Properties of the η and η' mesons from lattice QCD

Dissertation zur Erlangung
des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)
der Fakultät für Physik

der

Universität Regensburg

vorgelegt von

Jakob Fabian Simeth

aus Regensburg im Jahr 2022

Promotionsgesuch eingereicht am: 08. Juli 2020
Die Arbeit wurde angeleitet von: Prof. Dr. Gunnar S. Bali
Datum Promotionskolloquium: 22. Februar 2022

Prüfungsausschuss:

Vorsitzender:	Prof. Dr. Karsten Rincke
1. Gutachter:	Prof. Dr. Gunnar S. Bali
2. Gutachter:	Prof. Dr. Vladimir Braun
weiterer Prüfer:	Prof. Dr. Ferdinand Evers

Abstract

Properties of the η and η' mesons from lattice QCD

by Jakob Fabian Simeth

In this thesis we compute masses, decay constants and gluonic matrix elements of the flavour-diagonal pseudoscalar mesons η and η' from lattice QCD.

To control all relevant systematic errors we employ $N_f = 2 + 1$ flavour simulations along two distinct quark mass trajectories leading to and including the physical point. The continuum extrapolation is guided by four lattice spacings. The ensembles were generated within the coordinated lattice simulations initiative, and we set their relative scales in this work.

We discuss noise reduction techniques for the efficient calculation of the disconnected contributions that are important building blocks of the relevant correlation functions of the η and η' system. The physical states are no flavour eigenstates, and hence sophisticated analysis methods are required to extract them from the data. We develop a matrix generalization of the effective mass method which we employ in conjunction with additional techniques to determine masses and matrix elements.

The physical point extrapolation employs next-to-leading order large-N_c chiral perturbation theory, and we determine all relevant low energy constants. For the first time also their renormalization scale dependence is taken into account, and this provides an important check of the range of validity of this effective field theory with implications on many existing phenomenological analyses.

Our physical point results for the masses are in agreement with experimental values and read $M_\eta = 554.7(9.2)\,\text{MeV}$ and $M_{\eta'} = 930(21)\,\text{MeV}$. The determination of the four η and η' decay constants is the first from first principles and we obtain $F^8 = 115.0(2.8)\,\text{MeV}$ and $\theta_8 = -25.8(2.3)^\circ$ in the octet channel and $F^0 = 100.1(3.0)\,\text{MeV}$ and $\theta_0 = -8.1(1.8)^\circ$ for the singlet in the $\overline{\text{MS}}$ scheme at 2 GeV. These results are in excellent agreement with phenomenological determinations and at a similar level of precision.

Finally, we connect these axialvector decay constants with pseudoscalar and gluonic matrix elements to test the axial Ward identities, and predict the anomalous matrix elements to be $a_\eta = \langle\Omega|2\omega|\eta\rangle = 0.0170(10)\,\text{GeV}^3$ and $a_{\eta'} = \langle\Omega|2\omega|\eta'\rangle = 0.0381(84)\,\text{GeV}^3$ at the physical point and $\mu = 2\,\text{GeV}$.

Contents

List of Figures

List of Tables

1 Introduction

The discovery of the kaon in 1947 initiated a small revolution in particle physics: With the advent of accelerators more and more particles were found, and the vast number did not fit the hypothesis that all these are elementary. In 1961 Gell-Mann introduced his "Eightfold Way" [1] and Ne'eman [2] independently started to make sense of the ever-growing number of observed particles. They were able to classify hadrons in terms of representations of the SU(3) group, eventually leading to the concept of quarks. These are the constituents of the observed particles and the up, down and strange quark flavours are represented by the fundamental representation of the group. Only later this was completed with a colour gauge symmetry and formulated in a quantum field theory now known as quantum chromodynamics (QCD), an important building block in the standard model of particle physics. In that same year of the formulation of the Eightfold Way also the η meson was discovered [3] and mostly fitted into the classification scheme. Also, the Ω^- baryon [4] was successfully predicted and in fact most of the observed hadron spectrum could be explained to a reasonable precision.

The assigned quantum numbers could then be linked to the masses of the hadrons using the Gell-Mann–Okubo mass formula [1, 5, 6] using just a few proportionality constants. The observed lightness of pseudoscalar hadrons in comparison to their vector counterparts could be explained as a consequence of the spontaneous breaking of chiral symmetry. The pseudoscalar mesons then are the pseudo-Goldstone bosons of the breaking of the $\mathrm{SU}(3)_A$ axial component of this approximate global symmetry. In the decomposition of SU(3) into octet particles and a singlet one, however, only the octet ones had been observed, and there was no candidate for a sufficiently light pseudoscalar singlet state. A newly discovered state that would later be named η' [7] was first considered to be too heavy to align with the other pseudoscalar mesons and was excluded due to the observed decay $\eta' \to 3\pi^0$ that violates isospin symmetry [1].

This became known as the η–η' puzzle: The large mass difference between the η' and the octet particles could not be explained by their flavour contents. This problem could be linked to the breaking of the $\mathrm{U}(1)_A$ axial symmetry by anomalous quantum corrections [8, 9] that arise through the quantization of the theory and seemed artificial at first. Instantons — pseudo-particle solutions to the classical field equations in Euclidean time with non-trivial topology — were suggested to resolve this issue [10, 11]. The limit of an infinite number of colours and vanishing quark masses was studied and gave the Witten–Veneziano relation [12–14],

$$M_0^2 = \frac{2N_f}{F^2}\tau_0, \tag{1.1}$$

where M_0 is the singlet mass, F is the pseudoscalar decay constant in the chiral limit, N_f is the number of active flavours and τ_0 is the topological susceptibility in pure Yang-Mills theory, i.e., without dynamical quarks. The relation makes it very clear that the singlet component acquires a non-vanishing mass through the anomaly even in the chiral limit, leading to the observed mass difference of the states that have a singlet admixture relative to the pure octet states, i.e., predominantly to the η', but also to the η.

This means that both the η' and the η are superpositions of flavour octet and singlet combinations of quark fields. This so-called mixing is closely related to SU(3) flavour breaking — if

flavour symmetry was exact then the singlet and octet would be cleanly separated states — and so the flavour diagonal mesons π^0, η and η' are ideal to study differences between quark masses, for example the difference $m_u - m_d$ from the isospin violating decays $\eta \to 3\pi$, see the recent results by KLOE [15], for instance. The sensitivity to the topological structure of the vacuum through the flavour singlet components is what makes η/η' physics so interesting also for beyond the standard model searches. Axion-like particles a could, for example, be detected through violation of charge conjugation via processes like $\eta^{(\prime)} \to 2\pi^0 a \to 2\pi^0 l^+ l^-$, where $l^\pm$ are (anti-)leptons of any flavour [16]. The transition form factors $\gamma\gamma^* \to \eta^{(\prime)}$ are also important input for the hadronic contribution to the anomalous magnetic moment of the muon $(g-2)_\mu$, that attracts a lot of current interest due to the long-standing large discrepancy between the standard model prediction and the BNL experiment [17] that has recently been confirmed at Fermilab [18], see, e.g. [19] for a review.[1]

Experimental insights for such two-photon processes could be gained from the B-meson factories at BabaR and Belle [21–23]. There have also been extensive measurements from the KLOE and KLOE-2 [24] experiments and at BES-III [25]. Upcoming experiments at KEK, Belle II [26, 27], and Jefferson Lab, the "JLab Eta Factory" [28], will tremendously increase statistics and give results at higher, still mostly unexplored center-of-mass energies.

From the theory side a clean assessment of the η and η' system is complicated by the aforementioned mixing of the singlet and octet flavour combinations to form the physical states and the coupling to non-trivial topological configurations. The latter is a deeply non-perturbative phenomenon that defies the typical perturbative treatment and requires methods that are applicable in a low energy regime of QCD. Modern phenomenological analyses of η-η' mixing are largely based on large-N_c chiral perturbation theory (ChPT) [29–33] which allows for a unified treatment of the $\eta^{(\prime)}$-mesons together with the pseudo-Goldstone octet of the lightest pseudoscalars by simultaneously expanding around small quark masses and an infinite number of colours — a limit in which the anomaly vanishes. When combined with dispersion relations, this approach provides a quantitative description of a large variety of $\eta^{(\prime)}$ decays and low energy $\eta^{(\prime)}$ production processes, see, e.g., [16] and references therein. The convergence of this expansion, however, is hard to quantify, and additional input is needed to fix the low energy constants that are in general dependent on the energy scale of the process.

Ab initio lattice simulations are technically demanding due to the computationally expensive evaluation of disconnected contributions and the coupling to the topological charge, which results in large autocorrelation times and requires long time series to enable an adequate sampling of the topological sectors. Moreover, the extraction of ground state properties from correlation functions with a noise over signal ratio that increases rapidly in Euclidean time requires optimized methods.

Despite these challenges, steady progress has been made in computing the masses of the η and η' mesons, starting in the quenched approximation [34–36], and continuing with $N_f = 2$ mass-degenerate dynamical light quarks [35, 37–44]. In the latter case only one η meson exists, which is a pure singlet state, and no flavour mixing takes place. More realistic simulations of nature require an additional strange quark ($N_f = 2 + 1$) [45–48] (see also [49] for a different attempt using correlators of the topological charge density). More recently, $N_f = 2{+}1{+}1$ results [50–52] employing the twisted-mass fermion formulation, using several ensembles and lattice spacings, enabled a physical point extrapolation. In [53] the η' mass was calculated at non-zero temperature from topological charge density correlators. A very recent computation focused on $N_f = 1 + 1 + 1$ QCD+QED simulations [54], investigating the mass splitting of the η and the π^0 close to the unphysically heavy $N_f = 3$ flavour symmetric point and computing pseudoscalar matrix elements. These were also determined in [52]: Relating them to the four decay constants of the η/η' system made the first lattice determination possible, to a precision that is on par with phenomenological studies. Another lattice computation of these matrix elements was carried out in $N_f = 2 + 1$, in the context of a

[1] Recent results from the lattice, however, move the theory prediction closer to the experimental results and report a smaller deviation [20].

calculation of the semileptonic decay form factors $D_s \to \eta, \eta'$, albeit only on two ensembles at a single lattice spacing [48].

In this thesis we employ $N_f = 2 + 1$ lattice QCD to determine the masses of the η and η' mesons and, for the first time, also all four decay constants, and gluonic matrix elements extrapolated to the physical point and without relying on model assumptions. To this end twenty-one ensembles at various quark mass combinations, including the physical quark mass point and at four different lattice spacings are analysed. A combination of noise reduction techniques are employed, and refined analysis methods make it possible to construct the physical states reliably and lead to reduced statistical errors. A careful assessment of the systematic errors involved leads to realistic total error estimates and to final results that are currently the most precise theory predictions for the η/η' masses, decay constants and gluonic matrix elements.

The thesis is structured as follows: In ch. 2 we introduce the reader to QCD and its discretization making it amenable to simulations. In doing so, we focus on the peculiarities of the flavour diagonal η and η' mesons. Ch. 3 is devoted to the generation of ensembles within the coordinated lattice simulations (CLS) initiative that are being used in the analysis. Emphasis is put on taking the continuum limit and the determination of the gradient flow scale at the individual values of the lattice coupling. Using these ensembles, we elaborate on methods for an efficient estimation of stochastic loops contributing to the correlators of flavour diagonal mesons in ch. 4. There, we also make the connection from correlators of interpolating operators to the physical eigenstates of the η and η' mesons that allow us to extract masses and matrix elements. The continuum and chiral limit of these is taken in ch. 5 employing large-N_c ChPT as a parametrization of the mass dependence in the continuum. Low energy constants are determined that provide important input to other related calculations. In ch. 6 the Ward identities are tested, including the anomaly contribution to the singlet axialvector current. We determine the gluonic matrix elements in two ways, both indirectly through pseudoscalar matrix elements and taking the Ward identity for granted as well as directly by evaluating the anomaly matrix elements themselves. We finally conclude and summarize our findings in ch. 7.

2 Quantum Chromodynamics on and off the lattice

The standard model of particle physics is very successful in describing the interactions of matter on a microscopic level. It is formulated in terms of a field theory with a global gauge symmetry $\mathrm{U}(1) \times \mathrm{SU}(2) \times \mathrm{SU}(3)$, where $\mathrm{U}(1) \times \mathrm{SU}(2)$ is the subgroup of the unified electroweak interactions. Quantum chromodynamics is the $\mathrm{SU}(3)$ gauge theory of quarks and gluons and describes their strong interactions. Quarks carry colour-charge which is mediated by gluons, the gauge bosons of the theory. The number of colour charges, N_c, is three in nature and hence the number of gluons is $N_c^2 - 1 = 8$.

The quark fields appear in three families of quarks where each family consists of two quark flavours. In the context of this work the lightest three quarks are the most relevant degrees of freedom and these are the up, down and the strange quarks, while the heavier quarks effectively decouple from low-energy physics. Therefore, we usually set the number of active flavours to $N_f = 3$.

In this chapter, we will present the QCD Lagrangian and some of the features emerging from it in sec. 2.1. The presentation follows well-known text books, e.g. [55, 56], that should be consulted for more details and references. After briefly discussing asymptotic freedom and confinement in secs. 2.1.3 and 2.1.4, we focus on global symmetries of the Lagrangian, the axial anomaly and the implications on the hadron spectrum in 2.1.5. We discuss the quark flavour model predating QCD and its failure to describe the η and η' mesons in sec. 2.1.6.

Many of these observations are rooted in the deeply non-perturbative nature of QCD and this renders perturbative expansions — that proved successful in the electroweak sector of the standard model — useless in the low energy regime of QCD. We discuss a possible discretization of QCD on a four-dimensional space-time lattice in sec. 2.2 that allows ab initio simulations, i.e., without resorting to perturbation theory or model assumptions. This introduces lattice QCD, the main tool in this work, that is also well covered by a range of text books; in particular [57–59] proved useful. The process of generating gauge configurations with three dynamical quarks is described in sec. 2.2.4.

Sec. 2.3 is devoted to the topic of hadron spectroscopy on the lattice, going into more detail on the technicalities involved and aiming at the construction of η and η' physical states: We start by defining interpolating operators and how their overlap with the ground states can be improved in secs. 2.3.1 and 2.3.2, respectively. In the subsequent sections, we discuss the emergence and computation of disconnected contributions to propagators in sec. 2.3.3, the further analysis of two-point functions to extract masses in sec. 2.3.4 and local matrix elements in sec. 2.3.5.

Finally, in sec. 2.4, we then remove the ultraviolet regulator of the lattice theory by renormalization while taking the continuum limit and removing leading discretization errors by "improving" the currents. Thereby, we focus on simple quark-antiquark currents without derivative that are required for the calculation of the decay constants and emphasize the differences in the renormalization and improvement of flavour singlet and non-singlet combinations.

