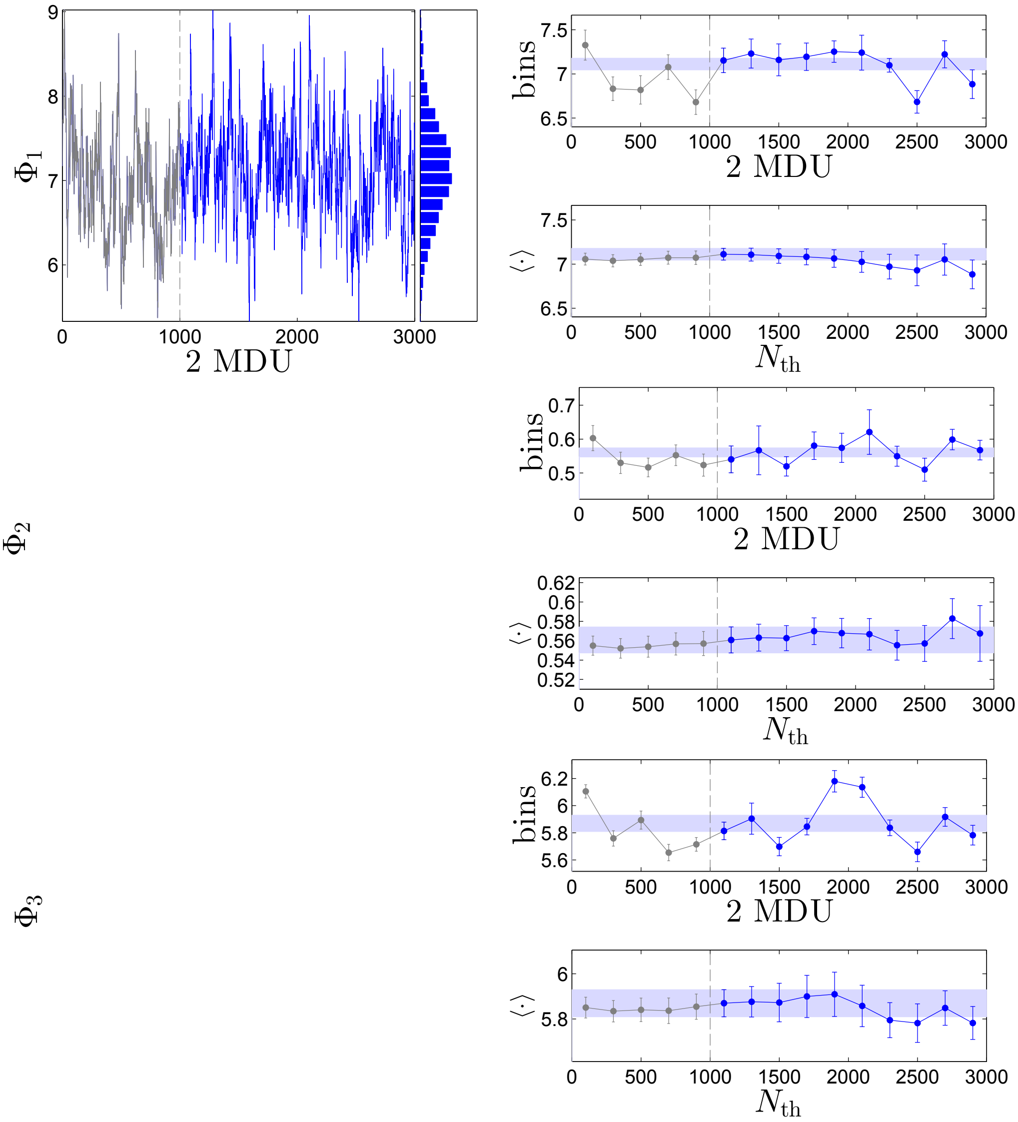


Figure I.2: MC histories and histograms of the primary observable Φ_1 *(left panel)*. The figures in the right panel show the expectation values of bins of 200 configurations *(top)* and the overall expectation value in dependence of N_{th} *(bottom)* for Φ_1, Φ_2 and Φ_3. The configurations which were in the end discarded due to the choice of N_{th} are shown in gray.

Improvement (IMPR)

- Algorithmic & Basic

N_{cf}	8x4000 (2 MDU)
$N_{\mathrm{th}}^{(\mathrm{rep})}$	0
N_1	9
S_τ	4-10

	Result	τ_{int}[2 MDU]	Q_{rep}
P_{acc}	(97.40 ± 0.02) %	0.50 ± 0.01	0.56
E_P	1.79768 ± 0.00002	1.92 ± 0.08	0.70
$\exp(-\Delta H)$	1.0003 ± 0.0004	0.49 ± 0.01	0.82

- Rational approximation

	n	$r_a - r_b$	spectrum `[ms2]`	$\max(W)$ `[ms1]`
D_s	8	5.00e-02 – 7.50e+00	7.47e-02 – 6.11e+00	1.74e-02
D_c	6	2.00e-01 – 7.60e+00	2.30e-01 – 6.20e+00	1.55e-02

- Topological charge & GF coupling

	$c = 0.3$		
	Result	τ_{int}[2 MDU]	Q_{rep}
Q_{top}	0.0001 ± 0.0001	1.08 ± 0.06	0.14
Q_{top}^2	0.0002 ± 0.0001	0.80 ± 0.04	0.20
$\delta^{\epsilon}_{Q_{\mathrm{top}},0}$	(100.0 ± 0.0) %	0.79 ± 0.03	0.14
$\langle t^2 E(t) \rangle$	$(8.52 \pm 0.01) \cdot 10^{-2}$	5.19 ± 0.60	0.31
$\langle t^2 E(t) \rangle \|_{Q_{\mathrm{top}}=0}$	$(8.52 \pm 0.01) \cdot 10^{-2}$	5.19 ± 0.60	0.31

- Main results

	Result	$T^\star/a \cdot$ Result	τ_{int}[2 MDU]	Q_{rep}
$aM^{ud}(T/2, T/4)$	0.0031 ± 0.0002	0.076 ± 0.004	1.11 ± 0.09	0.15
$a\Delta M^{ud}(3T/4, T/4)$	-0.0003 ± 0.0001	-0.007 ± 0.002	0.85 ± 0.06	0.80

Table I.2: Detailed results for the improvement run at $T^\star/a = 24$, $L/a = 12$, $c_{\mathrm{sw}} = 1.8$ and $g_0^2 = 1.6344$, $\kappa_l = 0.13515$, $\kappa_c = 0.13132$ (cf. tab. 9.15).

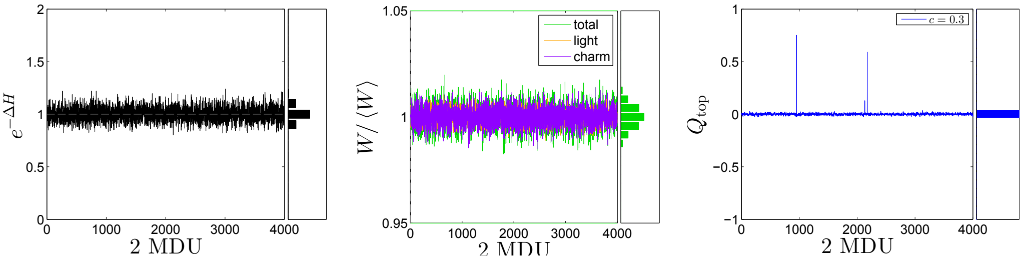


Figure I.3: MC histories and histograms of $e^{-\Delta H}$, normalized RHMC reweighting factors, and the topological charge Q_{top}.

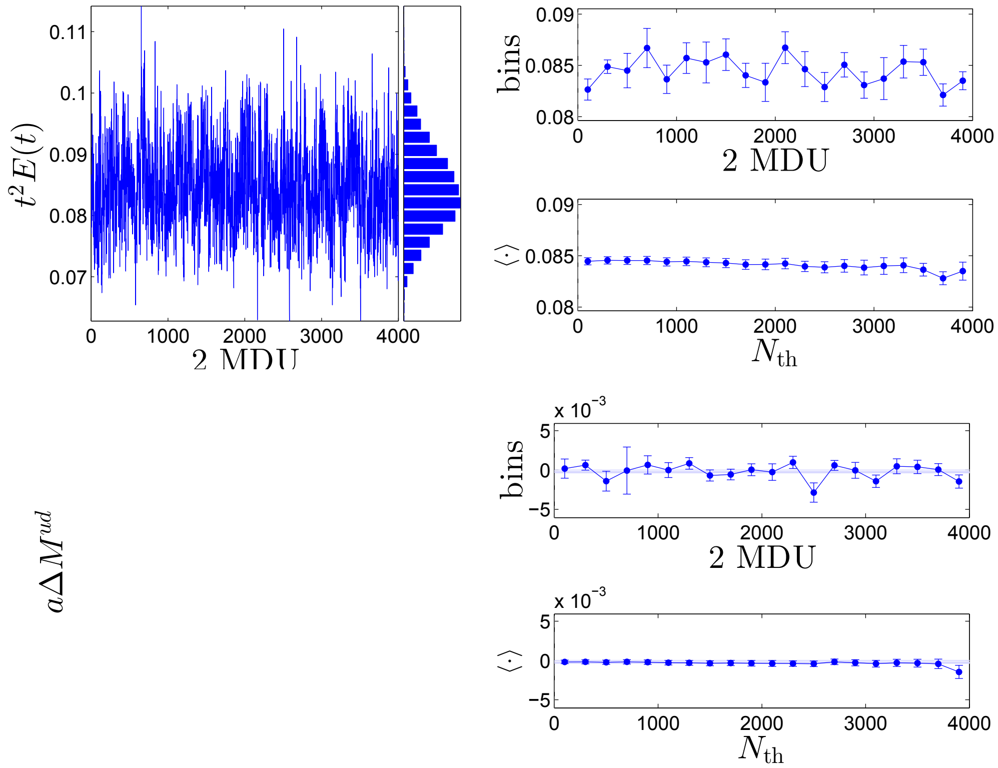


Figure I.4: MC history and histogram of $t^2E(t)$ for one replicum *(left panel)*. The figures in the right panel show the expectation values of bins of 200 configurations *(top)* and the overall expectation value in dependence of N_{th} *(bottom)* for $t^2E(t)$ and $a\Delta M^{ud}$ for one replicum.