Parts of the last section in this chapter, sec. 2.4, have been published in similar or, in places, verbatim form in [60].

2.1 The QCD Lagrangian and its properties in the continuum

The Lagrangian of quantum chromodynamics can be decomposed into

$$\mathcal{L}_{\mathrm{QCD}}(g_s, \{m_f\}) = \mathcal{L}_{\mathrm{YM}}(g_s) + \sum_f \mathcal{L}_f(g_s, m_f), \tag{2.1}$$

where $\mathcal{L}_{\mathrm{YM}}$ is the pure Yang-Mills part and $\mathcal{L}_f$ adds fermionic degrees of freedom for each of the N_f quark flavours $f = u, d, s, \ldots$. In the quantized theory, additional terms are added that fix the gauge and introduce the necessary Fadeev–Popov ghost fields that we do not discuss here.

The fermion masses $\{m_f\}$ and the strong coupling strength g_s are the only free parameters in this sector of the standard model[1] and can explain the mass spectrum of observed hadrons and their strong interactions to high accuracy. Corrections to QCD predictions from the electroweak sector of the standard model can be treated perturbatively due to the smallness of the electromagnetic gauge coupling.

2.1.1 Yang-Mills theory

Many of the non-perturbative features of QCD are already present in the purely gluonic (Yang-Mills) theory without interacting quarks. Its Lagrangian is

$$\mathcal{L}_{\mathrm{YM}} = -\frac{1}{4} F^a_{\mu\nu} F^{a,\mu\nu}, \tag{2.2}$$

where repeated indices are summed over and greek letters usually label space-time indices and the roman letters $a, b, c, \ldots$ are colour indices in the adjoint representation. The field strength tensor is defined as

$$F_{\mu\nu} = \frac{i}{g_s} \left[D_\mu, D_\nu \right]. \tag{2.3}$$

We define it in terms of the covariant derivative acting on a continuous function $f(x)$,

$$\hat{n}_\mu D_\mu f(x) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left(f(x + \epsilon\hat{n}) - \mathcal{U}(x + \epsilon\hat{n}, x) f(x) \right), \tag{2.4}$$

where $\hat{n}$ is a unit vector in some direction and $\mathcal{U}(x + \epsilon\hat{n}, x)$ is a gauge transporter that guarantees gauge invariance by connecting the points x and $x + \epsilon\hat{n}$. The defining property of the gauge transporters is their behaviour under local gauge transformations, $\Omega(x)$,

$$\mathcal{U}(y, x) \to \Omega(y)^\dagger \mathcal{U}(y, x) \Omega(x). \tag{2.5}$$

These gauge transformations are group elements of SU(N_c) and thus can be parametrized in terms of a continuous function $\theta : \mathbb{R}^4 \mapsto \mathbb{R}^{N_c^2 - 1}$ that maps a group element to every point in space-time,

$$\Omega(x) = \exp\left(i\theta^a(x) t^a \right), \tag{2.6}$$

where $t^a = \lambda^a/2$ with $a = 1, \ldots, N_c^2 - 1$ are the generators of SU(N_c) and in QCD $N_c = 3$. In this case, λ^a are the eight Gell-Mann matrices. The generators are traceless and Hermitian 3×3 matrices and satisfy the commutation algebra

$$\left[t^a, t^b \right] = i f^{abc} t^c, \tag{2.7}$$

where f^{abc} are the anti-symmetric structure constants of the group and the normalization corresponds to $\mathrm{tr}\, t^a t^b = \delta^{ab}/2$, see app. A.3.

[1] There is the theoretical possibility of a so-called theta term that will be briefly discussed in the next section.

The gauge transporter, eq. (2.5), can be defined as

$$\mathcal{U}(y,x) = \mathcal{P}\exp\left(ig_s\int_{C_{x,y}} \mathrm{d}s_\mu A_\mu(s)\right), \tag{2.8}$$

where $\mathcal{P}$ is a path ordering operator acting on paths $C_{x,y}$ that connect x and y. The gauge is mediated by the gauge bosons, and thus here gluon fields appear that transform according to the adjoint representation of the group, $A_\mu(x) = A_\mu^a t^a$. Trivially, for $y = x$, we obtain $U(x,x) = \mathbb{1}$ and for the limiting case of very small distances ϵ we can equally write

$$\mathcal{U}(x+\epsilon\hat{n},x) = \mathbb{1} + ig_s\epsilon\hat{n}_\mu A_\mu(x) + \mathcal{O}(\epsilon^2). \tag{2.9}$$

This allows us to write down an explicit form for the covariant derivative, eq. (2.4),

$$D_\mu = \partial_\mu - ig_s A_\mu(x), \tag{2.10}$$

and we can compute the gluonic field strength tensor using eq. (2.3),

$$F_{\mu\nu}^a(x) = \partial_\mu A_\nu^a(x) - \partial_\nu A_\mu^a(x) + g_s f^{abc} A_\mu^b(x) A_\nu^c(x). \tag{2.11}$$

From this and the gauge invariance of the Lagrangian, the transformation properties of the gluon field under gauge transformations immediately follow,

$$A_\mu(x) \to \Omega(x) A_\mu(x) \Omega^\dagger(x) + i(\partial_\mu \Omega(x))\Omega^\dagger(x). \tag{2.12}$$

The last term in eq. (2.11) arises from the commutator and is entirely due to the non-abelian nature of the gauge group. This is different to quantum electrodynamics where the gauge group is $\mathrm{U}(1)$ and thus abelian. Self-interaction of photons only takes place through virtual fermion loops and therefore is strongly suppressed. In QCD, however, gluons couple to another already at tree level. This gives rise to many of the non-perturbative phenomena and also to asymptotic freedom discussed in 2.1.3.

Gauge and Lorentz invariance allow for an additional topological contribution to eq. (2.2),

$$-\theta_0 N_f \frac{g_s^2}{16\pi^2} F_{\mu\nu}^a \widetilde{F}_{\mu\nu}^a, \tag{2.13}$$

where $\widetilde{F}_{\mu\nu} = \epsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}$ is the dual field strength tensor and $\epsilon_{\mu\nu\rho\sigma}$ is the totally anti-symmetric tensor with the convention $\epsilon_{0123} = 1$. If the coupling parameter θ_0 was non-zero, then this term would give rise to a violation of charge and parity (CP) symmetry that would manifest itself in a non-vanishing electric dipole moment of the neutron which has been ruled out experimentally to very high precision [61]. This gives rise to the strong CP problem, since in QCD there is no inherent reason why θ_0 should vanish.

2.1.2 Fermions

In nature we find six flavours of quarks with increasing masses. The up and down quarks are the lightest of all quarks. They have masses defined in the modified minimal subtraction ($\overline{\mathrm{MS}}$) scheme of about 2 MeV and 5 MeV, respectively [19]. The strange quark has a mass of $92.9(7)$ MeV, already about 27 times as heavy as the average mass of the up and down quarks. The charm quark mass is $1.280(13)$ GeV and heavier than most light hadrons. Therefore, its dynamical contribution is small [62]. At scales below the respective thresholds, effects from even heavier quarks can safely be neglected and the masses of these heavy quarks can effectively be set to infinite. This mass

hierarchy also allows for another approximation and treat some of the lightest quarks as massless. This enables expansions around small quark masses in the framework of chiral perturbation theory.

The fermionic part of the Lagrangian, eq. (2.1) is just the Dirac equation for every quark flavour $f = u, d, s, \ldots$ with mass m_f,

$$\mathcal{L}_f = \overline{\psi}_f (i\not{D} - m_f)\psi_f. \tag{2.14}$$

The fermion spinors, ψ_f, carry spin and colour degrees of freedom and the gauge fields enter the Dirac operator through the covariant derivative, eq. (2.10). Here, we use the Feynman-Slash notation $\not{D} = \gamma_\mu D_\mu$. The Dirac matrices γ_μ are 4×4 matrices and fulfil the anticommutator relation

$$\{\gamma_\mu, \gamma_\nu\} = 2g_{\mu\nu}, \tag{2.15}$$

where $g_{\mu\nu}$ is the metric tensor. Therefore, from counting the degrees of freedom we see that the particular choice of the representation is not unique and also it is different in Minkowskian or Euclidean metric, see also app. A.2 for their relations and definitions used in this work. Again, from gauge invariance we can immediately deduce that fermions transform trivially under gauge transformations,

$$\psi(x) \to \Omega(x)\psi(x), \qquad \overline{\psi}(x) \to \overline{\psi}(x)\Omega^\dagger(x). \tag{2.16}$$

2.1.3 Asymptotic freedom

Apart from the quark masses (and θ_0), the only free parameter of QCD is the gauge coupling g_s. As in other renormalizable field theories the definition of the coupling depends on the regularization and on the energy scale of the process under consideration. Therefore, values for the coupling are given at specific energies μ, e.g. at the mass of the Z-boson $m_Z \approx 91$ GeV. A recent lattice QCD determined the coupling at that scale to be $a_s(\mu^2 = m_Z^2) = 0.0375(32)$ [63]. The *running* of this value with the energy scale is determined by the β-function, which itself is parameter free and can be derived directly from the theory. It describes the change of the coupling with the energy scale and is defined as

$$\beta(\mu) = \mu^2 \frac{\mathrm{d}a_s(\mu)}{\mathrm{d}\mu^2}, \tag{2.17}$$

where $a_s = \alpha_s/\pi = g_s^2/(4\pi^2)$. It can be expressed as a perturbative series in the coupling itself,

$$\beta(\mu) = \sum_{i\geq 0} \beta_i a_s(\mu)^{i+2}. \tag{2.18}$$

The first two coefficients are "universal" and independent of the regularization employed,

$$\beta_0 = -\frac{1}{4}\left(\frac{11}{3}C_A - \frac{2}{3}N_f\right), \tag{2.19}$$

$$\beta_1 = -\frac{1}{16}\left[\frac{34}{3}C_A^2 - \left(\frac{10}{3}C_A + 2C_F\right)N_f\right], \tag{2.20}$$

where in QCD $C_A = 3$ and $C_F = 4/3$.

The sign of the first coefficient of the β-function, $\beta_0 < 0$ for $N_f \leq 16$, gives rise to the most prominent property of QCD, asymptotic freedom [64, 65], i.e., that the coupling strength decreases with increasing scale and eventually vanishes as the scale approaches infinity. For hard processes, corresponding to small distances, where $\mu > 100\,\mathrm{GeV}$, $a_s \lesssim 0.1$ and a perturbative treatment is possible. At low scales $\mu < 2\,\mathrm{GeV}$ non-perturbative effects dominate, the coupling becomes large and at a scale Λ_{QCD} the coupling eventually diverges. In the $\overline{\mathrm{MS}}$ scheme and for $N_f = 3$ flavours, $\Lambda_{\mathrm{QCD}} = 332(14)\,\mathrm{MeV}$ was determined from lattice QCD simulations [63] and corresponds to the value of a_s given above. In this low-energy regime of QCD close to $\mu \gtrsim \Lambda_{\mathrm{QCD}}$, perturbation

theory breaks down. Thus, non-perturbative methods are required, e.g. lattice QCD that will be introduced in sec. 2.2.

Asymptotic freedom is a direct consequence of the self-interaction of the gluons, i.e., that the gluon vertex receives gluonic loop corrections and these carry the colour charge themselves, leading to an anti-screening effect that is opposite to the charge screening of, e.g., QED. In fact, it can be shown that only non-Abelian gauge theories show this property of self-interaction of the gauge bosons and thus asymptotic freedom [66, 67]

2.1.4 Confinement

One of the prime examples of non-perturbative effects is string breaking. The force between two static quarks rises linearly with their distance, and is mediated by gluons. This phenomenon has been described in terms of (colour) flux tubes [68]. When the static energy is large enough the string "breaks" and a new quark-antiquark pair is created, forming another string that again is "colourless" — the anti-quark compensates the colour of the quark — and no free colour charges can be observed. This feature has been confirmed numerically, for example in [69].

From this example another feature of QCD becomes evident: No free quarks are observed in nature below the deconfinement temperature T_c. Since quarks carry colour charge, this follows from the postulate that no coloured (non-singlet) states exist and all hadrons are colourless:

$$\begin{array}{llrcl} |q\overline{q}\rangle & \text{(mesons):} & 3 \otimes \overline{3} & = & 1 \oplus 8, \\ |qqq\rangle & \text{(baryons):} & 3 \otimes 3 \otimes 3 & = & 1 \oplus 8 \oplus 8 \oplus 10, \end{array} \tag{2.21}$$

where the numbers on the right-hand side are the dimensions of the irreducible representations of the SU(3) product groups on the left-hand side. Some combinations of quarks and antiquarks like $|qq\rangle$, $|qqqq\rangle$, etc. do not possess singlet representations and have not been found experimentally, whereas for example tetra- ($|q\overline{q}q\overline{q}\rangle$) and pentaquarks ($|qqqq\overline{q}\rangle$) (that are close to meson-meson and meson-baryon states) have recently been confirmed experimentally [70, 71].

It is also possible to form colourless states purely from gluons through their self-interaction. Such bound states are called glueballs and predicted by the standard model [72–74], although their experimental detection is difficult [75]. Only very recently, a three-gluon state has been observed with more than 5 σ confidence [76].

2.1.5 Chiral symmetry and axial anomaly

The low energy regime of QCD can be effectively described by neglecting the heavier quarks, keeping only the lightest degrees of freedom that determine the vacuum structure of QCD. To this end it is helpful to study the massless (classical) Lagrangian of QCD that shows an extended global (space-time independent) symmetry of

$$\mathrm{SU}(N_f)_L \times \mathrm{SU}(N_f)_R \times \mathrm{U}(1)_V \times \mathrm{U}(1)_A, \tag{2.22}$$

where the subscripts denote the conserved currents, namely left, right, vector and axialvector currents that are described in the following.

To discuss the chiral properties of the classical Lagrangian, it is useful to define projectors

$$\mathcal{P}_L = \frac{1-\gamma_5}{2}, \quad \mathcal{P}_R = \frac{1+\gamma_5}{2}, \tag{2.23}$$

that have the expected properties $\mathcal{P}_{L/R}^2 = \mathcal{P}_{L/R}$ and $\mathcal{P}_L\mathcal{P}_R = \mathcal{P}_R\mathcal{P}_L = 0$. Their action on spinors define the left and right-handed fermion fields

$$\psi_L \coloneqq \mathcal{P}_L\psi, \quad \psi_R \coloneqq \mathcal{P}_R\psi. \tag{2.24}$$

Using $\psi = \psi_L + \psi_R$, the Dirac equation (2.14) can be written as

$$\mathcal{L}_f = i\overline{\psi}_{f,L}\not{D}\psi_{f,L} + i\overline{\psi}_{f,R}\not{D}\psi_{f,R} - m_f(\overline{\psi}_{f,L}\psi_{f,R} + \overline{\psi}_{f,R}\psi_{f,L}), \tag{2.25}$$

which shows that for $m_f = 0$ the left- and right-handed components decouple, while for $m_f \neq 0$ there exist terms that mix the two components of the spinors.