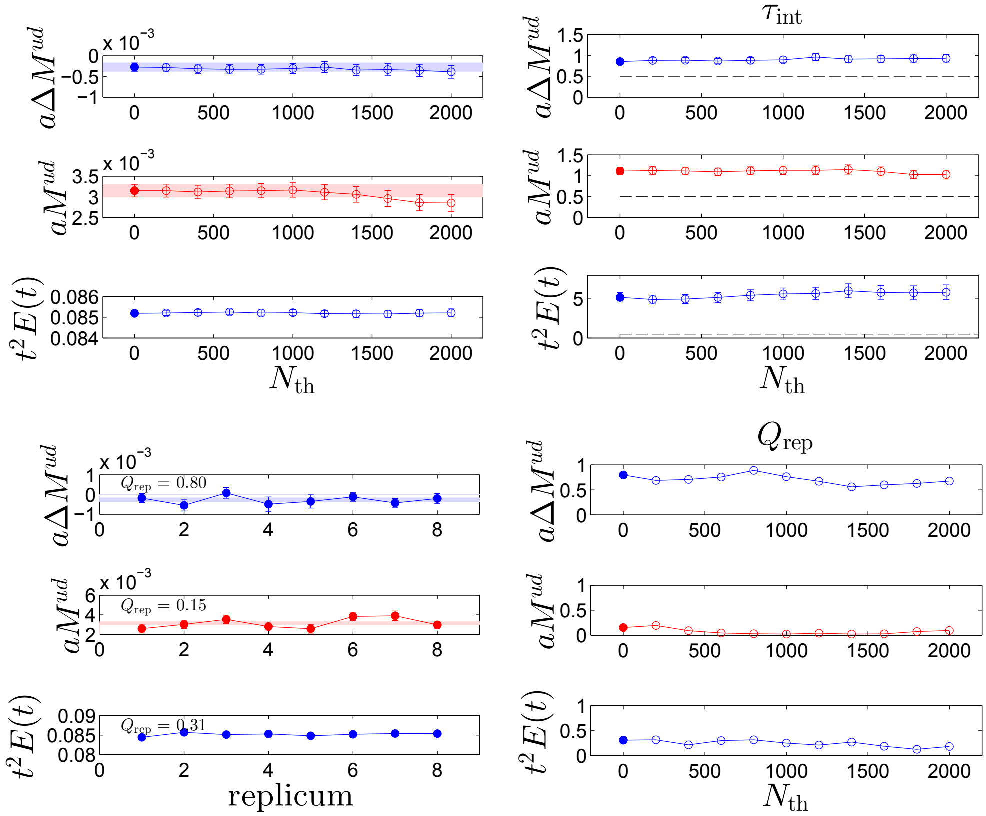


Figure I.5: The quantities $a\Delta M^{ud}$, aM^{ud} and $t^2E(t)$ *(top left)*, their autocorrelations $\tau_{\rm int}$ *(top right)*, and their replica qualities $Q_{\rm rep}$ *(bottom right)* in dependence of the thermalization $N_{\rm th}^{\rm (rep)}$. For the eventual choice of $N_{\rm th}^{\rm (rep)}$, filled circles are used, and the results of the single replica are compared *(bottom left)*.

I.2 Stability of results for $T^{\star}\Delta M^{ud}$ and $c_{\rm sw,I}$

Thermalization stability

First, we check the thermalization for the quantity $T^{\star}\Delta M^{ud}$. To this end, the results of the thermalization run ($n_{\rm rep} = 1$, $N_{\rm th}^{\rm (th)} \neq 0$) are compared with those of the replica runs without additional thermalization ($n_{\rm rep} = 8$, $N_{\rm th}^{\rm (rep)} = 0$), cf. the discussion in sec. 9.5.2. If both ensembles are fully thermalized, their results should be in agreement up to statistical deviations. In the left panel of fig. I.6, the difference

$$\delta^{\rm th,rep} = \frac{T^{\star}\Delta M^{ud,th} - T^{\star}\Delta M^{ud,rep}}{\sqrt{\left(\delta\left(T^{\star}\Delta M^{ud,th}\right)\right)^2 + \left(\delta\left(T^{\star}\Delta M^{ud,rep}\right)\right)^2}} \tag{I.1}$$

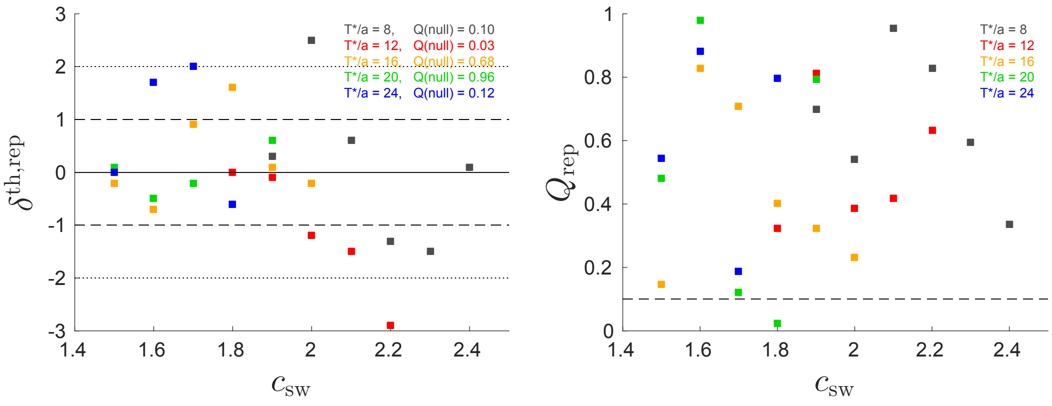


Figure I.6: Stability of the results for $T^\star \Delta M^{ud}$ without additional thermalization of the replica, $N_{\mathrm{th}}^{(\mathrm{rep})} = 0$. The *left panel* shows the deviation $\delta^{\mathrm{th,rep}}$ of the results from the thermalization and replica runs, (I.1), whereas the quality Q_{rep} of the replica distributions is displayed in the *right panel*.

is shown for all runs. The probability Q(null-hypothesis) for statistical deviations to be larger or equal to the ones observed are also displayed for each $T^\star/a$ in the figure. In the right panel, the quality of the replica distribution for $T^\star \Delta M^{ud}$ is shown. We emphasize again that the displayed results are all obtained *without* an additional thermalization of the replica, i.e. $N_{\mathrm{th}}^{(\mathrm{rep})} = 0$. In the cases[1] of $|\delta^{\mathrm{th,rep}}| > 2$ or $Q_{\mathrm{rep}} < 0.10$, this number was increased to exclude biased results due to incomplete thermalization.

Fit stability

We investigate the stability of the *individual* fits of $T^\star \Delta M^{ud}$ as a function of c_{sw}, shown in the left panel of fig. 9.13. To this end, we examine χ^2/n_{dof}, the fit quality Q_{fit} and the improvement point $c_{\mathrm{sw,I}}$ in the case where 1 or 2 data points are excluded from the fit. The results can be found in fig. I.7. The aforementioned quantities under consideration of all points in the fit are shown as squared dots (at $x = 3.5$). The circled dots, in contrast, correspond to the case of data points taken away from the margins, according to the table in fig. I.7. Hence, the further a circled dot is plotted on the left hand side in fig. I.7, i.e. the smaller the x, the more "left-leaning" is the fit, and vice versa. For the judgment of the data, one should keep in mind that, since the different fits ($\equiv$ different x) take into account a different number of data points, the meaning of χ^2/n_{dof} varies slightly, in the sense that the more data points are involved, the more likely it should be to get $\chi^2/n_{\mathrm{dof}} \approx 1$ (cf. (H.19)). This effect is taken into account

---

[1] Concretely, this concerns $(T^\star/a, c_{\mathrm{sw}}) = (8, 2.0), (12, 2.2), (20, 1.8)$, cf. fig. I.6.