In this limit of vanishing quark masses, the $\mathrm{SU}(3)_L \times \mathrm{SU}(3)_R$ part of the symmetry, eq. (2.22), describes the invariance of the Lagrangian under flavour rotations

$$\psi \to \exp\left(i\theta^a_{L/R}t^a\right)\mathcal{P}_{L/R}\psi, \tag{2.26}$$

where again $t^a (a = 1, \ldots, N_f^2 - 1)$ are the generators of $\mathrm{SU}(N_f)$. The conserved currents are

$$L^a_\mu = \frac{1}{2}\overline{\psi}\gamma_\mu t^a \mathcal{P}_L\psi = \frac{1}{2}\overline{\psi}_L\gamma_\mu t^a\psi_L, \quad R^a_\mu = \frac{1}{2}\overline{\psi}\gamma_\mu t^a\mathcal{P}_R\psi = \frac{1}{2}\overline{\psi}_R\gamma_\mu t^a\psi_R. \tag{2.27}$$

Their linear combinations are the non-singlet vector and axialvector currents,

$$V^a_\mu = L^a_\mu + R^a_\mu = \overline{\psi}\gamma_\mu t^a\psi, \quad A^a_\mu = R^a_\mu - L^a_\mu = \overline{\psi}\gamma_\mu\gamma_5 t^a\psi. \tag{2.28}$$

On the classical level, all these four currents are conserved and thus their divergences vanish. In the QCD vacuum, however, only the vector component seems to be conserved and the symmetry is spontaneously broken down to $\mathrm{SU}(N_f)_V$, giving rise to a non-vanishing chiral condensate $\langle\overline{\psi}\psi\rangle = \langle\overline{\psi}_L\psi_R + \overline{\psi}_R\psi_L\rangle \neq 0$ at low temperatures which may be used as an order parameter for the spontaneous symmetry breaking. While $\partial_\mu V_\mu = 0$, the divergence of the axialvector current reads

$$\partial_\mu A^a_\mu = \overline{\psi}\gamma_5\{M, t^a\}\psi, \tag{2.29}$$

where $M = \mathrm{diag}(m_u, m_d, m_s, \ldots)$ is the quark mass matrix. The pseudo-Goldstone bosons of this broken continuous symmetry are the eight light pseudoscalar mesons, i.e., the pions and kaons and the theoretical η_8. Implications of this on the quark model will be discussed in the next section. The fact that the divergences of the axialvector currents do not vanish also gives rise to pseudoscalar decays, i.e., a pseudoscalar meson $\mathcal{M}$ with the corresponding quantum numbers can be annihilated by the axialvector current and has a non-vanishing matrix element

$$\langle\Omega|A^a_\mu|\mathcal{M}(p)\rangle = ip_\mu F^a_\mathcal{M}, \tag{2.30}$$

which defines its decay constant $F^a_\mathcal{M}$. We denote the QCD vacuum state as $|\Omega\rangle$.

Similarly to the $\mathrm{SU}(N_f)_L \times \mathrm{SU}(N_f)_R$ subgroup of eq. (2.22), one can study the U(1) subgroups: Its vector part describes simple phase rotations of the form

$$\psi \to \exp\left(i\alpha_V\right)\psi, \tag{2.31}$$

and its corresponding Noether current is the vector current, $V^0_\mu = \overline{\psi}\gamma_\mu\psi$, which reflects baryon number conservation. Since the group is abelian, it commutes with the mass terms of eq. (2.14) and holds exactly even at finite quark masses.

The axialvector transformations rotate the two chiralities in opposite directions

$$\psi \to \exp\left(i\alpha_A\gamma_5\right)\psi = \exp\left(-i\alpha_A\right)\psi_L + \exp\left(i\alpha_A\right)\psi_R, \tag{2.32}$$

leading to a conservation at the classical level of only the derivative terms of equal chirality in eq. (2.25). The related current is the singlet axialvector current

$$A^0_\mu = \overline{\psi}\gamma_\mu\gamma_5\psi. \tag{2.33}$$

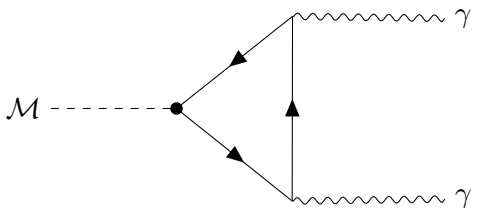


FIGURE 2.1: Feynman diagram relevant for the decay of a singlet pseudoscalar meson $\mathcal{M}$ (dashed line) into two photons (curly lines), $\mathcal{M} \to 2\gamma$ via a quark "triangle" (solid lines). The dot represents the axialvector current while the normal vertices are associated with vector currents. This so-called triangle diagram gives rise to the axial anomaly, since naively $\mathrm{tr}\,\gamma_5\gamma_\mu\gamma_\nu\gamma_\rho = 0$ and the non-vanishing result is due to the regularization prescription. To obtain the full result also the corresponding diagram with the external photons interchanged must be considered. Note that this diagram is only possible when the quarks within the meson interpolating function are of the same flavour, i.e., for a flavour diagonal meson $\mathcal{M} \in \{\pi^0, \eta, \eta'\}$.

This symmetry is broken in two ways: First, at non-vanishing masses it is broken explicitly. Second, it is destroyed by the famous chiral anomaly [77] that is often referred to as triangle or Adler-Bell-Jackiw anomaly [8, 9, 78]. The term "anomaly" means that it is a symmetry of the Lagrangian that is broken by quantization. As a result of the explicit and anomalous breaking of the $\mathrm{U}(1)_A$ symmetry, the divergence of the respective Noether current is not zero and becomes

$$\partial_\mu A^0_\mu = \frac{1}{\sqrt{2N_f}}\overline{\psi}\gamma_5 M\psi + \sqrt{2N_f}\omega, \tag{2.34}$$

where the first term is due to the explicit breaking by the masses, analogously to eq. (2.29). The second term arises from the anomaly and contains the topological susceptibility

$$\omega(x) = \frac{g_s^2}{32\pi^2}F^a_{\mu\nu}(x)\widetilde{F}^a_{\mu\nu}(x). \tag{2.35}$$

This can be derived from evaluating the so-called triangle diagram, cf. fig. 2.1 and paying attention to surface terms when shifting integration variables or, more elegantly, by Schwinger's point-split method, separating the fermion fields of the axialvector bilinear by an asymptotically small distance to regulate the divergence and ensuring gauge invariance, see [56, 79] and references therein for pedagogical derivations. Eq. (2.34) makes the connection to topology obvious: The contribution to eq. (2.34) in the chiral limit coming from the left- and right-handed fermion fields can be computed separately,

$$\partial_\mu A^0_\mu|_{M=0} = \partial_\mu R_\mu|_{M=0} - \partial_\mu L_\mu|_{M=0} = \sqrt{2N_f}(\omega_R - \omega_L), \tag{2.36}$$

and correspond to the number of left- or right-handed zero modes of the Dirac operator at a given field configuration,

$$n_{L/R} = \int \mathrm{d}^4x\omega_{L/R}(x). \tag{2.37}$$

The total topological charge is given by the integral over the anomaly equation (2.36),

$$Q_{\mathrm{top}} = n_R - n_L, \tag{2.38}$$

which is integer valued when the domain of integration is such that the integrand vanishes at the boundaries. The value of the topological charge labels the different topological sectors of the QCD vacuum, i.e., each has the same ground state energy.

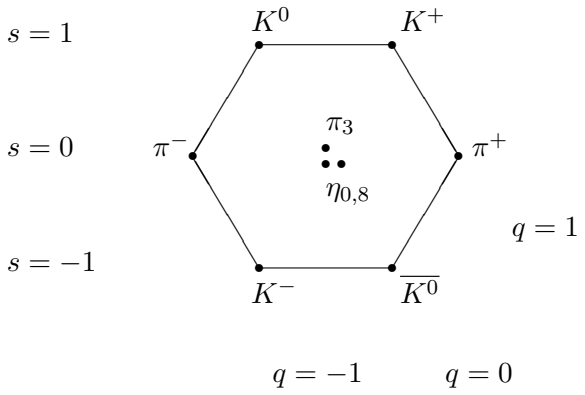


FIGURE 2.2: Light pseudoscalar mesons ($J^{PC} = 0^{-(+)}$), consisting of only light and strange quarks as predicted by the "Eightfold Way" and the quark model. Strangeness increases from bottom ($s = -1$) to top ($s = 1$) while the electric charge is constant along diagonal lines from top-left to bottom-right and increases from bottom-left ($q = -1$) to top-right ($q = 1$). The π_3, η_8 and the η_0 are (predicted) flavour diagonal mesons and the quark bilinears consist of a quark and an anti-quark of the same flavour. Consequently, these carry neither strangeness nor electric charge. The η^0 is the singlet and obtains parts of its mass from the axial anomaly.

By defining $t^0 = \frac{1}{\sqrt{2N_f}}\mathbb{1}$, the axial Ward identities (AWIs) eqs. (2.34) and (2.29) can be combined and written in a compact way,

$$\partial_\mu A^a_\mu = \overline{\psi}\gamma_5\{M, t^a\}\psi + \sqrt{2N_f}\delta^{a0}\omega, \quad a = 0, \ldots, N_f^2 - 1, \tag{2.39}$$

and we will refer to this equation also by the name of the partially conserved axial current (PCAC) relation.

2.1.6 Quark model and the η/η'

Before the Lagrangian of QCD itself and in particular the idea of colour gauge symmetry was established, the "zoo" of hadrons was classified in terms of effective quantum numbers, describing the degrees of freedom of the theory that are consequences of the symmetries discussed in the previous section. A successful approach and important milestone was the development of the "Eightfold Way" by Gell-Mann [80]. Hadronic bound states can be classified in terms of flavour quantum numbers like isospin (I), strangeness (S), etc. for the heavier quarks and quantum numbers that label total spin (J), parity (P) and charge conjugation (C), usually written as J^{PC}.

For the pseudoscalar mesons considered in this work we have $J^{PC} = 0^{-+}$. Flavour symmetry predicts eight non-singlet and one singlet particle, since for mesons that consist of a light ($q \in \{u, d, s\}$) quark and anti-quark, $3 \otimes \overline{3} = 8 \oplus 1$. These eight non-singlet pseudoscalar mesons are the three pions π_0, π^+, π^-, four kaons, $K^0, \overline{K}^0, K^+$ and K^- and one predicted octet particle that is called η_8. These can be organized in terms of their "leading" flavour content by pseudoscalar operators whose hermitian conjugate can create such states from the vacuum,

$$P^a = \overline{\psi}\gamma_5 t^a \psi, \tag{2.40}$$

using $t^0 = \frac{1}{\sqrt{2N_f}}\mathbb{1}$, $a = 0, \ldots, 8$ and $\overline{\psi} = (\bar{u}, \bar{d}, \bar{s})$. The naming convention is

$$\begin{aligned}
\pi^+ &\sim (P^1 + iP^2)^\dagger = u\gamma_5\bar{d}, & \pi^- &\sim (P^1 - iP^2)^\dagger = d\gamma_5\bar{u},\\
\pi_3 &\sim P^{3\dagger} = \tfrac{1}{2}\left(u\gamma_5\bar{u} - d\gamma_5\bar{d}\right), & &\\
K^+ &\sim (P^4 + iP^5)^\dagger = u\gamma_5\bar{s}, & K^- &\sim (P^4 - iP^5)^\dagger = s\gamma_5\bar{u},\\
K^0 &\sim (P^6 + iP^7)^\dagger = d\gamma_5\bar{s}, & \overline{K}^0 &\sim (P^6 - iP^7)^\dagger = s\gamma_5\bar{d},\\
\eta_8 &\sim P^{8\dagger} = \tfrac{1}{\sqrt{12}}\left(u\gamma_5\bar{u} + d\gamma_5\bar{d} - 2s\gamma_5\bar{s}\right), & &\\
\eta_0 &\sim P^{0\dagger} = \tfrac{1}{\sqrt{6}}\left(u\gamma_5\bar{u} + d\gamma_5\bar{d} + s\gamma_5\bar{s}\right). & &
\end{aligned} \tag{2.41}$$

These nine mesons are listed and arranged by their quantum numbers in fig. 2.2. The eight non-singlet mesons are the Goldstone-bosons of the breaking of $\mathrm{SU}(3)_L \times \mathrm{SU}(3)_R \to \mathrm{SU}(3)_V$ discussed in the previous section and are indeed relatively light compared to other (non-pseudoscalar) mesons with the same quark content. The unflavoured vector mesons, called the ρ mesons, have masses around $M_\rho \approx 770\,\mathrm{MeV}$ which is large compared to the pseudoscalar pions, $M_{\pi^\pm} \approx 139\,\mathrm{MeV}$. In the limit of vanishing quark masses eq. (2.30) implies that all non-singlet mesons ($a = 1 \ldots 8$) become massless:

$$p_\mu \langle \Omega | A^a_\mu | \mathcal{M}(p)\rangle = i p^2(p) F^a_{\mathcal{M}} \xrightarrow{m_f \to 0} 0, \tag{2.42}$$

due to $\partial_\mu A^a_\mu = 0$ in that case. Since $F^a_{\mathcal{M}} \neq 0$, it follows that $p^2 = 0$ and the mass in the chiral limit $M_{\mathcal{M}} = 0$ as is expected for the Goldstone bosons. We denote the state of a pseudoscalar meson $\mathcal{M}$ with four-momentum p as $|\mathcal{M}(p)\rangle$ and imply that the current is chosen such that the matrix element does not vanish at finite quark masses (e.g., $a = 3$ for $\mathcal{M} = \pi^0$, etc.).

The situation is different for the (hypothetical) singlet that arises from the breaking of the axial $\mathrm{U}(1)_A$ symmetry and for which a quark mass independent contribution persists in the chiral limit, cf. eq. (2.39),

$$p_\mu \langle \Omega | A^0_\mu | \eta_0 \rangle = i p^2 F^0_{\eta_0} \xrightarrow{m_f \to 0} \sqrt{6}\langle \Omega | \omega | \eta_0 \rangle, \tag{2.43}$$

where we identify the rest mass of the singlet $M_0^2 = p_0^2$ in the chiral limit. For this reason, the chiral anomaly increases the mass of the singlet η_0 compared to the other non-singlet pseudoscalar mesons.