x	left	right
1	0	2
2	0	1
3.5	1	1
5	1	0
6	2	0

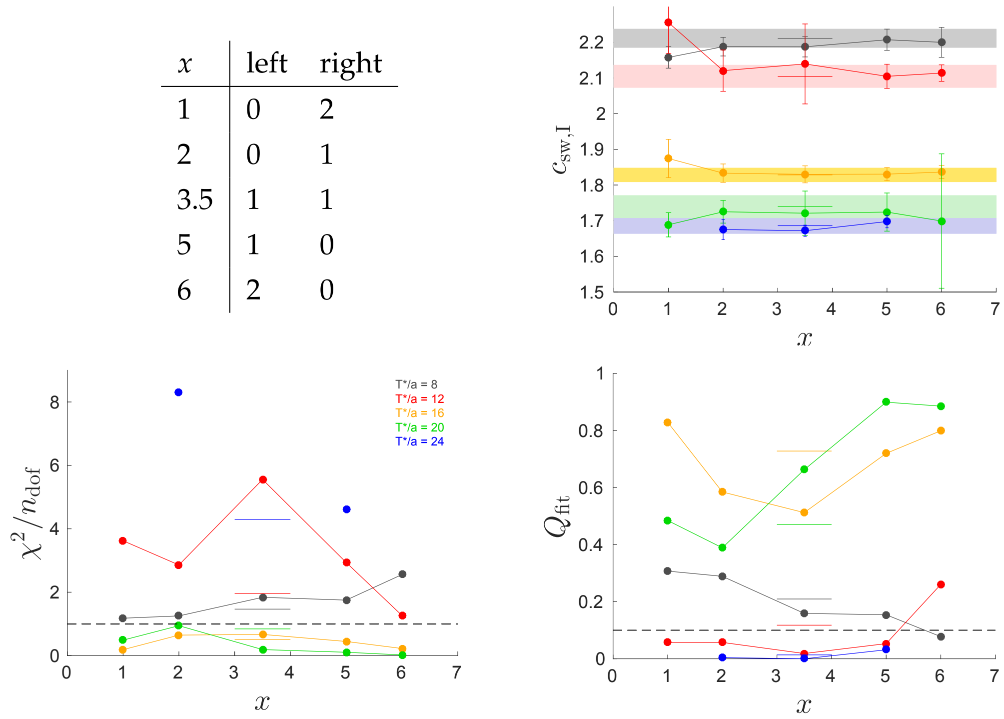


Figure I.7: The table shows the number of points taken away from the left (small c_{sw}) and right (large c_{sw}) side in the fits of $T^{\star}\Delta M^{ud}$, and the corresponding value of x at which the results can be found as circled dots in the neighboring figures. These illustrate the stability of the fits, with respect to $c_{\mathrm{sw,I}}$ *(top right)*, χ^2/n_{dof} *(bottom left)*, Q_{fit} *(bottom right)*. The original results with the respective full data sets are shown as solid lines at $x = 3.5$, and error bands in the case of $c_{\mathrm{sw,I}}$.

by Q_{fit}. We find that both the fit quality and the stability are mostly satisfying. The exception at $T^{\star}/a = 24$ is partly due to fact that only four data points enter the fit, and is a reason why the *global* fit (9.124b) is used in the end, cf. sec. 9.6.2.

I.3 Time dependence of Schrödinger functional correlation functions

In fig. I.8 and fig. I.9, we compare the time dependences of the SFCFs f_A and f_P and quantities derived thereof, for the cases of

$$T/a = 8, \quad c_{\mathrm{sw}} = 2.4, \, g_0^2 = 1.9069 \tag{I.2a}$$

$$\text{and} \qquad T/a = 16, \, c_{\mathrm{sw}} = 1.7, \, g_0^2 = 1.6741 \, . \tag{I.2b}$$

For the sake of brevity, we constrain ourselves to the *improvement* simulations. The first and second derivative of f_A and f_P are shown since they enter r^{ud}, s^{ud} and thus $m^{ud}_{\mathrm{PCAC,I}}$ (9.111) with perturbative c_A (3.115), as well as c_A (9.112) and M^{ud} (9.113). The respective primed counterparts are illustrated as well. One observes that the SFCFs and their derivatives are affected by boundary effects. At $y_0 = T/4$, which is relevant for the imposition of the improvement condition (9.114), this concerns the coarsest lattice $T/a = 8$ the most, and is responsible for the weak occurrence of a plateau in the plots of $m^{ud}_{\mathrm{PCAC,I}}$ and M^{ud}. However, we generally find

$$r^{ud}(x_0) \approx r'^{ud}(x_0) \qquad \text{and} \qquad m^{ud}_{\mathrm{PCAC,I}}(x_0) \approx m'^{ud}_{\mathrm{PCAC,I}}(x_0) \tag{I.3}$$

for both $x_0 = T/4$ and $x_0 = 3T/4$. The imposition of the improvement conditions on c_A (9.109) and c_{sw} (9.114) lead to

$$M^{ud}(x_0) = M'^{ud}(x_0) \; . \tag{I.4}$$

at $x_0 = T/4$ and $x_0 = 3T/4$, respectively. This can be seen in the case of the $T/a = 8$ simulation, where $c_{\mathrm{sw}} \approx c_{\mathrm{sw,I}}$, along with the approximate fulfillment of (9.116a). Consequently, the transition from $c_A = 0$ (in r^{ud}) or the perturbative use of c_A (in $m^{ud}_{\mathrm{PCAC,I}}$) to the non-perturbatively improved $c_{A,\mathrm{I}}$ (in M^{ud}) is a rather small step, ensuring that the order of magnitude of $c_{A,\mathrm{I}}$ is roughly similar to the perturbative estimate. Hence, the $O(a)$ effects that accompany our choice of the non-perturbative improvement condition are reasonably small. The same holds for $c_{\mathrm{sw,I}}$.

In contrast, the values of r^{ud} and $m^{ud}_{\mathrm{PCAC,I}}$ differ significantly between $x_0 = T/4$ and $x_0 = 3T/4$, especially for large lattice spacings ($T/a = 8$):

$$r^{ud}(x_0) \neq r^{ud}(y_0) \qquad \text{and} \qquad m^{ud}_{\mathrm{PCAC,I}}(x_0) \neq m^{ud}_{\mathrm{PCAC,I}}(y_0) \; , \tag{I.5}$$

and similarly for the primed counterparts. Hence, the above argumentation is unlikely to be true for an alternative improvement condition that enforces M^{ud} to be equal at different time slices, like e.g.

$$M^{ud}(x_0) = M^{ud}(y_0) \; . \tag{I.6}$$

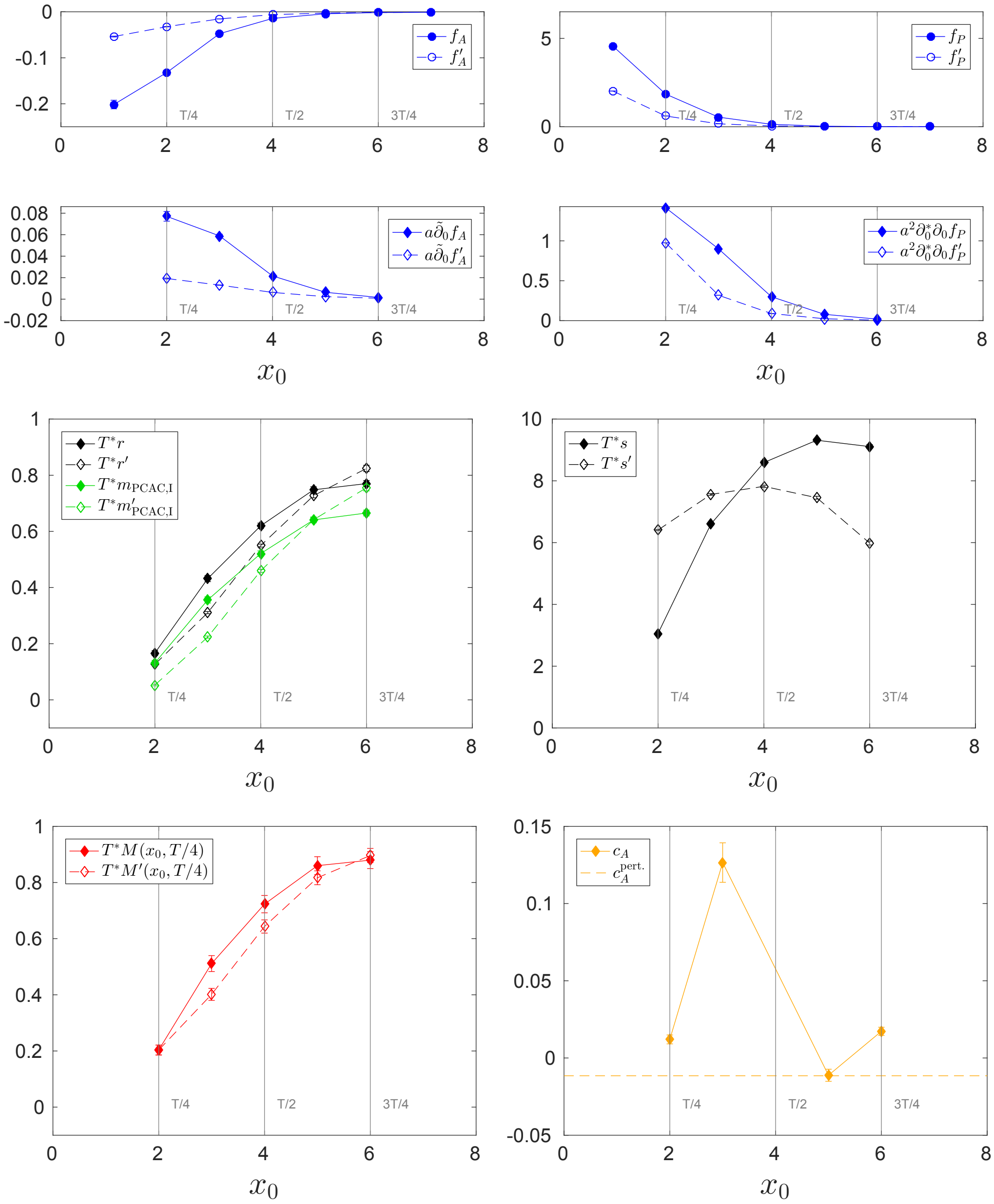


Figure I.8: x_0-dependence of Schrödinger functional correlation functions and derived observables that are relevant for the imposition of the improvement condition on c_{sw}, for the case of $T^\star/a = 8$ and $c_{\mathrm{sw}} = 2.4$. Note that all quantities employ the flavor combination *ud*.