Therefore, in the $N_f = 3$ chiral limit — or the symmetric limit, where all quark masses are equal and $M = m\mathbb{1}$ commutes with any other matrix — the π^0 is generated solely by the $a = 3$ pseudoscalar current, $|\pi^0\rangle = P^{3\dagger}|\Omega\rangle = \psi\gamma_5 t^3 \overline{\psi}|\Omega\rangle$, the $|\eta\rangle = |\eta_8\rangle = P^{8\dagger}|\Omega\rangle$ is a pure octet and the $|\eta'\rangle = |\eta_0\rangle = P^{0\dagger}|\Omega\rangle$ is a pure singlet state. Elsewhere in the quark mass plane, the remaining $\mathrm{SU}(3)$ flavour symmetry is explicitly broken, since the strange quark is much heavier than the light (up and down) quarks and the following complication arises: The π_3, η_8 and η_0 all have the same flavour quantum numbers and are linear combinations of the quark model predictions

$$\begin{pmatrix} \pi^0 \\ \eta \\ \eta' \end{pmatrix} = \begin{pmatrix} R(\theta_\pi) & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & R(\theta_P) \end{pmatrix} \begin{pmatrix} \pi_3 \\ \eta_8 \\ \eta_0 \end{pmatrix}, \tag{2.44}$$

where $\theta_\pi = \mathcal{O}(m_d^2 - m_u^2)$ and $\theta_P = \mathcal{O}(2m_s^2 - m_u^2 - m_d^2)$ are the angles in the 2×2 orthogonal transformation

$$R(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}. \tag{2.45}$$

Since the isospin breaking effects $\mathcal{O}(m_d^2 - m_u^2)$ are small relative to the strange quark mass, the neutral pion receives almost no contributions from the octet and singlet currents which involve

strange quarks and $\pi^0 \approx \pi_3$ to a very high precision. For the η and η', however, SU(3) breaking is more significant and this has the consequence that the physical (mass eigen-) states observed in nature are far away from the expected flavour states. Hence, the simple quark model fails to predict their masses, i.e., $\theta_P \neq 0$.

Historically, this became known as "mixing" of the η and η' but it should not be confused with state mixing as, for example takes place between the K^0 and the $\overline{K}^0$ which both are physical states (with different flavour quantum numbers) and can oscillate between these by physical (weak) processes. In contrast, the physical particles here are distinct η and η' states and both contain singlet (η_0) and octet (η_8) contributions, if $m_s \neq m_\ell$. In this thesis, we frequently adopt this jargon of mixing, but it should be understood as described here.

There is a further limitation of the quark model: It does not predict gluonic and exotic (tetra- and pentaquark) states and contributions from heavier quarks. Glueballs — gluonic bound states — can have the same quantum numbers and, again, contribute to the physical states. Also, heavier quarks contribute to the physical η and η' states, but their effect is small. The argument is the same as why the neutral pion does not "see" the strange quark: the mass of the charm is much heavier than the valence light quarks.

2.2 Discretization

As discussed in sec. 2.1.3, the physics in a low energy regime of QCD cannot be studied using perturbation theory, i.e., by expanding in the coupling constant. This prevents the study of, e.g., vacuum effects and hadron spectroscopy. Among the non-perturbative approaches to QCD, the lattice formulation, based on early ideas by Kenneth Wilson [81], is the most popular as it enables ab initio calculations of QCD observables starting from the Lagrangian. To this end, the fields are put on a lattice — fermion fields on the sites and the gluonic link fields on the connections between the sites — with finite lattice spacing and extent. This serves as a UV-regulator of the theory and enables computations by using only a finite number of space-time points. The regulator must be removed in the end by taking the continuum limit and renormalization. This topic will be discussed in sec. 2.4.

In this section the main concepts of lattice QCD are introduced, starting from the path integral formulation and possible discretizations of the gauge fields and fermions. Although the focus is put on the actions used for the generation of the ensembles in this work, we mainly discuss the bare fundamentals that can also be found in text books, e.g. [57–59]. The discussion of the unique features of the simulation efforts undertaken in the coordinated lattice simulation (CLS) initiative and a presentation of the ensembles is deferred to ch. 3.

2.2.1 Path integral formulation of QCD

The path integral formalism [82, 83] provides a prescription to quantize a classical Lagrangian $\mathcal{L}$, and allows the calculation of expectation values $\langle O \rangle$ of arbitrary (usually non-local), gauge invariant operators O that depend on general fields $\phi, \dots$ (for the moment we ignore complications that arise from fermion and non-abelian gauge fields),

$$\langle O \rangle = \frac{1}{Z} \int \mathrm{D}\phi \dots \, O(\phi, \dots) \, \exp\left(-S[\phi, \dots]\right), \tag{2.46}$$

with the partition function

$$Z = \int \mathrm{D}\phi \dots \, \exp\left(-S[\phi, \dots]\right), \tag{2.47}$$

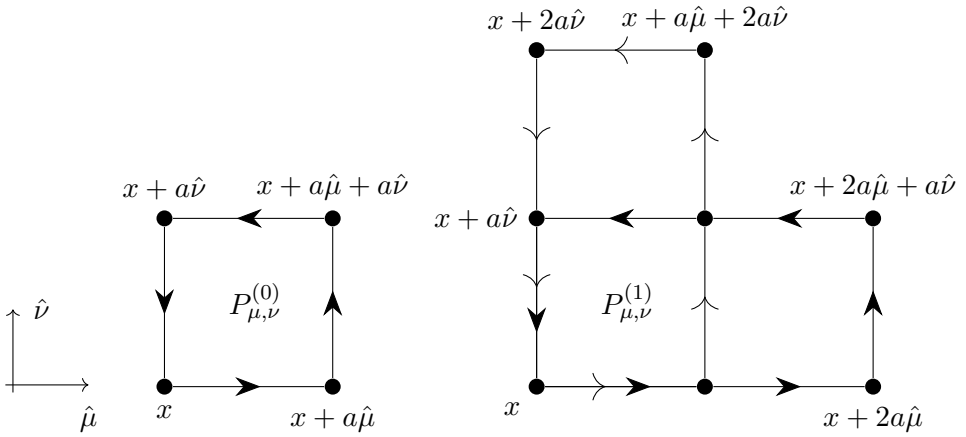


FIGURE 2.3: Gauge loops in the $\hat{\mu} - \hat{\nu}$-plane that contribute to the Wilson ($P^{(0)}_{\mu,\nu}$) and Lüscher-Weisz ($P^{(0)}_{\mu,\nu}$ and $P^{(1)}_{\mu,\nu}$) gauge actions. The two rectangular loops with different arrow shapes on the right-hand side are summed over and both contribute to $P^{(1)}_{\mu,\nu}$, see eq. (2.57).

where we have already performed the Wick rotation to imaginary times $x_0 \mapsto -ix_0$ to render the integrand real and finite[2]. Integration over group valued fields is according to the corresponding Haar measure and over all possible field configurations, weighted by the corresponding Boltzmann factor $\exp(-S[\phi, \dots])$. The direct evaluation of this integral is impossible and can only be estimated stochastically by using Monte-Carlo methods, which will be discussed below.

The action S is a suitable discretization of the QCD action in Euclidean space-time so that all coordinates have been replaced by integer multiples of the lattice spacing a,

$$x_\mu = an_\mu, \qquad 0 \leq n_\mu < N_\mu = L_\mu/a, \quad n_\mu \in \mathbb{Z}, \tag{2.48}$$

where L_μ is the lattice extent in direction $\hat{\mu}$ and N_μ the respective number of sites. We frequently refer to the time direction with $t = x_0$. In order to work with a finite number of lattice points, boundary conditions on the fields are put in place. In this work, we will freely replace an_μ by x_μ, even when the discretized space-time coordinates are meant.

The computation of the lattice action then amounts to summing over the lattice and recovers the (Euclidean) continuum action S^{cont} in the continuum limit, $a \to 0$,

$$S[\phi(n), \dots] = a^{-4} \sum_n \mathcal{L}[\phi(a, n), \dots] \to S^{\text{cont}} = \int \mathrm{d}^4 x \mathcal{L}^{\text{cont}}[\phi(x), \dots]. \tag{2.49}$$

Just like the Lagrangian, eq. (2.1), the action can be split into a gluonic and fermionic parts for each flavour,

$$S[U, \{\psi_f\}] = S_G[U] + \sum_f S_f[U, \psi_f]. \tag{2.50}$$

The choices for S_G and S_f are not unique and different choices may differ by their discretization and how they approach the continuum limit. In the following suitable choices for each of the terms will be discussed.

2.2.2 Gauge invariance, link fields and gauge action

We now turn to the concrete form of lattice actions, starting with the gauge part. Clearly, any action must obey gauge invariance and the straight-forward way is to discretize the (continuum) gauge transporters, eq. (2.8),

$$G(x,y) = \mathcal{P}\exp\left(i\int_{\mathcal{C}(x,y)} A_\mu \mathrm{d}s_\mu\right), \tag{2.51}$$

where $\mathcal{P}$ is the path ordering operator, the (bare) coupling g_0 has now been absorbed into the definition of the gluon field, $A_\mu \mapsto \frac{1}{g_0}A_\mu$, and $\mathcal{C}(x,y)$ are paths connecting space-time points x and y. Any closed loop is invariant under gauge transformations Ω due to the unitarity and the behaviour of the gauge transporters under gauge transformations, eq. (2.5).

On the lattice, this can be achieved by introducing oriented variables, linking two sites x and $x + a\hat{\mu}$,

$$U_\mu(x) = \exp(iaA_\mu(x + a\hat{\mu}/2)) \in \mathrm{SU}(3), \tag{2.52}$$

where the path ordering is irrelevant for $a \to 0$ and $A_\mu(x + a\hat{\mu}/2)$ is the discretized and rescaled version of the gluon field. Just as its continuum counterpart, eq. (2.8), the gauge field $U_\mu(x)$ is $\mathrm{SU}(3)$ group-valued and it "lives" on the links between every lattice point, i.e., the link field consists of $N_d \times N_c^2 \times \prod N_\mu$ complex numbers, where $N_d = 4$ is the dimension of space-time.

The simplest closed loop — going one hop forward and then back — is

$$U_\mu(x)U_{-\mu}(x + a\hat{\mu}) = \bullet\!\!\rightarrow\!\!\leftarrow\!\!\bullet = \mathbb{1} \tag{2.53}$$

and thus

$$U_{-\mu}(x + a\hat{\mu}) = U_\mu^{-1}(x) = U_\mu^\dagger(x). \tag{2.54}$$

The first non-trivially closed path is the plaquette

$$P^{(0)}_{\mu,\nu}(x) = U_\mu(x)U_\nu(x + a\hat{\mu})U_\mu^\dagger(x + a\hat{\nu})U_\nu^\dagger(x), \tag{2.55}$$

depicted in the left panel of fig. 2.3.

This enables us to write down the simplest discretization of the Yang-Mills action (assuming periodic boundary conditions), the Wilson gauge action

$$S_{G,W}[U] = \frac{\beta}{6}\sum_x \sum_{\mu<\nu} \mathrm{Re}\,\mathrm{tr}\left(\mathbb{1} - P^{(0)}_{\mu\nu}(x)\right), \tag{2.56}$$

where $\beta = 6/g_0^2$.

It can be shown that this definition is already free of $\mathcal{O}(a)$ effects and only higher order deviations need to be removed when taking the continuum limit. Discretization effects can be reduced [84] by including planar loops that extend over two lattice spacings

$$\begin{aligned} P^{(1)}_{\mu,\nu}(x) = {} & U_\mu(x)U_\nu(x + a\hat{\mu})U_\nu(x + a\hat{\nu} + a\hat{\mu})U_\mu^\dagger(x + 2a\hat{\nu})U_\nu^\dagger(x + a\hat{\nu})U_\nu^\dagger(x) \\ & + U_\mu(x)U_\mu(x + a\hat{\mu})U_\nu(x + 2a\hat{\mu})U_\mu^\dagger(x + a\hat{\mu} + a\hat{\nu})U_\mu^\dagger(x + a\hat{\nu})U_\nu^\dagger(x). \end{aligned} \tag{2.57}$$

[2] When studying physics at, e.g., non-zero chemical potential or theta term, the action after Wick rotation to Euclidean times acquires an imaginary contribution and gives rise to the sign problem, making it hard to integrate numerically using Monte-Carlo methods due to the oscillatory integrand. We will not be concerned with these simulations in this thesis.

The inclusion of these rectangles in the action is known as Lüscher-Weisz action

$$S_{G,LW}[U] = \frac{\beta}{6}\sum_x \sum_{\mu<\nu} \left[c_0 \,\mathrm{tr}\left(\mathbb{1} - P^{(0)}_{\mu\nu}(x)\right) + c_1 \,\mathrm{tr}\left(\mathbb{1} - P^{(1)}_{\mu\nu}(x)\right)\right], \tag{2.58}$$

where c_i are improvement parameters that need to be normalized such that $c_0 + 8c_1 = 1$ and $c_0 > 0$. The case $c_0 = 1$ and $c_1 = 0$ corresponds to the Wilson plaquette action $S_{G,W}$. The perturbatively determined tree-level values are $c_0 = \frac{5}{3}$ and $c_1 = -\frac{1}{12}$ [85]. This removes the bulk of $\mathcal{O}(a^2)$ effects but for further improvement a non-perturbative determination is needed. There exists a choice based on a renormalization group analysis [86]. However, any modification of the gluonic action will also change the renormalization of fermionic observables as it changes the gluon propagator and, hence, to avoid this dependency, we stick with the more generally applicable tree-level coefficients.

2.2.3 Fermion discretization

The action for a fermion of flavour f within eq. (2.50) reads

$$S_f[U,\psi] = a^4 \sum_x \overline{\psi}_f(x) D_f \psi_f(x), \tag{2.59}$$

with a suitable discretization of the massive Dirac operator D_f that acts on fermion fields ψ_f. Other than the gauge links, fermion spinors are positioned on the sites of the lattice and have colour and Dirac components. Their field on a finite lattice consists of $N_c \times 4 \times \prod N_\mu$ complex numbers. They are Grassmann valued to ensure Fermi-Dirac statistics, i.e., they anticommute

$$\left\{\psi, \psi'\right\} = \left\{\overline{\psi}, \overline{\psi}'\right\} = \left\{\overline{\psi}, \psi'\right\} = 0, \tag{2.60}$$

where ψ and ψ' are arbitrary fermion spinors.

One possible discretization of the Dirac operator and a particularly popular choice is given by the Wilson formulation[81] with an additional Sheikholeslami-Wohlert term [87] for $\mathcal{O}(a)$ improvement,

$$D_f = \frac{1}{2}\gamma_\mu(\nabla^*_\mu + \nabla_\mu) - \frac{a}{2}\nabla^*_\mu\nabla_\mu + \frac{a}{4}c_{\mathrm{SW}}\sigma_{\mu\nu}\widehat{F}_{\mu\nu} + m_{0,f}, \tag{2.61}$$

where $\sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu]$. The forward and backward linear approximations ∇_μ and ∇^*_μ of the covariant derivative, eq. (2.10), are defined in terms of their actions on fermion spinors

$$\begin{aligned}\nabla_\mu\psi(x) &= \frac{1}{a}\left(U_\mu(x)\psi(x+a\hat{\mu}) - \psi(x)\right),\\ \nabla^*_\mu\psi(x) &= \frac{1}{a}\left(\psi(x) - U^\dagger_\mu(x-\hat{\mu})\psi(x-a\hat{\mu})\right).\end{aligned} \tag{2.62}$$

The first sum in eq. (2.61) and the mass term corresponds to the straight-forward discretization of the continuum Dirac operator. This naive discretization alone would, however, generate fifteen fermion doublers as can be shown by looking at the degeneracy of the zero modes. To cure this problem the first term in the second sum is introduced that shifts the doublers away — making them infinitely heavy — when approaching the continuum limit. This, however, breaks chiral symmetry explicitly and results in an additive renormalization of the quark masses. At the same time, this introduces $\mathcal{O}(a)$ discretization artefacts that can be removed by tuning the improvement coefficient c_{SW} that multiplies the last term of eq. (2.61). To leading order in perturbation theory we have $c_{\mathrm{SW}} = 1$, but higher orders depend on the gauge action and coupling. A non-perturbative determination for the tree-level improved Lüscher-Weisz gauge action is available in [88].

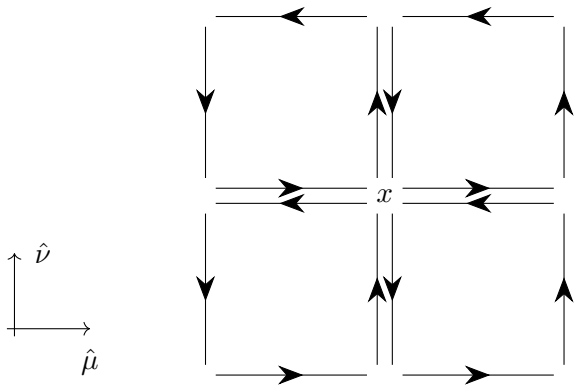


FIGURE 2.4: "Clover leaf" products that add as individual terms to $Q_{\mu\nu}$. Sites are not drawn, and each link extends one lattice unit.

The construction of the c_{SW} term follows the Symanzik improvement description and is obtained by including the lowest contributing $\mathcal{O}(a)$ terms. This amounts to a symmetrized "cloverleaf" shaped sum of gauge loops,

$$\widehat{F}_{\mu\nu}(x) = \frac{-i}{8a^2}\left(Q_{\mu\nu}(x) - Q_{\nu\mu}(x)\right), \tag{2.63}$$

where the leafs are symbolized by the loops in fig. 2.4 and can be written in terms of plaquette sums, see eq. (2.55),

$$Q_{\mu\nu}(x) = P^{(0)}_{\mu,\nu}(x) + P^{(0)}_{-\mu,\nu}(x) + P^{(0)}_{-\mu,-\nu} + P^{(0)}_{\mu,-\nu}. \tag{2.64}$$

It should be noted that the Dirac operator depends on the flavour only through the mass term. The Dirac operator (setting $c_{\mathrm{SW}} = 0$ for the moment) can be rewritten using a flavour independent hopping term

$$H(x)\psi(x) = \sum_{\mu}\left[(\mathbb{1}-\gamma_\mu)U_\mu(x)\psi(x+a\hat{\mu}) + (\mathbb{1}+\gamma_\mu)U^\dagger_\mu(x-a\hat{\mu})\psi(x-a\hat{\mu})\right], \tag{2.65}$$

that collects all terms of the Dirac operator that are one "hop" away from x. The mass dependence can then be absorbed in constants $\alpha_f = m_{0,f} + 4/a$ and $\kappa_f = (2am_{0,f}+8)^{-1}$. With these abbreviations the massive Dirac operator of flavour f reads

$$D_f = \alpha_f(\mathbb{1} - \kappa_f H). \tag{2.66}$$

The global prefactor α_f can be absorbed in the definition of the fermion fields of that particular flavour f. Then the hopping parameter κ_f is the only mass-dependent parameter and used to tune the quark mass. The form of the Dirac operator in eq. (2.66) can be used to expand around small hopping parameters (corresponding to large quark masses) and will be of use later on.

Due to the explicit breaking of chiral symmetry, the quark masses defined, for example, through the PCAC relation, eq. (2.39), do not vanish when $m_{0,f} = 0$ but at the critical quark mass $m_{0,f} = m_{\mathrm{cr},f}$. The value of the critical mass is a priori unknown and depends on the lattice spacing. It has been determined for our action and range of lattice spacings in [89, 90].

These complications would be absent in fermion discretizations that preserve chiral symmetry. However, it has been shown by Nielsen and Ninomiya [91] that there is in fact no fermion discretization in four dimensions that preserves chiral symmetry and is free of doublers, i.e., there is no discretized Dirac operator that obeys $\gamma_5 D + D\gamma_5 = 0$ exactly at a finite lattice spacing. This

turned out to be the lattice equivalent to the Adler–Bell–Jackiw anomaly, i.e., a fermion discretization that has these desired properties would not reproduce the anomaly in the continuum; the anomaly would exactly be cancelled by the doublers [92]. However, there exist other fermion formulations that are chiral in the sense that they obey the Ginsparg-Wilson equation [93],

$$D_f\gamma_5 + \gamma_5 D_f = aD_f\gamma_5 D_f \tag{2.67}$$

and thus allow to assess chiral properties even at finite lattice spacing. This is particularly important when studying, e.g., spontaneous chiral symmetry breaking. Prominent discretizations are those by Neuberger [94] and domain wall fermions[95].

All of these discretizations, including the Wilson one discussed above, exhibit another important symmetry that can easily be shown from, e.g., eq. (2.61), and that is γ_5-hermiticity,

$$D_f^\dagger = \gamma_5 D_f \gamma_5. \tag{2.68}$$

In consequence, the Dirac operator has only complex conjugate pairs of eigenvalues and the quark determinant is real, which is an important property for the simulation of dynamical QCD. Furthermore, this property allows for the cheap and efficient computation of point-to-all quark propagators, inverting the Dirac operator on an (extended) source as will be discussed in sec. 2.3.

2.2.4 Dynamical simulations with three quarks

Having suitable discretized actions at hand the actual simulations need to be performed. In this section we briefly sketch the main complications that arise from dynamical simulations with three quarks (that is, two degenerate light quarks and one typically heavier strange quark) and the methods that are used to solve these issues. Many of the technical details can be found in [96], but important detail is also given in [97].

The numerical task is to evaluate the integral in eq. (2.46) and the partition function to compute expectation values of observables in the end. This is done by importance sampling, generating gauge configurations according to the Boltzmann weight. In the case of pure gauge theories corresponding to no dynamical — infinitely heavy — quarks, this is relatively simple, because the gauge action is local and updates can be performed more easily, e.g. using the heatbath algorithm [98]. The matter becomes more complicated with fermions. Since these are Grassmann valued objects they cannot be simulated directly but need to be integrated out [99, 100]

$$Z = \int D[U] \prod_f D\left[\psi_f\right] D\left[\overline{\psi}_f\right] \exp\left(-S[U,\{\psi_f\}]\right) = \int D\left[U\right] \prod_f \det\left(D_f\right) \exp\left(-S_G\right), \tag{2.69}$$

where for the moment we assume that all Dirac operators have only eigenvalues with positive real parts and the fermionic parts have become fermion determinants so that only the gauge action remains in the exponent. Naturally, the determinants involve all entries of the Dirac matrices and thus are intrinsically non-local.

Because $\det(D_f)$ is real, the quantity

$$p[U] = \frac{1}{Z}\exp(-S[U,\{\psi_f\}]), \quad \text{where} \quad S[U] = S_G - \sum_f \ln\det\left(D_f[U]\right) \tag{2.70}$$

can be interpreted as a probability density that can be integrated out to obtain expectation values of observables

$$\langle\mathcal{O}\rangle = \int D[U]p[U]\mathcal{O}[U]. \tag{2.71}$$

In practice this integral is approximated by averaging over a large-enough number of representative gauge configurations. The accuracy of this approximation can be estimated by computing statistical errors, where the autocorrelation in Monte-Carlo time among measurements must be taken into account, cf. app. B.

The evaluation of the fermion determinants can be done cost-efficiently by stochastic estimation, resorting to pseudo-fermions ϕ_i that are no longer Grassmann numbers, but Gaussian distributed random variables. The combined fermion determinant for two mass-degenerate quarks can be evaluated using

$$|\det D_f[U]|^2 \propto \int D[\phi] \exp\left[-\phi^\dagger (D_f^\dagger[U] D_f[U])^{-1} \phi\right], \tag{2.72}$$

where $D[\phi] = \prod_{x,a,\alpha} \mathrm{d}\phi_{\alpha,a}(x) \mathrm{d}\phi^\dagger_{\alpha,a}(x)$ and thus it can be approximated by random sampling. This form can also be used to derive fermion forces that are needed in hybrid Monte-Carlo (HMC) simulations [101], where one integrates the molecular-dynamics equations, following the trajectory given by the canonical momentum and field coordinates over some molecular dynamics time τ. At the end of each such trajectory, the proposed field update is accepted or rejected by a Metropolis step, evaluating the true change in action.

Still, the force originating from the fermionic sector requires inversions of the Dirac operator and is more expensive to compute than that from the gauge sector while giving only comparably small corrections; therefore, it is advantageous to simulate the different parts of the action with different time step resolutions [102]. One can go further and split the same flavour into two determinants evaluated with separate pseudofermions [103]

$$\left|\det D_f\right|^2 = \det(D_f^\dagger D_f + \mu^2) \det \frac{D_f^\dagger D_f}{D_f^\dagger D_f + \mu^2}, \tag{2.73}$$

with an artificial "twisted" mass μ. The same relation can be used recursively to factorize the first determinant further, allowing arbitrarily many frequency splittings that need to be evaluated with separate pseudofermions but with decreasing frequency in the trajectory. If chosen appropriately, the first determinant is almost insensitive to the true quark mass, its force dominates in the HMC trajectories and suppresses eigenvalue fluctuations towards very small values [104], thereby avoiding so-called exceptional configurations and stabilizing the simulations. In practice, a slightly different factorization turned out to have improved properties which we do not quote here but refer to the literature [104]. Since the fluctuations of the second determinant are small, it is also possible to evaluate it after the simulation, reweighting the generated ensembles with this factor during analysis, cf. app. B. Using the above factorization in a multiple timestep setup in conjunction with efficient solvers for the Dirac equation that make use of the local coherence [105] of low modes like algebraic multigrid [106, 107] or domain decomposition [108, 109] allow for an efficient sampling of the field space.

The integration of a third non-degenerate flavour raises additional difficulties: Both for the probabilistic interpretation and for the rewriting into pseudofermions it is essential that the fermion determinant is positive. This is not ensured for Wilson fermions at least at small quark masses. The inclusion of the strange quark is built upon the assumption that the spectral gap of the hermitian operator $Q_s = \gamma_5 D_s$ (with real eigenvalues) is large and positive and the physical strange quark mass is large enough to prevent the eigenvalues from "tunnelling" through the spectral gap to negative eigenvalues.

In the evaluation of the fermion determinant, $|Q_s| = \sqrt{Q_s^2} = \sqrt{D_s^\dagger D_s}$ is used, and an approximation is needed for its inversion, i.e., a function $R(Q_s^2)$ that approximates the inverse square

root and $|Q_s|R(Q_s^2) \approx \mathbb{1}$ within some known range and precision. The Zolotarev rational approximation is a popular choice, since it allows splitting up the individual terms,

$$R(X) = A \prod_{i=1}^{n} \frac{X + a_{2i-1}}{X + a_{2i}}, \tag{2.74}$$

where n is the degree of the polynomials in numerator and denominator and a_i and A are fixed (analytically known) parameters with $a_{i+1} < a_i$. Eq. (2.74) can be reexpanded in a series $R(X) = A\left(1 + \sum_{i=1}^{n} r_i (X + a_{2i})^{-1}\right)$ and thus $R(Q_s^2)$ can be computed by solving the set of equations

$$\left(Q_s^2 + \mu_k^2\right) \psi_k = \phi \quad \text{for } k = 1, \ldots, n, \tag{2.75}$$

where μ_k and the exact form and normalization of the source ψ_k depend on the parameters of the approximation. Since the right-hand side is constant, a multishift-solver can be used to solve for all k at once. The strange quark determinant then can be computed using two pseudo-fermions,

$$\det |Q_s| \propto \int \mathrm{D}[\phi_1]\mathrm{D}[\phi_2] \exp\left(-\phi_1^\dagger \left(|Q_s|R(Q_s^2)\right)^{-1} \phi_1 + \phi_2^\dagger R(Q_s^2)\phi_2\right), \tag{2.76}$$

where within the trajectory only the second (ϕ_2) term needs to be evaluated and the first corrects only for the approximation error. This can be done in an accept-reject step or again by means of reweighting the simulated action to the true action. The strange reweighting factor corresponding to the first term is given by

$$w_s = \det\left(Q_s R(Q_s^2)\right) \tag{2.77}$$

and requires stochastic evaluation, see [97] for details. The evaluation in practice assumes a positive spectrum due to the large strange quark mass for efficiency reasons but the reweighting factor is not protected against negative eigenvalue fluctuations of Q_s. In actual simulations configurations with negative reweighting factors have been observed and in this case the reweighting factor needs to be corrected by flipping the sign accordingly [110].

Adding the additional non-degenerate quark in the HMC by using the rational approximation is called RHMC and allows to simulate also single (heavy) quark flavours instead of only mass degenerate pairs of quarks. This enables the efficient simulation of $N_f = 2 + 1$ flavour QCD to generate gauge configurations that can be analysed in many contexts.

2.3 Hadron spectroscopy on the lattice

In this section, we describe the steps necessary to compute hadron masses and local matrix elements on the individual configurations that were generated in the Monte-Carlo process. We emphasize peculiarities of flavour diagonal mesons like the η and η': First, suitable interpolating operators are introduced that can excite (both the) η and η' states, a fact that is often referred to as "mixing" of the η and η'. After Wick contractions, propagators are built that can be used to extract masses and amplitudes. Subsequently, we set up our conventions and definitions for the four independent decay constants of the η/η' system and discuss how these are related to local matrix elements extracted from the propagators.

2.3.1 Interpolating operators

We consider interpolators $\hat{O}_{\vec{p}}$ that destroy hadron states with the desired quantum numbers and spatial momentum $\vec{p}$. Their hermitian conjugates create such states from the vacuum,

$$\hat{O}_{\vec{p}}^\dagger |\Omega\rangle = |\Phi\rangle, \tag{2.78}$$

where $|\Phi\rangle$ are field configurations that generally do not correspond to a particular mass eigenstate but rather to a linear combination of states with the same quantum numbers.