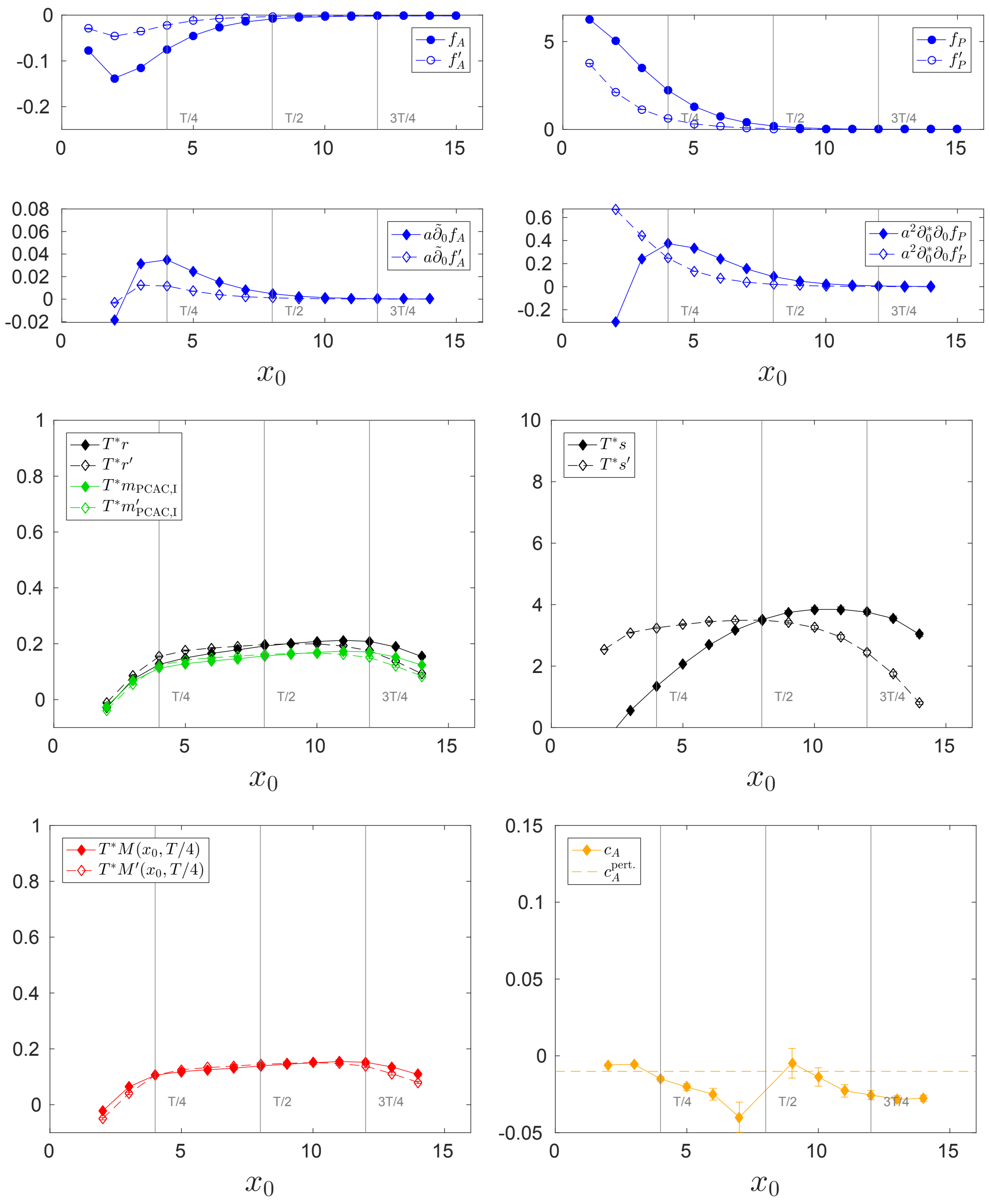


Figure I.9: x_0-dependence of Schrödinger functional correlation functions and derived observables that are relevant for the imposition of the improvement condition on c_{sw}, for the case of $T^\star/a = 16$ and $c_{\mathrm{sw}} = 1.7$. Note that all quantities employ the flavor combination *ud*.

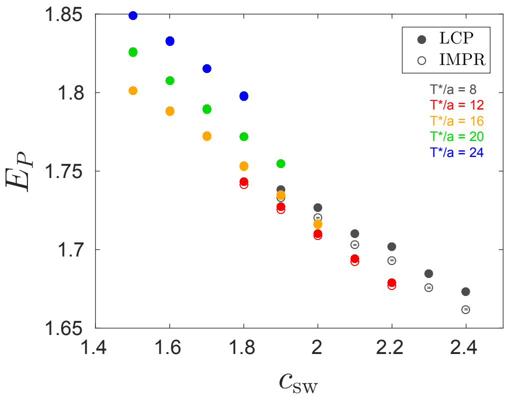


Figure I.10: Results for the average plaquette E_P (I.7) in the case of the tuning (LCP, filled circles) and improvement (IMPR, empty circles) runs.

I.4 Additional quantities (E_P, Q_{top}, c_A)

Average plaquette

The average plaquette[2]

$$E_P = \langle S_P \rangle \, / \, (6|X|) \tag{I.7}$$

is defined as the expectation value of the sum of plaquette variables (cf. (3.30))

$$S_P = \sum_x \sum_{\mu<\nu} \mathrm{Re}\,\mathrm{tr}[U_{\mu\nu}(x)] \;, \tag{I.8}$$

as it also appears in the Wilson gauge action (3.36), divided by the number $6|X|$ of terms summed over (cf. (3.14)). The results of our simulations are compiled in fig. I.10.

Topological charge

The results for the expectation values of the topological charge Q_{top} (6.31), its second moment Q^2_{top} (cf. (6.29)), and the fraction $\delta^{\epsilon}_{Q_{\mathrm{top}},0}$ (7.5) of configurations that are attributed to the trivial topological sector, are shown for $c = 0.3$ in fig. I.11, while the corresponding autocorrelations can be found in fig. I.12.

We find that the expectation value of Q_{top} is compatible with zero in all cases (except at $T/a = 12, L/a = 12, c_{\mathrm{sw}} = 2.2$). Apart from the tuning simulations at $T/a = 8, 12$, virtually all configurations of the respective ensembles belong to the trivial sector.

---

[2] Note that while the normalization of E_P is a matter of convention, we follow the definition employed in the `openQCD` package.

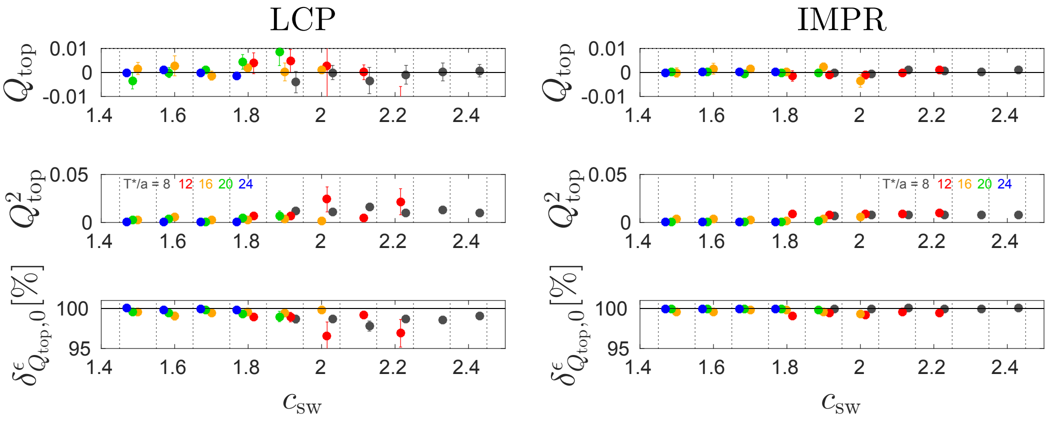


Figure I.11: Q_{top}, Q^2_{top} and $\delta^{\epsilon}_{Q_{\text{top}},0}$ with $\epsilon = 0.5$ for the tuning *(left panel)* and improvement *(right panel)* simulations and $c = 0.3$. The horizontal black line in the plots at the bottom corresponds to 100 %.

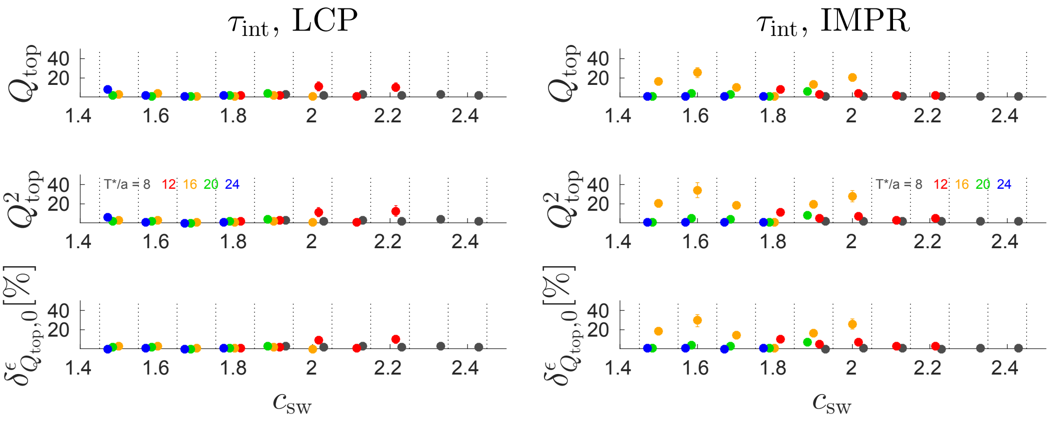


Figure I.12: Autocorrelations of the quantities shown in fig. I.11.

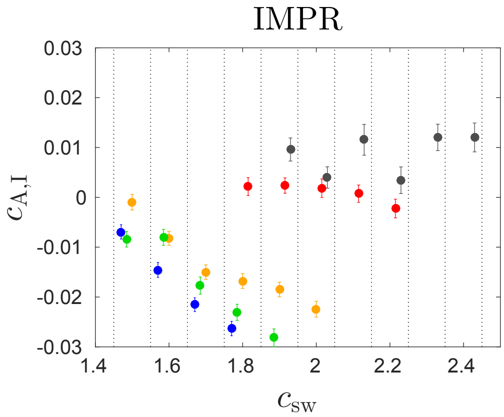


Figure I.13: Results from the improvement simulations for $c_{A,\mathrm{I}}$ as given by (9.112). The points are shifted horizontally, cf. fig. 9.9.

Improvement coefficient $c_{A,\mathrm{I}}$

The improvement coefficient $c_{A,\mathrm{I}}$, defined by (9.112), is shown in fig. I.13. Note that strictly speaking, the displayed $c_{A,\mathrm{I}}$ were evaluated at values of c_{sw} that do not correspond to $c_{\mathrm{sw,I}}$. Hence, the unique real improvement coefficient $c_{A,\mathrm{I}} = c_{A,\mathrm{I}}|_{c_{\mathrm{sw,I}}}$, that together with $c_{\mathrm{sw,I}}$ fulfills both the equations (9.112) and (9.115) is obtained by interpolation to $c_{\mathrm{sw,I}}$. Without going into further details here, one may anticipate that those $c_{A,\mathrm{I}}|_{c_{\mathrm{sw,I}}}$ lie in a reasonable range below the perturbative estimates for $T \geq 16$, while they are rather far off for $T < 16$. These are supposedly cutoff effects that are rooted in the very definition of $c_{A,\mathrm{I}}$ used for the determination of $c_{\mathrm{sw,I}}$. For a proper determination of the improvement coefficient c_A, it is more useful to adapt a strategy similar to the one described in app. F.2.

Bibliography

[1] **Wikipedia**
Standard Model — Wikipedia, The Free Encyclopedia (2016). [Online; accessed 11-May-2016], `https://en.wikipedia.org/w/index.php?title=Standard_Model&oldid=707260009`

[2] **G. Aad *et al.*** (ATLAS)**.** Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B716, 1–29 (2012).* `arXiv:1207.7214`.

[3] **S. Chatrchyan *et al.*** (CMS)**.** Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B716, 30–61 (2012).* `arXiv:1207.7235`.

[4] **F. Englert and R. Brout.** Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett. 13, 321–323 (1964).*

[5] **P. W. Higgs.** Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett. 13, 508–509 (1964).*

[6] **R. Sommer and U. Wolff.** Non-perturbative computation of the strong coupling constant on the lattice. *Nucl. Part. Phys. Proc. 261-262, 155–184 (2015).* `arXiv:1501.01861`.

[7] **K. Symanzik.** Some topics in quantum field theory. *Mathematical problems in theoretical physics, eds. R. Schrader et al., Lecture Notes in Physics, Springer. New York 153 (1980).*

[8] **K. Symanzik.** Continuum Limit and Improved Action in Lattice Theories. 1. Principles and ϕ^4 Theory. *Nucl.Phys. B226, 187 (1983).*

[9] **K. Symanzik.** Continuum Limit and Improved Action in Lattice Theories. 2. O(N) Nonlinear Sigma Model in Perturbation Theory. *Nucl.Phys. B226, 205 (1983).*

[10] **R. G. Edwards, U. M. Heller and T. R. Klassen.** The Effectiveness of nonperturbative O(a) improvement in lattice QCD. *Phys. Rev. Lett. 80, 3448–3451 (1998).* `arXiv:hep-lat/9711052`.

[11] **J. Garden, J. Heitger, R. Sommer and H. Wittig** (ALPHA, UKQCD)**.** Precision computation of the strange quark's mass in quenched QCD. *Nucl. Phys. B571, 237–256 (2000).* `arXiv:hep-lat/9906013`.

[12] **R. G. Edwards, U. M. Heller and T. R. Klassen.** Accurate scale determinations for the Wilson gauge action. *Nucl. Phys. B517, 377–392 (1998).* `arXiv:hep-lat/9711003`.

[13] **R. Sommer.** A new way to set the energy scale in lattice gauge theories and its applications to the static force and alpha-s in SU(2) Yang-Mills theory. *Nucl. Phys. B411, 839–854 (1994).* `arXiv:hep-lat/9310022`.

[14] **S. Aoki *et al.*** (CP-PACS)**.** Light hadron spectrum and quark masses from quenched lattice QCD. *Phys. Rev. D67, 034503 (2003).* `arXiv:hep-lat/0206009`.

[15] **B. Sheikholeslami and R. Wohlert.** Improved Continuum Limit Lattice Action for QCD with Wilson Fermions. *Nucl.Phys. B259, 572 (1985).*

[16] **M. Lüscher, S. Sint, R. Sommer, P. Weisz and U. Wolff.** Nonperturbative O(a) improvement of lattice QCD. *Nucl. Phys. B491, 323–343 (1997).* `arXiv:hep-lat/9609035`.

[17] **K. Jansen and R. Sommer** (ALPHA)**.** O(a) improvement of lattice QCD with two flavors of Wilson quarks. *Nucl.Phys. B530, 185–203 (1998).* `arXiv:hep-lat/9803017`.

[18] **N. Yamada *et al.*** (JLQCD, CP-PACS)**.** Non-perturbative O(a)-improvement of Wilson quark action in three-flavor QCD with plaquette gauge action. *Phys.Rev. D71, 054505 (2005).* `arXiv:hep-lat/0406028`.

[19] **S. Aoki *et al.*** (CP-PACS, JLQCD)**.** Nonperturbative O(a) improvement of the Wilson quark action with the RG-improved gauge action using the Schrödinger functional method. *Phys. Rev. D73, 034501 (2006).* `arXiv:hep-lat/0508031`.

[20] **N. Cundy, M. Gockeler, R. Horsley, T. Kaltenbrunner, A. Kennedy *et al.*.** Non-perturbative improvement of stout-smeared three flavour clover fermions. *Phys. Rev. D79, 094507 (2009).* `arXiv:0901.3302`.

[21] **F. Tekin, R. Sommer and U. Wolff** (ALPHA)**.** Symanzik improvement of lattice QCD with four flavors of Wilson quarks. *Phys.Lett. B683, 75–79 (2010).* `arXiv:0911.4043`.

[22] **J. Bulava and S. Schaefer.** Improvement of $N_f = 3$ lattice QCD with Wilson fermions and tree-level improved gauge action. *Nucl. Phys. B874, 188–197 (2013).* `arXiv:1304.7093`.

[23] **S. L. Adler.** Axial vector vertex in spinor electrodynamics. *Phys. Rev. 177, 2426–2438 (1969).*

[24] **J. Bell and R. Jackiw.** A PCAC puzzle: $\pi^0 \to \gamma\gamma$ in the sigma model. *Nuovo Cim. A60, 47–61 (1969).*

[25] **D.J.Gross and R.Jackiw.** Effect of Anomalies on Quasi-Renormalizable Theories. *Phys. Rev. D6, 477 (1972).*

[26] **K. A. Olive *et al.*** (Particle Data Group)**.** Review of Particle Physics. *Chin. Phys. C38, 090001 (2014).*

[27] **T. Ohl.** Drawing Feynman diagrams with Latex and Metafont. *Comput. Phys. Commun. 90, 340–354 (1995).* `arXiv:hep-ph/9505351`.

[28] **I. J. R. Aitchison and A. J. G. Hey.** Gauge theories in particle physics: A practical introduction. Vol. 2: Non-Abelian gauge theories: QCD and the electroweak theory. *Bristol, UK: IOP (2004).*

[29] **D. J. Gross and F. Wilczek.** Ultraviolet Behavior of Nonabelian Gauge Theories. *Phys.Rev.Lett. 30, 1343–1346 (1973).*

[30] **H. D. Politzer.** Reliable Perturbative Results for Strong Interactions? *Phys.Rev.Lett. 30, 1346–1349 (1973).*

[31] **C. Gattringer and C. Lang.** Quantum Chromodynamics on the Lattice: An Introductory Presentation. Lecture Notes in Physics. *Springer (2010).*

[32] **S. Necco and R. Sommer.** The $N_f = 0$ heavy quark potential from short to intermediate distances. *Nucl. Phys. B622, 328–346 (2002).* `arXiv:hep-lat/0108008`.