We excite pseudoscalar states with $J^{\mathrm{PC}} = 0^{-+}$ using the corresponding pseudoscalar interpolators,

$$\mathcal{P}^a_{\vec{p}}(t) = a^3 \sum_{\vec{x}} \exp(-i\vec{p}\cdot\vec{x})\mathcal{P}^a(t,\vec{x}), \quad \text{where} \quad \mathcal{P}^a(x) = \sqrt{2}\overline{\widetilde{\psi}}(x)t^a\gamma_5\widetilde{\psi}(x) \tag{2.79}$$

where we keep a notation similar to the continuum eq. (2.40) and $\widetilde{\psi}(x)$ is a usually spatially-extended functional of the local fermion fields, organized in a vector $\psi = (u,d,s)^\top$ and centered around $x = (t,\vec{x})$. The construction of extended operators will be described in the subsequent section.

The quark flavour structure is determined by U(3) flavour irreducible representations, see eq. (2.41) for the flavour structures corresponding to individual pseudoscalar meson states. The flavour diagonal combinations are

$$\mathcal{P}^3 = \frac{1}{\sqrt{2}}\left(\mathcal{P}^u - \mathcal{P}^d\right), \tag{2.80}$$

$$\mathcal{P}^8 = \frac{1}{\sqrt{6}}\left(\mathcal{P}^u + \mathcal{P}^d - 2\mathcal{P}^s\right), \tag{2.81}$$

$$\mathcal{P}^0 = \frac{1}{\sqrt{3}}\left(\mathcal{P}^u + \mathcal{P}^d + \mathcal{P}^s\right). \tag{2.82}$$

where $\mathcal{P}^q = \overline{\widetilde{q}}\gamma_5\widetilde{q}$ and $\widetilde{q} \in \{\widetilde{u},\widetilde{d},\widetilde{s}\}$. Within our simulations, the up and down quarks are set to equal masses $m_u = m_d$ and we refer to them as "light" quarks (compared to the strange quark) and write $\mathcal{P}_\ell = \frac{1}{\sqrt{2}}\left(\mathcal{P}^u + \mathcal{P}^d\right)$. As a result of this approximation of exact isospin symmetry, any correlators that emerge from Wick contractions with the triplet interpolator $\mathcal{P}^3$ will immediately cancel out.

2.3.2 Quark smearing

To increase the overlap of the fermion sources and sinks with the physical states, spatially extended operators are employed, denoted by calligraphic letters, e.g., $\mathcal{P}^a$. These are related to local currents by replacing the local fermion fields ψ with smeared quark fields, $\widetilde{\psi} = \phi^s\psi$, where ϕ is the Wuppertal smearing kernel [111]

$$\phi(x,y) = \frac{1}{1+6\delta}\left(\delta(x,y) + \delta\sum_{j=\pm1}^{\pm3}\widetilde{U}_j(x)\delta(x+a\hat{j},y)\right). \tag{2.83}$$

The width of the smearing is determined by the parameter δ and a common choice is to keep it fixed at $\delta = 0.25$. The smearing radius is then controlled by the number s of applications of the smearing kernel to the fermion vector. An optimal smearing radius ensures dominance of the states of interest and thus the optimal choice of s depends on the lattice spacing, the quark masses and the particle state that should be extracted.

In eq. (2.83), $\widetilde{U}_\mu(x)$ is a spatially APE smeared [112] version of the gauge field

$$\widetilde{U}_\mu(x) = \mathrm{argmax}_{U'_\mu}\,\mathrm{Re}\left(U'_\mu V^\dagger_\mu(x)\right), \tag{2.84}$$

which projects a smoothed gauge link $V_\mu(x)$ back onto SU(3). This is necessary since $V_\mu(x)$ is no longer in the group manifold because it is defined in terms of a sum over neighbouring link

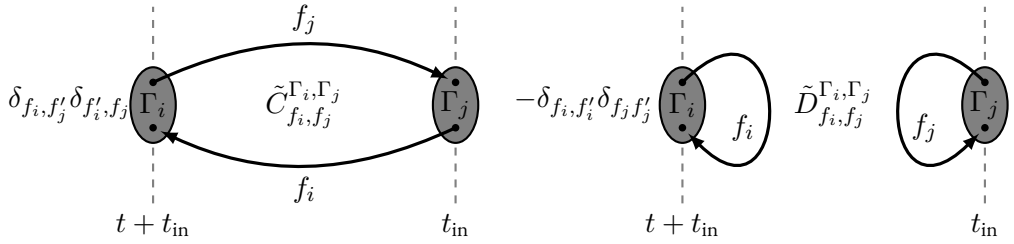


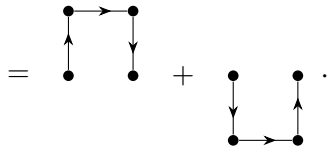
FIGURE 2.5: Graphical representation of connected and disconnected contributions to meson two-point functions, cf. eq. (2.88).

variables,

$$V_\mu(x) = (1-\alpha)U_\mu(x) + \frac{\alpha}{6}\sum_{\nu\neq\mu} C_{\mu\nu}(x), \tag{2.85}$$

where α is a weight parameter and $C_{\mu\nu}(x)$ is a sum over staples,

$$\begin{aligned} C_{\mu\nu}(x) &= U_\nu(x)U_\mu(x+a\hat{\nu})U_\nu^\dagger(x+a\hat{\mu}) \\ &\quad + U_\nu^\dagger(x-a\hat{\nu})U_\mu(x-a\hat{\nu})U_\nu(x-a\hat{\nu}+a\hat{\mu}) \end{aligned} \tag{2.86}$$



The smeared gauge fields are only used in the construction of extended quark sources and sinks but not in the Dirac operator, since this would change the details of the fermion action and imply a different set of (unknown) improvement parameters.

2.3.3 Propagators and disconnected contributions

Having suitable operators at hand, propagators can be constructed by evaluating the fermionic expectation value

$$\langle \mathcal{B}_i(t+t_{\text{in}})\mathcal{B}_j^\dagger(t_{\text{in}})\rangle, \tag{2.87}$$

creating states at some initial time t_{in} and then destroying them at a later time $t+t_{\text{in}}$. For the discussion in this section, we drop spatial coordinates and momentum projection and focus on the Euclidean times. The interpolators $\mathcal{B}_i$ may be any combination of suitable operators, e.g., the pseudoscalar interpolators $\mathcal{P}^a$ or $\mathcal{P}^q$. We will also compute local matrix elements and in these cases local axialvector currents are chosen at the sink time $t+t_{\text{in}}$.

Using Wick's theorem, two-point functions as eq. (2.87) can be related to traces over inverses of the Dirac operators, making use of the Grassmannian nature of the fermion fields. If the interpolating operators $\mathcal{B}_i$ are single biquark fields, the following simple relation is obtained:

$$\begin{aligned} \langle \mathcal{B}_i(t+t_{\text{in}})\mathcal{B}_j^\dagger(t_{\text{in}})\rangle &= \langle q_{f_i}(t+t_{\text{in}})\Gamma_i\bar{q}_{f_i'}(t+t_{\text{in}})\bar{q}_{f_j'}(t_{\text{in}})\Gamma_j q_{f_j}(t_{\text{in}})\rangle \\ &= q^{f_i}(t+t_{\text{in}})\Gamma_i\bar{q}^{f_i'}(t+t_{\text{in}})\bar{q}^{f_j'}(t_{\text{in}})\Gamma_j q^{f_j}(t_{\text{in}}) \\ &= \delta_{f_i,f_j'}\delta_{f_i',f_j}\tilde{C}^{\Gamma_i,\Gamma_j}_{f_i,f_i'}(t+t_{\text{in}},t_{\text{in}}) - \delta_{f_i,f_i'}\delta_{f_j,f_j'}\tilde{D}^{\Gamma_i,\Gamma_j}_{f_i,f_j}(t+t_{\text{in}},t_{\text{in}}), \end{aligned} \tag{2.88}$$

where we refer to the quark flavours of interpolator $\mathcal{B}_i$ with $f_i^{(\prime)} \in \{u,d,s\}$. Likewise, we label the respective Dirac structures with $\Gamma_i \in \{\gamma_5, \gamma_\mu\gamma_5\}$.

If the interpolators are flavour diagonal, $f_i = f_i'$ and $f_j = f_j'$, disconnected contributions $\tilde{D}$ arise that are built by correlating two disconnected loops (one-point functions),

$$\tilde{D}^{\Gamma_i,\Gamma_j}_{\substack{s_i,s_j\\f_i,f_j}}(t+t_{\text{in}},t_{\text{in}}) = L^{\Gamma_i,f_i}_{s_i}(t+t_{\text{in}})L^{\Gamma_j,f_j}_{s_j}(t_{\text{in}}), \tag{2.89}$$

where s_i label the smearing applied to the sources and sinks. Each of these loops contains a trace over space, spin and colour degrees of freedom, so that the quark line closes at the same space-time point,

$$L^{\Gamma,f}_{s}(t,\vec{x}) = a^3 \sum_{\vec{x},\vec{y},\vec{z}} \operatorname{tr}\Gamma\phi^s(x,y)D_f^{-1}(y,z)\phi^s(z,x), \tag{2.90}$$

where ϕ is the smearing kernel defined in eq. (2.83). The large size of the Dirac matrix, D_f, makes an explicit inversion in actual simulations prohibitively expensive and practically impossible and so the above trace is estimated stochastically. This and additional noise reduction techniques are the subject of sec. 4.1.

The expression (2.88) also contains the connected correlation function $\tilde{C}$. This part can be computed very efficiently by exploiting the γ_5-Hermiticity of the Dirac operator, eq. (2.68), and inverting on a smeared point source $\Delta^s_{x_{\text{in}},\alpha_{\text{in}},a_{\text{in}}}(x) = \phi^s(x,x_\in)\delta(\alpha,\alpha_{\text{in}})\delta(a,a_{\text{in}})$. The point-to-all propagator $M^s_f(x) = D_f^{-1}(x,y)\Delta^s(y)$ is a matrix in Dirac and colour space but only a vector in coordinate space and can be used to construct connected correlation functions, starting at a fixed source position x_{in}:

$$\tilde{C}^{\Gamma_i,\Gamma_j}_{\substack{f_i,f_j\\s_i,s_j}}(t,\vec{p}) = a^3V_3 \sum_{\vec{x}} e^{-i\vec{p}\cdot(\vec{x}-\vec{x}_{\text{in}})} \left\langle \Gamma_i\gamma_5\phi^{s_i}M^{s_i}_{f_i}(x)\gamma_5\Gamma_j\phi^{s_j}M^{s_j}_{f_j}(x)\right\rangle, \tag{2.91}$$

where $a^3V_3 = a^3N_1N_2N_3$ is the spatial volume of the lattice. We suppress the dependence of $\tilde{C}$ on the source coordinates x_{in} and in general the signal can be improved by averaging over several source positions to take better advantage of large lattice volumes.

We simulate $N_f = 2+1$ flavours, and so we make use of isospin symmetry by setting $\tilde{C}^{i,j}_{\ell,\ell} := \tilde{C}^{i,j}_{u,u} = \tilde{C}^{i,j}_{u,d} = \tilde{C}^{i,j}_{d,u} = \tilde{C}^{i,j}_{d,d}$ (combining the irrelevant smearing and Dirac indices at the source and sink to multi-indices i,j), and analogously for the disconnected contributions $\tilde{D}$. For generic interpolators $\mathcal{J}^q_i$, for example $\mathcal{J} = \mathcal{P}$ for the pseudoscalar interpolators that are frequently used in this work and in the flavour basis, we obtain

$$\langle \mathcal{J}^\ell_i \mathcal{J}^\ell_j\rangle = \tilde{C}^{i,j}_{\ell,\ell} - 2\tilde{D}^{i,j}_{\ell,\ell}, \tag{2.92a}$$

$$\langle \mathcal{J}^\ell_i \mathcal{J}^s_j\rangle = -\frac{2}{\sqrt{2}}\tilde{D}^{i,j}_{\ell,s}, \tag{2.92b}$$

$$\langle \mathcal{J}^s_i \mathcal{J}^s_j\rangle = \tilde{C}^{i,j}_{s,s} - \tilde{D}^{i,j}_{s,s}, \tag{2.92c}$$

where the flavour combinations are those defined after eqs. (2.80) to (2.82). It is straight-forward to apply eq. (2.88) to sums of biquark fields, for example to the octet and singlet combinations $\mathcal{J}^8$ and $\mathcal{J}^0$,

$$\langle \mathcal{J}^8_i \mathcal{J}^8_j\rangle = \frac{1}{3}\left(\tilde{C}^{i,j}_{\ell,\ell} + 2\tilde{C}^{i,j}_{s,s} - 2\tilde{D}^{i,j}_{\ell,\ell} + 2\tilde{D}^{i,j}_{\ell,s} + 2\tilde{D}^{i,j}_{s,\ell} - 2\tilde{D}^{i,j}_{s,s}\right), \tag{2.93a}$$

$$\langle \mathcal{J}^8_i \mathcal{J}^0_j\rangle = \frac{\sqrt{2}}{3}\left(\tilde{C}^{i,j}_{\ell,\ell} + \tilde{C}^{i,j}_{s,s} - 2\tilde{D}^{i,j}_{\ell,\ell} + 2\tilde{D}^{i,j}_{\ell,s} - \tilde{D}^{i,j}_{s,\ell} + \tilde{D}^{i,j}_{s,s}\right), \tag{2.93b}$$

$$\langle \mathcal{J}^0_i \mathcal{J}^0_j\rangle = \frac{1}{3}\left(2\tilde{C}^{i,j}_{\ell,\ell} + \tilde{C}^{i,j}_{s,s} - 4\tilde{D}^{i,j}_{\ell,\ell} - 2\tilde{D}^{i,j}_{\ell,s} - 2\tilde{D}^{i,j}_{s,\ell} - \tilde{D}^{i,j}_{s,s}\right). \tag{2.93c}$$

It will turn out to be useful to build correlators mixing the octet/singlet and the flavour bases. These reduce to

$$\langle \mathcal{J}_i^8 \mathcal{J}_j^\ell \rangle = \frac{\sqrt{3}}{3}\left(\tilde{C}_{\ell,\ell}^{i,j} - 2\tilde{D}_{\ell,\ell}^{i,j} + 2\tilde{D}_{\ell,s}^{i,j}\right), \tag{2.94a}$$

$$\langle \mathcal{J}_i^0 \mathcal{J}_j^\ell \rangle = \sqrt{\frac{2}{3}}\left(\tilde{C}_{\ell,\ell}^{i,j} - 2\tilde{D}_{\ell,\ell}^{i,j} - \tilde{D}_{\ell,s}^{i,j}\right), \tag{2.94b}$$

$$\langle \mathcal{J}_i^8 \mathcal{J}_j^s \rangle = \frac{1}{\sqrt{3}}\left(\tilde{C}_{s,s}^{i,j} - \tilde{D}_{s,s}^{i,j} - 2\tilde{D}_{s,\ell}^{i,j}\right), \tag{2.94c}$$

$$\langle \mathcal{J}_i^0 \mathcal{J}_j^s \rangle = -\sqrt{\frac{2}{3}}\left(\tilde{C}_{s,s}^{i,j} - \tilde{D}_{s,s}^{i,j} + \tilde{D}_{s,\ell}^{i,j}\right). \tag{2.94d}$$

2.3.4 Analysis of two-point functions

The explicit forms in the previous section allow to evaluate correlators on the lattice, employing connected and disconnected contributions of light and strange flavours to the correlation functions. In this section the connection to physical eigenstates and their analysis is made. To this end correlators can be written in terms of the transfer operator T which shifts all field operators by one lattice unit in the positive time direction [113]. It is related to the Hamiltonian by $T = \exp(-aH)$ with eigenvalues $\exp(-aE_n)$ and has eigenstates denoted by $|n\rangle$ for $n \in \mathbb{Z}_+$. We choose the following normalization

$$\langle m, \vec{p}\,'|n, \vec{p}\,\rangle = 2E_n(\vec{p})V_3\delta_{\vec{p},\vec{p}\,'}\delta_{m,n}, \tag{2.95}$$

where $\vec{p}$ and $\vec{p}\,'$ are the spatial momenta of the states.