[33] **M. Lüscher.** Lattice QCD: From quark confinement to asymptotic freedom. *Annales Henri Poincare 4, S197–S210 (2003).* `arXiv:hep-ph/0211220`.

[34] **K. G. Wilson.** Confinement of Quarks. *Phys.Rev. D10, 2445–2459 (1974).*

[35] **H. B. Nielsen and M. Ninomiya.** No Go Theorem for Regularizing Chiral Fermions. *Phys. Lett. B105, 219 (1981).*

[36] **H. B. Nielsen and M. Ninomiya.** Absence of Neutrinos on a Lattice. 1. Proof by Homotopy Theory. *Nucl. Phys. B185, 20 (1981).* [Erratum: Nucl. Phys.B195,541(1982)].

[37] **H. B. Nielsen and M. Ninomiya.** Absence of Neutrinos on a Lattice. 2. Intuitive Topological Proof. *Nucl. Phys. B193, 173 (1981).*

[38] **M. Bruno *et al.*.** Simulation of QCD with $N_f = 2 + 1$ flavors of non-perturbatively improved Wilson fermions. *JHEP 02, 043 (2015).* `arXiv:1411.3982`.

[39] **R. Sommer.** Scale setting in lattice QCD. *PoS LATTICE2013, 015 (2014).* `arXiv:1401.3270`.

[40] **M. Lüscher, R. Sommer, P. Weisz and U. Wolff.** A Precise determination of the running coupling in the SU(3) Yang-Mills theory. *Nucl. Phys. B413, 481–502 (1994).* `arXiv:hep-lat/9309005`.

[41] **A. Bode *et al.*** (ALPHA)**.** First results on the running coupling in QCD with two massless flavors. *Phys. Lett. B515, 49–56 (2001).* `arXiv:hep-lat/0105003`.

[42] **M. Della Morte *et al.*** (ALPHA)**.** Computation of the strong coupling in QCD with two dynamical flavors. *Nucl.Phys. B713, 378–406 (2005).* `arXiv:hep-lat/0411025`.

[43] **M. Dalla Brida, P. Fritzsch, T. Korzec, A. Ramos, S. Sint *et al.*.** Towards a new determination of the QCD Lambda parameter from running couplings in the three-flavour theory. *PoS LATTICE2014, 291 (2014).* `arXiv:1411.7648`.

[44] **M. Dalla Brida, P. Fritzsch, T. Korzec, A. Ramos, S. Sint and R. Sommer** (ALPHA). A status update on the determination of $\Lambda_{\overline{\mathrm{MS}}}^{N_f=3}$ by the ALPHA collaboration. *PoS LATTICE2015, 248 (2016).* `arXiv:1511.05831`.

[45] **M. Della Morte, R. Hoffmann, F. Knechtli, J. Rolf, R. Sommer, I. Wetzorke and U. Wolff** (ALPHA). Non-perturbative quark mass renormalization in two-flavor QCD. *Nucl. Phys. B729, 117–134 (2005).* `arXiv:hep-lat/0507035`.

[46] **M. Lüscher, S. Sint, R. Sommer and P. Weisz.** Chiral symmetry and O(a) improvement in lattice QCD. *Nucl.Phys. B478, 365–400 (1996).* `arXiv:hep-lat/9605038`.

[47] **M. Lüscher and P. Weisz.** On-Shell Improved Lattice Gauge Theories. *Commun. Math. Phys. 97, 59 (1985).*

[48] **M. Lüscher and P. Weisz.** Computation of the Action for On-Shell Improved Lattice Gauge Theories at Weak Coupling. *Phys. Lett. B158, 250 (1985).*

[49] **R. Wohlert.** Improved continuum limit lattice action for quarks. *DESY preprint 87-069, unpublished (1987).*

[50] **M. Lüscher and P. Weisz.** O(a) improvement of the axial current in lattice QCD to one loop order of perturbation theory. *Nucl. Phys. B479, 429–458 (1996).* `arXiv:hep-lat/9606016`.

[51] **S. Aoki, R. Frezzotti and P. Weisz.** Computation of the improvement coefficient c_{sw} to one loop with improved gluon actions. *Nucl. Phys. B540, 501–519 (1999).* `arXiv:hep-lat/9808007`.

[52] **G. Heatlie, G. Martinelli, C. Pittori, G. C. Rossi and C. T. Sachrajda.** The improvement of hadronic matrix elements in lattice QCD. *Nucl. Phys. B352, 266–288 (1991).*

[53] **S. Sint and P. Weisz.** Further results on O(a) improved lattice QCD to one loop order of perturbation theory. *Nucl. Phys. B502, 251–268 (1997).* `arXiv:hep-lat/9704001`.

[54] **M. Della Morte, R. Hoffmann and R. Sommer.** Non-perturbative improvement of the axial current for dynamical Wilson fermions. *JHEP 0503, 029 (2005).* `arXiv:hep-lat/0503003`.

[55] **G. de Divitiis and R. Petronzio.** Nonperturbative renormalization constants on the lattice from flavor nonsinglet Ward identities. *Phys.Lett. B419, 311–316 (1998).* `arXiv:hep-lat/9710071`.

[56] **M. Guagnelli, R. Petronzio, J. Rolf, S. Sint, R. Sommer and U. Wolff** (ALPHA). Nonperturbative results for the coefficients b_m and $b_A - b_P$ in O(a) improved lattice QCD. *Nucl. Phys. B595, 44–62 (2001).* `arXiv:hep-lat/0009021`.

[57] **J. Heitger and J. Wennekers** (ALPHA). Effective heavy light meson energies in small volume quenched QCD. *JHEP 02, 064 (2004).* `arXiv:hep-lat/0312016`.

[58] **P. Fritzsch, J. Heitger and N. Tantalo.** Non-perturbative improvement of quark mass renormalization in two-flavour lattice QCD. *JHEP 1008, 074 (2010).* `arXiv:1004.3978`.

[59] **W. Bietenholz *et al.*.** Tuning the strange quark mass in lattice simulations. *Phys. Lett. B690, 436–441 (2010).* `arXiv:1003.1114`.

[60] **P. Matthews and A. Salam.** The Green's functions of quantized fields. *Nuovo Cimento 12, 563–565 (1954).*

[61] **P. Matthews and A. Salam.** Propagators of quantized fields. *Nuovo Cimento pages 120–134 (1955).*

[62] **H. J. Rothe.** Lattice gauge theories: An Introduction. *World Sci. Lect. Notes Phys. 43, 1–381 (1992).* [World Sci. Lect. Notes Phys.82,1(2012)].

[63] **N. Metropolis and S. Ulam.** The Monte Carlo Method. *Journal of the American Statistical Association 44, 335–341 (1949).*

[64] **M. Creutz.** Monte Carlo Study of Quantized SU(2) Gauge Theory. *Phys. Rev. D21, 2308–2315 (1980).*

[65] **S. L. Adler.** An Overrelaxation Method for the Monte Carlo Evaluation of the Partition Function for Multiquadratic Actions. *Phys. Rev. D23, 2901 (1981).*

[66] **C. Whitmer.** Overrelaxation methods for Monte Carlo simulations of quadratic and multiquadratic actions. *Phys. Rev. D29, 306–311 (1984).*

[67] **M. Creutz.** Microcanonical Monte Carlo simulation. *Phys. Rev. Lett. 50, 1411 (1983).*

[68] **G. Parisi and Y.-S. Wu.** Perturbation Theory Without Gauge Fixing. *Sci. Sin. 24, 483 (1981).*

[69] **P. H. Damgaard and H. Huffel.** Stochastic Quantization. *Phys. Rept. 152, 227 (1987).*

[70] **S. Duane, A. D. Kennedy, B. J. Pendleton and D. Roweth.** Hybrid Monte Carlo. *Phys. Lett. B195, 216–222 (1987).*

[71] **P. H. Ginsparg and K. G. Wilson.** A Remnant of Chiral Symmetry on the Lattice. *Phys. Rev. D25, 2649 (1982).*

[72] **D. Weingarten and D. Petcher.** Monte Carlo integration for lattice gauge theories with fermions. *Phys. Let B99, 333–338 (1981).*

[73] **R. Frezzotti and K. Jansen.** A Polynomial hybrid Monte Carlo algorithm. *Phys. Lett. B402, 328–334 (1997).* `arXiv:hep-lat/9702016`.

[74] **A. D. Kennedy, I. Horvath and S. Sint.** A new exact method for dynamical fermion computations with nonlocal actions. *Nucl. Phys. Proc. Suppl. 73, 834–836 (1999).* `arXiv:hep-lat/9809092`.

[75] **M. A. Clark and A. D. Kennedy.** Accelerating dynamical fermion computations using the rational hybrid Monte Carlo (RHMC) algorithm with multiple pseudofermion fields. *Phys. Rev. Lett. 98, 051601 (2007).* `arXiv:hep-lat/0608015`.

[76] **M. Lüscher and F. Palombi.** Fluctuations and reweighting of the quark determinant on large lattices. *PoS LATTICE2008, 049 (2008).* `arXiv:0810.0946`.

[77] `http://cern.ch/luscher/openQCD`.

[78] **M. Lüscher.** Computational Strategies in Lattice QCD. *Modern perspectives in lattice QCD: Quantum field theory and high performance computing. Proceedings, International School, 93rd Session, Les Houches, France, August 3-28, 2009, pages 331–399 (2010).* `arXiv:1002.4232`.

[79] **Y. Saad.** Iterative Methods for Sparse Linear Systems. *SIAM, Philadelphia (2003).*

[80] **B. Jegerlehner.** Krylov space solvers for shifted linear systems. *(1996).* `arXiv:hep-lat/9612014`.

[81] **M. Hasenbusch.** Speeding up the hybrid Monte Carlo algorithm for dynamical fermions. *Phys. Lett. B519, 177–182 (2001).* `arXiv:hep-lat/0107019`.

[82] **M. Hasenbusch and K. Jansen.** Speeding up lattice QCD simulations with clover improved Wilson fermions. *Nucl. Phys. B659, 299–320 (2003).* `arXiv:hep-lat/0211042`.

[83] **C. Urbach, K. Jansen, A. Shindler and U. Wenger.** HMC algorithm with multiple time scale integration and mass preconditioning. *Comput. Phys. Commun. 174, 87–98 (2006).* `arXiv:hep-lat/0506011`.

[84] **D. H. Weingarten and J. C. Sexton.** Hamiltonian evolution for the hybrid Monte Carlo algorithm. *Nucl. Phys. Proc. Suppl. 26, 613–616 (1992).*

[85] **I. Omelyan, I. Mryglod and R. Folk.** Symplectic analytically integrable decomposition algorithms: classification, derivation, and application to molecular dynamics, quantum and celestial mechanics simulations. *Comp. Phys. Comm 151, 272–314 (2003).*

[86] **M. Lüscher.** Schwarz-preconditioned HMC algorithm for two-flavour lattice QCD. *Comput. Phys. Commun. 165, 199–220 (2005).* `arXiv:hep-lat/0409106`.

[87] **N. Madras and A. D. Sokal.** The Pivot algorithm: a highly efficient Monte Carlo method for selfavoiding walk. *J.Statist.Phys. 50, 109–186 (1988).*

[88] **U. Wolff** (ALPHA)**.** Monte Carlo errors with less errors. *Comput. Phys. Commun. 156, 143–153 (2004).* [Erratum: Comput. Phys. Commun.176,383(2007)]. `arXiv:hep-lat/0306017`.

[89] **S. Schaefer, R. Sommer and F. Virotta** (ALPHA)**.** Critical slowing down and error analysis in lattice QCD simulations. *Nucl. Phys. B845, 93–119 (2011).* `arXiv:1009.5228`.

[90] **M. Lüscher, R. Narayanan, P. Weisz and U. Wolff.** The Schrödinger functional: A Renormalizable probe for nonAbelian gauge theories. *Nucl. Phys. B384, 168–228 (1992).* `arXiv:hep-lat/9207009`.

[91] **S. Sint.** On the Schrödinger functional in QCD. *Nucl. Phys. B421, 135–158 (1994).* `arXiv:hep-lat/9312079`.

[92] **R. Sommer.** Non-perturbative QCD: Renormalization, O(a)-improvement and matching to Heavy Quark Effective Theory. *Workshop on Perspectives in Lattice QCD. Nara, Japan, October 31-November 11, 2005 (2006).* `arXiv:hep-lat/0611020`.

[93] **R. Narayanan and U. Wolff.** Two loop computation of a running coupling lattice Yang-Mills theory. *Nucl. Phys. B444, 425–446 (1995).* `arXiv:hep-lat/9502021`.

[94] **A. Bode, P. Weisz and U. Wolff** (ALPHA)**.** Two loop computation of the Schrödinger functional in lattice QCD. *Nucl. Phys. B576, 517–539 (2000).* `arXiv:hep-lat/9911018`.

[95] **P. Fritzsch, F. Knechtli, B. Leder, M. Marinkovic, S. Schaefer *et al.*.** The strange quark mass and Lambda parameter of two flavor QCD. *Nucl.Phys. B865, 397–429 (2012).* `arXiv:1205.5380`.

[96] **S. Sint and R. Sommer.** The Running coupling from the QCD Schrödinger functional: A One loop analysis. *Nucl. Phys. B465, 71–98 (1996).* `arXiv:hep-lat/9508012`.

[97] **A. Bode, U. Wolff and P. Weisz** (ALPHA)**.** Two loop computation of the Schrödinger functional in pure SU(3) lattice gauge theory. *Nucl. Phys. B540, 491–499 (1999).* `arXiv:hep-lat/9809175`.

[98] **M. Lüscher and S. Schaefer.** Lattice QCD without topology barriers. *JHEP 07, 036 (2011).* `arXiv:1105.4749`.

[99] **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. *Nucl. Phys. Proc. Suppl. 129, 408–410 (2004).* `arXiv:hep-lat/0309160`.

[100] **S. Capitani, M. Guagnelli, M. Lüscher, S. Sint, R. Sommer, P. Weisz and H. Wittig.** Non-perturbative quark mass renormalization. *Nucl. Phys. Proc. Suppl. 63, 153–158 (1998).* `arXiv:hep-lat/9709125`.

[101] **M. Lüscher.** Properties and uses of the Wilson flow in lattice QCD. *JHEP 1008, 071 (2010).* `arXiv:1006.4518`.

[102] **M. Lüscher and P. Weisz.** Perturbative analysis of the gradient flow in non-abelian gauge theories. *JHEP 02, 051 (2011).* `arXiv:1101.0963`.

[103] **M. Lüscher.** Chiral symmetry and the Yang–Mills gradient flow. *JHEP 1304, 123 (2013).* `arXiv:1302.5246`.

[104] **A. Ramos.** The Yang-Mills gradient flow and renormalization. *PoS LATTICE2014, 017 (2015).* `arXiv:1506.00118`.

[105] **P. Fritzsch and A. Ramos.** The gradient flow coupling in the Schrödinger Functional. *JHEP 1310, 008 (2013).* `arXiv:1301.4388`.

[106] **M. Lüscher.** Step scaling and the Yang-Mills gradient flow. *JHEP 06, 105 (2014).* `arXiv:1404.5930`.

[107] **A. Bazavov *et al.* (MILC).** Symanzik flow on HISQ ensembles. *PoS LATTICE2013, 269 (2014).* `arXiv:1311.1474`.

[108] **Z. Fodor, K. Holland, J. Kuti, D. Nogradi and C. H. Wong.** The Yang-Mills gradient flow in finite volume. *JHEP 1211, 007 (2012).* `arXiv:1208.1051`.

[109] **O. Bär and M. Golterman.** Chiral perturbation theory for gradient flow observables. *Phys. Rev. D89, 034505 (2014).* `arXiv:1312.4999`.

[110] **A. Ramos and S. Sint.** On O(a^2) effects in gradient flow observables. *PoS LATTICE2014, 329 (2015).* `arXiv:1411.6706`.

[111] **A. Ramos and S. Sint.** Symanzik improvement of the gradient flow in lattice gauge theories. *Eur. Phys. J. C76(1), 15 (2016).* `arXiv:1508.05552`.

[112] **A. M. Polyakov.** Quark Confinement and Topology of Gauge Groups. *Nucl. Phys. B120, 429–458 (1977).*

[113] **F. Bruckmann.** Topological objects in QCD. *Eur. Phys. J. ST 152, 61–88 (2007).* `arXiv:0706.2269`.

[114] **A. A. Belavin, A. M. Polyakov, A. S. Schwartz and Yu. S. Tyupkin.** Pseudoparticle Solutions of the Yang-Mills Equations. *Phys. Lett. B59, 85–87 (1975).*

[115] **G. 't Hooft.** Computation of the Quantum Effects Due to a Four-Dimensional Pseudoparticle. *Phys. Rev. D14, 3432–3450 (1976).* [Erratum: Phys. Rev.D18,2199(1978)].

[116] **D. Diakonov.** Chiral symmetry breaking by instantons. *Proc. Int. Sch. Phys. Fermi 130, 397–432 (1996).* `arXiv:hep-ph/9602375`.

[117] **S. Vandoren and P. van Nieuwenhuizen.** Lectures on instantons. *(2008).* `arXiv:0802.1862`.

[118] **L. Del Debbio, L. Giusti and C. Pica.** Topological susceptibility in the SU(3) gauge theory. *Phys. Rev. Lett. 94, 032003 (2005).* `arXiv:hep-th/0407052`.

[119] **M. F. Atiyah and I. M. Singer.** The Index of elliptic operators. 5. *Annals Math. 93, 139–149 (1971).*

[120] **M. E. Peskin and D. V. Schroeder.** An Introduction to quantum field theory. *Reading, USA: Addison-Wesley (1995).*

[121] **E. Witten.** Current Algebra Theorems for the U(1) Goldstone Boson. *Nucl. Phys. B156, 269–283 (1979).*

[122] **G. Veneziano.** U(1) Without Instantons. *Nucl. Phys. B159, 213–224 (1979).*

[123] **M. Lüscher.** Topological effects in QCD and the problem of short distance singularities. *Phys. Lett. B593, 296–301 (2004).* `arXiv:hep-th/0404034`.

[124] **P. Hasenfratz, V. Laliena and F. Niedermayer.** The Index theorem in QCD with a finite cutoff. *Phys. Lett. B427, 125–131 (1998).* `arXiv:hep-lat/9801021`.

[125] **M. Lüscher and F. Palombi.** Universality of the topological susceptibility in the SU(3) gauge theory. *JHEP 09, 110 (2010).* `arXiv:1008.0732`.