Using the multiple applications of the transfer operator, correlators like eq. (2.87) can be spectrally decomposed into

$$\langle \mathcal{B}_i(t)\mathcal{B}_j^\dagger(0)\rangle = \langle \hat{\mathcal{B}}_i T^{t/a}\hat{\mathcal{B}}_j^\dagger\rangle = \sum_{n=1}^{\infty}\langle\Phi_i|n\rangle\exp(-tE_n)\langle n|\Phi_j\rangle, \tag{2.96}$$

where we implicitly made use of translational invariance to average over all available initial times t_{in} and the hat indicates that the operators are time independent and defined in the Heisenberg picture. Hence, all time dependence is contained in the exponential factors, the eigenvalues of the transfer operator. Since the currents considered in this work do not possess a vacuum expectation value, i.e. $\langle\Phi_i|\Omega\rangle = 0$, we do not consider the vacuum state in the counting and identify the lowest energy eigenstate with the ground state energy of the physical state that possesses the corresponding quantum numbers, for example $|0\rangle = |\eta\rangle$, $|1\rangle = |\eta'\rangle$, etc.

To extract energies of states, the left-hand side of eq. (2.96) is evaluated on the lattice and then fitted to effective parametrizations using the right-hand side, necessarily truncating the sum at some finite state N. A precise extraction of the ground state (and also higher lying states) is challenging if there exist near-by excitations and the achievable precision depends crucially on both the energy differences and the relative magnitudes of the time-independent amplitudes $\langle\Phi|n\rangle$. These describe the overlap of the created interpolating function with the n-th state and depend on the details of the lattice calculation, i.e., on the exact choice of the interpolator and particularly also on the quark smearing employed. This allows for some tuning, enhancing some states over others that would otherwise disturb the extraction of the states of interest. Such improvements can be achieved by finding an optimal smearing radius and by combining several interpolating functions $\mathcal{B}_i$. In that case eq. (2.96) is a matrix equation and extracting physical states $|n\rangle$ and their energies E_n amounts to a diagonalization of this matrix. In actual computations, the infinite sum must be truncated at some finite highest resolvable state N leading to truncation errors of $\mathcal{O}\left(\exp(-(E_{N+1} - E_N)t)\right)$. The optimal truncation N is crucial for a reliable extraction of the lowest states: On the one hand, if N is too small, higher states spoil the lower states. On the other hand, if N is too large the

multi-exponential cannot be reliably resolved from the finite statistics of the data. We will focus on techniques for precise extractions of physical states in sec. 4.2.3.

2.3.5 Local matrix elements and decay constants

Physical states can also be destroyed by local currents and their matrix elements describe the coupling to the corresponding downstream decay processes. For example, in the case of charged pseudoscalar meson states, a flavour changing vector boson $W_\mu^\pm$ can be emitted that may further decay into a lepton-neutrino pair. Flavour diagonal pseudoscalar mesons can decay strongly, i.e., through annihiliation into gluons, and contain a singlet component opening the door to anomalous decays, see fig. 2.1.

We recall the definition eq. (2.30) of decay constants $F^a_{\mathcal{M}}$ of a pseudoscalar $\mathcal{M}(p)$ with four-momentum p,

$$\langle \Omega | A^a_\mu | \mathcal{M}(p) \rangle = i p_\mu F^a_{\mathcal{M}}, \tag{2.97}$$

where for $\mathcal{M} = \eta$ and $\mathcal{M} = \eta'$ in the isospin limit singlet ($a = 0$) and octet ($a = 8$) currents contribute and the axialvector currents are defined in the continuum in eq. (2.28). We will discuss the corresponding renormalized and improved current definitions for the lattice discretization in sec. 2.4.2.

Eq. (2.97) means that the η/η' system is described by in total four independent decay constants, a singlet and an octet one for each of the particles. The singlet and octet flavour combinations of currents are given by

$$J^8 = \frac{1}{\sqrt{12}} \left(J^u + J^d - 2J^s \right) = \frac{1}{\sqrt{6}} J^\ell - \frac{1}{\sqrt{3}} J^s, \tag{2.98}$$

$$J^0 = \frac{1}{\sqrt{6}} \left(J^u + J^d + J^s \right) = \frac{1}{\sqrt{3}} J^\ell + \frac{1}{\sqrt{6}} J^s, \tag{2.99}$$

where $J \in \{A_\mu, P\}$ and $J^\ell = \frac{1}{\sqrt{2}} \left(J^u + J^d \right)$. Note that these differ from the extended quark interpolators eqs. (2.80) to (2.82) by $1/\sqrt{2}$.

This normalization corresponds to $F_{\pi^0} \approx 92\,\text{MeV}$ at the physical point and the definition is related to f_π by $F_{\pi^0} = f_\pi/\sqrt{2}$ which is typically more popular in the lattice literature. Since there is practically no strong mixing to the neutral pion, $F_{\pi^0} \approx F^3_{\pi^0}$. For the charged pions, $\mathcal{M} = \pi^\pm$, $F_{\pi^+} = \sqrt{2} F^1_{\pi^+} = -\sqrt{2} i F^2_{\pi^+} = -\sqrt{2} F^1_{\pi^-} = \sqrt{2} i F^2_{\pi^-}$, cf. eq. (2.41), and similarly for the kaons. Along the symmetric line of exact SU(3) flavour symmetry, where also the strange quark is mass degenerate, $m_u = m_d = m_s$, this also implies $F_{\pi^0} = F^8_\eta = F_{\eta_8}$, whereas the singlet decouples due to the axial $\mathrm{U}(1)_A$ anomaly, $F^0_{\eta'} = F_{\eta_0}$ and $F^0_\eta = F^8_{\eta'} = 0$.

In general the four decay constants can also be reparameterized in terms of two decay constants F^8, F^0 and two mixing angles θ_8 and θ_0 for each flavour combination,

$$\begin{pmatrix} F^8_\eta & F^0_\eta \\ F^8_{\eta'} & F^0_{\eta'} \end{pmatrix} = \begin{pmatrix} F^8 \cos\theta_8 & -F^0 \sin\theta_0 \\ F^8 \sin\theta_8 & F^0 \cos\theta_0 \end{pmatrix}, \tag{2.100}$$

where in the exact SU(3) flavour limit, $\theta_8 = \theta_0 = 0$.

As will be discussed in detail in sec. 2.4, the octet and singlet axialvector currents differ in their anomalous dimensions and renormalize differently. For this reason this definition is the more natural, whereas from a flavour symmetry perspective another popular choice is

$$\left\langle \Omega \left| A^q_\mu \right| \mathcal{M}(p) \right\rangle = i\sqrt{2} F^q_{\mathcal{M}} p_\mu, \tag{2.101}$$

where $q \in \{\ell, s\}$ labels the flavour. An analogous parametrization to eq. (2.100) can also be introduced in this so-called flavour basis

$$\begin{pmatrix} F_\eta^\ell & F_\eta^s \\ F_{\eta'}^\ell & F_{\eta'}^s \end{pmatrix} = \begin{pmatrix} F^\ell \cos\phi_\ell & -F^s \sin\phi_s \\ F^\ell \sin\phi_\ell & F^s \cos\phi_s \end{pmatrix}. \tag{2.102}$$

The flavour basis is more suitable when the renormalization scale is a priori unknown, e.g. in phenomenological studies and chiral pertubation theory predictions fitting to experiments. The reason is that the difference between ϕ_ℓ and ϕ_s is formally an $1/N_c$ effect and can be neglected for most phenomenological extractions [114, 115]. The two definitions of the decay constants are related by an orthogonal transformation

$$\begin{pmatrix} F_\mathcal{M}^8 \\ F_\mathcal{M}^0 \end{pmatrix} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & -\sqrt{2} \\ \sqrt{2} & 1 \end{pmatrix} \begin{pmatrix} F_\mathcal{M}^\ell \\ F_\mathcal{M}^s \end{pmatrix}. \tag{2.103}$$

2.4 Renormalization and improvement

Since the lattice spacing acts as an ultraviolet regulator for the quantized theory, bare results $J(a)$ usually diverge when approaching the continuum limit, $a \to 0$. As in continuum calculations, the regulator must be removed by renormalizing the respective quantities and matching to a continuum scheme as, for example, the popular $\overline{\text{MS}}$ scheme. Then, the renormalized combination $\widehat{J} = Z_J(a)J(a)$ remains finite in the continuum.

In many cases, results in the continuum still depend on the renormalization scale μ of the process under consideration. This dependence is described by the — renormalization scheme dependent — anomalous dimension of the current J,

$$\gamma_J(a_s) = \mu^2 \frac{\mathrm{d}Z_J}{\mathrm{d}\mu^2} = -\sum_{n\geq 0} \gamma_J^{(n)} a_s^{n+1}. \tag{2.104}$$

On the lattice renormalization factors Z_J depend not only on the value of the lattice spacing (through the lattice coupling), but also on the details of the discretization, i.e., on the exact gauge and fermion actions and on the exact renormalization prescription. Properly renormalized, physical and measurable results are finite and must agree in the continuum limit but at any non-zero value of the lattice spacing, differences between renormalization schemes can be expected. The renormalization of lattice results and the removal of $\mathcal{O}(a)$ lattice spacing effects determines how the continuum limit is approached and how reliable that limit can be taken. The numerical values of the renormalization factors and improvement coefficients presented in this section are specific to the action used in the simulations that are presented in ch. 3, but the general discussion is of broader use and the discussion of linear improvement coefficients is universal for all simulations with Wilson Clover fermions independent of the gauge action used (which induces only higher $\mathcal{O}(a^2)$ lattice spacing effects).

In this section we present the renormalization and $\mathcal{O}(a)$ improvement of the quantities that we focus on in this work, that is local axialvector matrix elements required for the computation of decay constants. We stress the different renormalization and improvement of singlet ($a = 0$) and non-singlet ($a \neq 0$) currents. The determination of the relevant matrix elements of the η and η', see sec. 2.3.5, will be described in sec. 4.3. Later in this work we will also determine meson masses and gluonic matrix elements. The former are renormalization group invariant and do not require renormalization whereas the renormalization of the latter is more complicated and deferred to ch. 6.

2.4.1 Improvement of the coupling and taking the continuum limit

The lattice spacing in simulations of QCD is tuned via the lattice coupling. Typically, dimensionless combinations of observables with the lattice spacing are determined, renormalized and converted to physical units by dividing out the known lattice spacing for that coupling $a^*(\beta)$. Then, using these combinations the continuum limit is taken,

$$\widehat{J}(a,\{m_f\} = Z_J(a^*)\left[a^n J(a,\{m_f\})\right]\left[a^*(\beta)\right]^{-n} \xrightarrow{a\to 0} \widehat{J}(\mu,\{m_f\}), \tag{2.105}$$

where $\{m_f\}$ represents the (valence and sea) quark mass dependence of the observable, and we assume a multiplicative renormalization of the observable.

However, there are mass-dependent corrections to the lattice coupling [116, 117],

$$\tilde{g}^2 = \tilde{g}^2(g^2, a\overline{m}) = g^2\left[1 + b_g(g^2)a\overline{m}\right] + \mathcal{O}(a^2), \tag{2.106}$$

where $a\overline{m} = \sum m_f/N_f$. Therefore, the point in the quark mass plane where the lattice spacing is defined matters, when we want to keep the lattice spacing fixed over a series of ensembles that vary in their sea quark masses, see also sec. 3.4.1. In the above equation, $b_g(g^2)$ is an as yet unknown improvement coefficient function. However, the one-loop coefficient is known [117],

$$b_g(g^2) = 0.012000(2)N_f g^2 + \mathcal{O}(g^4). \tag{2.107}$$

In order to work at a fixed lattice spacing, we would need to increase β when approaching the physical point. Without a more precise knowledge of $b_g(g^2)$ and possible higher order effects, this seems impractical and instead, we write down an $\mathcal{O}(a)$-"improved" lattice spacing,

$$a(g^2, a\overline{m}) = a\left(g^2\left(1 + b_g a\overline{m}\right)\right) = a(g^2, 0)\left(1 + b_a a\overline{m}\right) \coloneqq a_0\left(1 + b_a a\overline{m}\right). \tag{2.108}$$

The relation between the improvement coefficient b_a and b_g can be derived by integrating out the lattice β-function for our action [90], and yields in $N_f = 3$ in $N_f = 3$ to leading order,

$$b_a^{(1)}(g^2) = -\frac{2\pi^2}{\beta_0 g^2} b_g^{(1)} = 0.31583(5) + \mathcal{O}(g^2), \tag{2.109}$$

where β_0 is the first coefficient of the β-function as defined in sec. 2.1.3.