[126] **A. Phillips and D. Stone.** Lattice Gauge Fields, Principal Bundles and the Calculation of Topological Charge. *Commun. Math. Phys. 103, 599–636 (1986).*

[127] **M. Lüscher.** Topology of Lattice Gauge Fields. *Commun. Math. Phys. 85, 39 (1982).*

[128] **M. Lüscher.** Trivializing maps, the Wilson flow and the HMC algorithm. *Commun. Math. Phys. 293, 899–919 (2010).* `arXiv:0907.5491`.

[129] **M. Bruno, S. Schaefer and R. Sommer** (ALPHA)**.** Topological susceptibility and the sampling of field space in $N_f = 2$ lattice QCD simulations. *JHEP 08, 150 (2014).* `arXiv:1406.5363`.

[130] **M. Lüscher and S. Schaefer.** Non-renormalizability of the HMC algorithm. *JHEP 04, 104 (2011).* `arXiv:1103.1810`.

[131] **P. Fritzsch, A. Ramos and F. Stollenwerk.** Critical slowing down and the gradient flow coupling in the Schrödinger functional. *PoS Lattice2013, 461 (2014).* `arXiv:1311.7304`.

[132] **O. Bär, G. Rupak and N. Shoresh.** Chiral perturbation theory at O(a^2) for lattice QCD. *Phys. Rev. D70, 034508 (2004).* `arXiv:hep-lat/0306021`.

[133] **T. Bhattacharya, R. Gupta, W. Lee, S. R. Sharpe and J. M. Wu.** Improved bilinears in lattice QCD with non-degenerate quarks. *Phys. Rev. D73, 034504 (2006).* `arXiv:hep-lat/0511014`.

[134] **M. Kurth and R. Sommer.** An Explicit example of HQET at one loop order of perturbation theory. *Nucl. Phys. Proc. Suppl. 106, 355–357 (2002).* `arXiv:hep-lat/0110096`.

[135] **M. Bruno, J. Finkenrath, F. Knechtli, B. Leder and R. Sommer** (ALPHA). Effects of Heavy Sea Quarks at Low Energies. *Phys. Rev. Lett. 114(10), 102001 (2015).* `arXiv:1410.8374`.

[136] **F. Tekin, R. Sommer and U. Wolff** (ALPHA). The Running coupling of QCD with four flavors. *Nucl.Phys. B840, 114–128 (2010).* `arXiv:1006.0672`.

[137] **S. Aoki *et al.*** (PACS-CS). Precise determination of the strong coupling constant in $N_f = 2 + 1$ lattice QCD with the Schrödinger functional scheme. *JHEP 0910, 053 (2009).* `arXiv:0906.3906`.

[138] **P. Fritzsch, R. Sommer, F. Stollenwerk and U. Wolff.** Determination of $c_{\rm sw}$ in $N_f = 3 + 1$ Lattice QCD with massive Wilson fermions. *PoS LATTICE2014, 293 (2015).* `arXiv:1501.06753`.

[139] **Z. Fodor, K. Holland, J. Kuti, S. Mondal, D. Nogradi *et al.*.** The lattice gradient flow at tree-level and its improvement. *JHEP 1409, 018 (2014).* `arXiv:1406.0827`.

[140] **M. Guagnelli, J. Heitger, R. Sommer and H. Wittig** (ALPHA). Hadron masses and matrix elements from the QCD Schrödinger functional. *Nucl. Phys. B560, 465–481 (1999).* `arXiv:hep-lat/9903040`.

[141] **J. Bulava, M. Della Morte, J. Heitger and C. Wittemeier** (ALPHA). Non-perturbative improvement of the axial current in N_f=3 lattice QCD with Wilson fermions and tree-level improved gauge action. *Nucl. Phys. B896, 555–568 (2015).* `arXiv:1502.04999`.

[142] **B. Blossier *et al.*** (ALPHA). Parameters of Heavy Quark Effective Theory from $N_f = 2$ lattice QCD. *JHEP 1209, 132 (2012).* `arXiv:1203.6516`.

[143] **F. Bernardoni *et al.*.** The b-quark mass from non-perturbative $N_f = 2$ Heavy Quark Effective Theory at $O(1/m_h)$. *Phys. Lett. B730, 171–177 (2014).* `arXiv:1311.5498`.

[144] **B. Blossier, M. Della Morte, N. Garron and R. Sommer.** HQET at order $1/m$: I. Non-perturbative parameters in the quenched approximation. *JHEP 1006, 002 (2010).* `arXiv:1001.4783`.

[145] **J. Heitger and A. Jüttner.** Lattice cutoff effects for F_{D_s} with improved Wilson fermions: A Final lesson from the quenched case. *JHEP 0905, 101 (2009).* `arXiv:0812.2200`.

[146] **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. *Phys.Rev. D82, 034512 (2010).* `arXiv:1004.4285`.

[147] **J. Heitger, G. M. von Hippel, S. Schaefer and F. Virotta.** Charm quark mass and D-meson decay constants from two-flavour lattice QCD. *PoS LATTICE2013, 475 (2013).* `arXiv:1312.7693`.

[148] **M. Della Morte, R. Sommer and S. Takeda.** On cutoff effects in lattice QCD from short to long distances. *Phys. Lett. B672, 407–412 (2009).* `arXiv:0807.1120`.

[149] **M. Dalla Brida, P. Fritzsch, T. Korzec, A. Ramos, S. Sint and R. Sommer** (ALPHA). *Slow running of the Gradient Flow coupling from 200 MeV to 4 GeV in $N_\mathrm{f} = 3$ QCD (2016).* `arXiv:1607.06423`.

[150] **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. *Phys. Lett. B372, 275–282 (1996).* `arXiv:hep-lat/9512009`.

[151] **K. Jansen and R. Sommer.** $\mathcal{O}$(a) improvement of lattice QCD with two flavors of Wilson quarks *pages `[hep-lat/9803017v3]` (2002).*

[152] **C. Wittemeier.** Implementation of a Program for QCD and HQET correlation Functions in the Schrödinger Functional. *diploma thesis (2012).*

[153] **A. Vladikas.** Three Topics in Renormalization and Improvement. *Modern perspectives in lattice QCD: Quantum field theory and high performance computing. Pro-*

ceedings, International School, 93rd Session, Les Houches, France, August 3-28, 2009, pages 161–222 (2011). `arXiv:1103.1323`.

[154] **M. Testa.** Some observations on broken symmetries. *JHEP 04, 002 (1998).* `arXiv:hep-th/9803147`.

[155] **L. H. Karsten and J. Smit.** Lattice Fermions: Species Doubling, Chiral Invariance, and the Triangle Anomaly. *Nucl. Phys. B183, 103 (1981).*

[156] **M. Bochicchio, L. Maiani, G. Martinelli, G. C. Rossi and M. Testa.** Chiral Symmetry on the Lattice with Wilson Fermions. *Nucl. Phys. B262, 331 (1985).*

[157] **S. Dürr and M. Della Morte.** Exploring two nonperturbative definitions of c_A. *Nucl. Phys. Proc. Suppl. 129, 417–419 (2004).* `arXiv:hep-lat/0309169`.

[158] **T. Kaneko, S. Aoki, M. Della Morte, S. Hashimoto, R. Hoffmann and R. Sommer** (ALPHA, CP-PACS/JLQCD)**.** Non-perturbative improvement of the axial current with three dynamical flavors and the Iwasaki gauge action. *JHEP 04, 092 (2007).* `arXiv:hep-lat/0703006`.

[159] **U. Wolff, P. Galler and T. Korzec.** Lecture Notes on Computational Physics I [in german], Humboldt University, Berlin. *(2014).* `https://www.physik.hu-berlin.de/de/com/teachingandseminars/previousCPI`.

[160] **D. Savoiu and G. Quast.** KAFE - Karlsruhe Fit Environment. `http://www.ekp.kit.edu/~quast`.

[161] **R. Brun and F. Rademakers.** ROOT - An Object Oriented Data Analysis Framework,. *Proceedings AIHENP'96 Workshop, Lausanne, Sep. 1996, Nucl. Inst. & Meth. in Phys. Res. A 389 (1997) 81-86.* `http://root.cern.ch/`.

[162] **E. Jones, T. Oliphant, P. Peterson *et al.*.** SciPy: Open source scientific tools for Python. *(2001–).* `http://www.scipy.org/`.

List of Figures

List of Tables

Acknowledgements

I would like to thank the following people who helped me one way or the other to finish this thesis:

- Ulli Wolff and Rainer Sommer,
  for giving me the opportunity to work on an interesting subject in my favorite branch of physics and favorite city, as well as their guidance, advice and patience.
- Patrick Fritzsch and Tomasz Korzec,
  for their willingness to always answer my questions, occasional encouragement and their continuous support, which was far beyond what can be taken for granted.
- Alberto Ramos and Oliver Bär,
  for collaboration, sharing insights and especially for their talent and effort to make things understandable for novices of the field
- Till Martini and Christoph Meyer,
  for letting me forget about the fact that I was the only PhD student in my research group for most of the time.
- My incompetent teachers back in school,
  who made me repeat 10th grade because of insufficient performances in physics, which gave me an extra piece of motivation to become a physicist.
- Last but most importantly, my family for everything.

Selbständigkeitserklärung

Ich erkläre, dass ich die Dissertation selbständig und nur unter Verwendung der von mir gemäß § 7 Abs. 3 der Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät, veröffentlicht im Amtlichen Mitteilungsblatt der Humboldt-Universität zu Berlin Nr. 126/2014 am 18.11.2014 angegebenen Hilfsmittel angefertigt habe.

Berlin, den 14.12.2016,

Felix Stollenwerk