When taking the continuum limit, it is thus beneficial to extrapolate dimensionless combinations of two observables O and P instead of combinations with the lattice spacing, eq. (2.105):

$$\frac{Z_J(a)a^n J(a,\{m_f\})}{Z_K(a)a^n K(a,\{m_f\})} \xrightarrow{a\to 0} \frac{\widehat{J}(\mu,\{m_f\})}{\widehat{K}(\mu,\{m_f\})}, \tag{2.110}$$

where it is assumed that J and K have the same mass dimension n or that the auxiliary observable K has been transformed accordingly. The linear corrections, eq. (2.108), cancel from these ratios and if the lattice measurements $a^n K$ and $a^n J$ both have no other linear cut-off effects, then the combination on the left-hand side is also free of linear discretization effects. In this work we are interested in masses that are automatically free of $\mathcal{O}(a)$ effects and decay constants of which the complete linear improvement will be discussed in the following sections. As auxiliary observable, we will use the (renormalization group invariant) Wilson flow scale, $a^n K = \left(a^2/t_0\right)^{n/2}$, that will be discussed further in sec. 3.4. We then extrapolate dimensionless combinations of the Wilson flow scale with hadron masses $\sqrt{t_0}M_{\mathcal{M}}$ or decay constants $\sqrt{t_0}F^a_{\mathcal{M}}$. Both the observable and t_0/a^2 have been measured on the same ensemble and thus at the same lattice spacing $a(g^2, a\overline{m})$ of eq. (2.108). Due to the cancellations mentioned above, we then obtain results that are free of $\mathcal{O}(a)$ effects that stem from the coupling and only need to parametrize the remaining cut-off effects when

taking the continuum limit.

2.4.2 Improvement of η and η' decay constants

We define unrenormalized pseudoscalar and axialvector local fermion bilinear operators on the lattice,

$$\tilde{P}^a = \overline{\psi}\gamma_5 t^a \psi, \qquad \tilde{P}^0 = \overline{\psi}\gamma_5 t^0 \psi + a g_P \operatorname{tr} F_{\mu\nu}\widetilde{F}_{\mu\nu}, \tag{2.111}$$

$$\tilde{A}^a_\mu = \overline{\psi}\gamma_\mu\gamma_5 t^a \psi + a c_A \partial_\mu \overline{\psi}\gamma_5 t^a \psi, \qquad \tilde{A}^0_\mu = \overline{\psi}\gamma_\mu\gamma_5 t^0 \psi + a c^s_A \partial_\mu \overline{\psi}\gamma_5 \psi, \tag{2.112}$$

where in this case $a = 1, \ldots, 8$ and c_A, c^s_A and g_P are the mass independent improvement coefficients for our action [118]. While the non-singlet improvement coefficient c_A has been determined non-perturbatively [119], its singlet equivalent c^s_A is unknown, and we parametrize it as $c^s_A = c_A + \delta c_A$, where δc_A is of order g^4. Also a value of g_P is not known for our action. We denote the singlet current where we set $\delta c_A = 0$ by $\check{A}^0_\mu$ and the singlet pseudoscalar current setting $g_P = 0$ by $\check{P}^0$.

Using the above currents, we first determine unrenormalized and only partially improved decay constants

$$\tilde{F}^8_n E_n(p) = \langle \Omega | \tilde{A}^8_\mu | n(p) \rangle, \tag{2.113}$$

$$\check{F}^0_n E_n(p) = \langle \Omega | \check{A}^0_\mu | n(p) \rangle, \tag{2.114}$$

see eq. (2.97) and the energies $E_n(p)$ of the states $|n(p)\rangle$ are defined as the corresponding eigenvalues of the transfer operator, eq. (2.96). The technical details of the extraction of physical states $|\eta\rangle$ and $|\eta'\rangle$ will be given in ch. 4 and in particular sec. 4.3 for the construction of such matrix elements. It is also useful to define the pseudoscalar matrix elements,

$$\tilde{H}^8_n = \langle \Omega | \tilde{P}^8 | n \rangle, \tag{2.115}$$

$$\check{H}^0_n = \langle \Omega | \check{P}^0 | n \rangle. \tag{2.116}$$

While $\tilde{H}^8_n$ is fully $\mathcal{O}(a)$ improved in the chiral limit, $\check{H}^0_n$ is still incomplete, cf. eq. (2.111):

$$\tilde{H}^0_n = \check{H}^0_n + a c^s_P \left\langle \Omega \left| \partial_\mu A^0_\mu \right| n \right\rangle, \tag{2.117}$$

where we have replaced $a g_P \operatorname{tr} F_{\mu\nu}\widetilde{F}_{\mu\nu} \mapsto a c^s_P \left\langle \Omega \left| \partial_\mu A^0_\mu \right| n \right\rangle$ by virtue of the singlet axial Ward identity in the chiral limit, eq. (2.39). Note that this substitution changes mass dependent linear improvement terms of the pseudoscalar matrix elements.

The derivative acting on the pseudoscalar matrix elements in eqs. (2.112) and (2.117) can be carried out in time direction and this yields, when acting on a correlation function of that state as discussed in sec. 2.3.4,

$$\begin{aligned} \langle \Omega | \partial_t J^a(t) \mathcal{B}_i(0) | \Omega \rangle &= \frac{1}{2a} \langle \Omega | (J(t+a) - J(t-a)) \mathcal{B}_i(0) | \Omega \rangle \\ &= -\frac{\sinh(aE_n)}{a} \langle \Omega | J^a(t) \mathcal{B}_i(0) | \Omega \rangle \\ &\approx -E_n \langle \Omega | J^a(t) \mathcal{B}_i(0) | \Omega \rangle, \end{aligned} \tag{2.118}$$

where J is one of currents in eqs. (2.111) and (2.112). In this sense, we can substitute $\langle \Omega | \partial_t P^a | n \rangle$ for $-E_n H^a_n$, up to $\mathcal{O}(a^2)$ corrections.

Using these abbreviations, we can write the singlet decay constants as

$$\tilde{F}^0_n = \check{F}^0_n + a\, \delta c_A \check{H}^0_n, , \tag{2.119}$$

leaving the parameter δc_A as a free parameter that will enter our continuum limit fit. The use of the incompletely improved $\check{H}_n^0$ instead of, e.g., $\tilde{H}_n^0$ leads to different $\mathcal{O}(a^2)$ terms but is irrelevant for the removal of linear lattice spacing effects.

Regarding the mass dependent improvement of the octet current, we use eq. (15) of [118], where full $\mathcal{O}(a)$ improvement has been worked out for generic bilinear quark operators employing Wilson fermions. For the corresponding decay constants this translates into

$$F^8_{\eta^{(\prime)}} = Z_A \Big[(1+3a\tilde{b}_A\overline{m})\tilde{F}^8_{\eta^{(\prime)}} + \frac{a}{\sqrt{3}} b_A \Big(m_\ell \check{F}^\ell_{\eta^{(\prime)}} - \sqrt{2} m_s \check{F}^s_{\eta^{(\prime)}} \Big) - \sqrt{2} a f_A \delta m \check{F}^0_{\eta^{(\prime)}} \Big] + \mathcal{O}(a^2), \tag{2.120}$$

where Z_A is the renormalization factor for non-singlet currents on which we give more details in the following section. The mass dependent improvement terms are parametrized by $\tilde{b}_A$, b_A and f_A, while the mass independent improvement coefficient c_A is already included in $\tilde{F}^8_{\eta^{(\prime)}}$. It turns out (see again eq. (15) of [118]) that even for the improvement of the octet axialvector current, singlet currents are required. Note that within the $\mathcal{O}(a)$ improvement terms we can replace any unimproved lattice decay constant by either $\tilde{F}$ or $\check{F}$ since the difference will only have an $\mathcal{O}(a^2)$ effect on the result. These replacements are convenient for performing the continuum extrapolation, as will be discussed in sec. 5.2. Subsequently, $\check{F}^\ell_{\eta^{(\prime)}}$ and $\check{F}^s_{\eta^{(\prime)}}$ are obtained from $\check{F}^0_{\eta^{(\prime)}}$ and $\check{F}^8_{\eta^{(\prime)}}$ via the rotation (2.103). The quark masses, their averages and differences that appear in eq. (2.120) are given by

$$am_{0,f} = \frac{1}{2}\left(\frac{1}{\kappa_f} - \frac{1}{\kappa_{\text{cr}}}\right), \quad a\overline{m} = \frac{a}{3}\left(2m_\ell + m_s\right), \quad a\delta m = am_s - am_\ell. \tag{2.121}$$

The critical hopping parameter $\kappa_{\text{cr}}(\beta)$ was determined for our action and lattice spacings in [89, 90].

The improvement coefficients b_A and $\tilde{b}_A$ have been determined non-perturbatively in [120, 121]. The sea quark coefficient multiplying $a\overline{m}$ slightly differs from the definition of $\bar{b}_A$ in [118], where the lattice spacing is assumed to be fixed. As discussed in sec. 2.4.1 we work at a fixed value of the coupling and correspondingly absorb the effect of b_g into $\tilde{b}_A$, see the definition and discussion in [120]. The corresponding coefficients for our β values from [121] read

$$\begin{aligned} \tilde{b}_A(\beta = 3.4) &= -0.11(13), \qquad \tilde{b}_A(\beta = 3.46) = 0.10(11), \\ \tilde{b}_A(\beta = 3.55) &= -0.04(12), \qquad \tilde{b}_A(\beta = 3.70) = -0.05(8). \end{aligned} \tag{2.122}$$

For the valence quark coefficient b_A we use the parametrization [121]

$$b_A(g^2) = 1 + 0.0881\, C_F g^2 + b\, g^4, \quad \text{where} \quad b = 0.0113(44) \tag{2.123}$$

and $g^2 = 6/\beta$ and $C_F = 4/3$. The improvement term in eq. (2.120) that is proportional to the difference of the quark masses is only present in flavour diagonal currents. Its coefficient, f_A, is unknown, and formally it is of $\mathcal{O}(g^6)$ [118, 122]. This is the only unknown parameter needed to achieve full $\mathcal{O}(a)$ improvement of the octet decay constants, and we incorporate it into the functional form of the continuum extrapolation, see sec. 5.2.

Regarding the improvement of the singlet decay constants, utilizing eq. (23) of [118], we obtain

$$F^0_{\eta^{(\prime)}} = Z^s_A \left[\left(1 + 3a\tilde{d}_A\overline{m}\right) \tilde{F}^0_{\eta^{(\prime)}} + \frac{1}{\sqrt{3}} a d_A \left(\sqrt{2} m_l \check{F}^\ell_{\eta^{(\prime)}} + m_s \check{F}^s_{\eta^{(\prime)}} \right) \right]. \tag{2.124}$$

Again, we replaced the lattice decay constants within the $\mathcal{O}(a)$ terms by partially improved ones and replace $\bar{d}_A \mapsto \tilde{d}_A$. The renormalization factor $Z^s_A \neq Z_A$ is discussed in the next subsection.

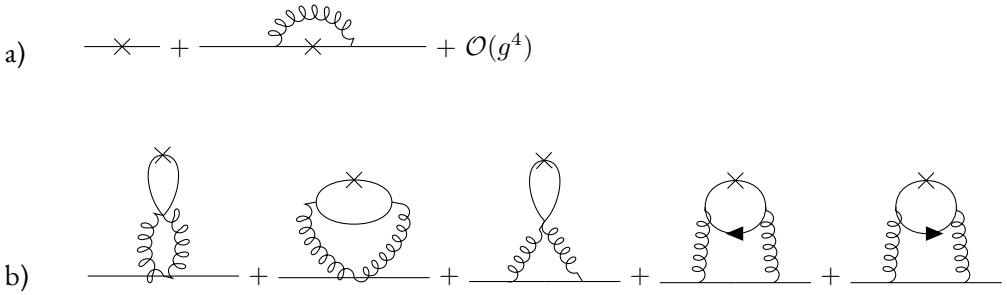


FIGURE 2.6: Contributions to the full operator in lattice perturbation theory. The cross represents the current insertion. The upper row (a) is truncated at $\mathcal{O}(g^2)$ and the second row (b) lists all two-loop contributions ($\mathcal{O}(g^4)$) that appear exclusively in the singlet case [125] and that are the leading contributions to the difference $Z_A^s - Z_A$. These arise solely through "disconnected" quark loops — only connected via gluon propagators — that are only possible for flavour diagonal currents. Note that the "tadpole"-like graphs that involve a two-quark-two-gluon vertex (the first three) have no continuum analogue.

Unfortunately, both improvement coefficients $d_A = b_A + \mathcal{O}(g^4)$ and $\tilde{d}_A = \mathcal{O}(g^4)$ are only known to $\mathcal{O}(g^2)$ in perturbation theory. In analogy to f_A, we will include these parameters in the continuum extrapolation formulae (along with δc_A, see eq. (2.119)) in ch. 5.2.

2.4.3 Renormalization of the singlet and non-singlet axialvector currents

We now turn to the renormalization factors Z_A and Z_A^s in eqs. (2.120) and (2.124). In the continuum theory vector and non-singlet axialvector currents are both conserved in massless QCD (see eq. (2.29)) and so they do not acquire an anomalous dimension, i.e., they do not depend on the renormalization scale μ. The breaking of chiral symmetry via the Wilson term leads to non-trivial renormalization factors on the lattice, and for the non-singlet axialvector currents we use the non-perturbatively determined values of $Z_A = Z_A^8$ for our action that can be found in the $Z^l_{A,\text{sub}}$ column of tab. 6 of [123].

In the continuum $\overline{\text{MS}}$ scheme, the singlet axialvector current acquires an anomalous dimension through the axial anomaly [77, 124],

$$\mu^2 \frac{\mathrm{d}}{\mathrm{d}\mu^2} \langle \Omega | \widehat{A}^0_\mu(\mu) | \mathcal{M} \rangle = -\frac{N_f}{2} \frac{\alpha_s^2}{\pi^2} \langle \Omega | \widehat{A}^0_\mu(\mu) | \mathcal{M} \rangle + \mathcal{O}(\alpha_s^3). \tag{2.125}$$

The scale dependence only enters as a two-loop effect and this can be understood by looking at the Feynman diagrams that contribute exclusively to the singlet renormalization, cf. fig. 2.6: Since the current is diagonal in flavour space, the current insertion can also take place in a fermionic loop that is connected through gluons to the external legs. The additional contributions are thus a pure sea-quark effect.

Correspondingly, to one loop $\gamma^{(0)}_{A^a_\mu} = \gamma^{(0)}_{A^0_\mu} = 0$, and the first three coefficients of the γ-function eq. (2.104) for the singlet axialvector current are [77, 124][3]

$$\gamma^{(0)}_{A^0_\mu} = 0, \tag{2.126}$$

$$\gamma^{(1)}_{A^0_\mu} = \frac{3}{8} C_F N_f, \tag{2.127}$$

$$\gamma^{(2)}_{A^0_\mu} = \frac{1}{64} \left[\left(\frac{142}{3} C_F C_A - 18 C_F^2 \right) N_f - \frac{4}{3} C_F N_f^2 \right], \tag{2.128}$$

[3] Recently, $\gamma^{(3)}_{A^0_\mu}$ has been computed, too [126].

where in QCD $C_A = 3$ and $C_F = 4/3$.

From the β- and γ-functions, eqs. (2.17) and (2.104), one can easily derive the scale evolution of local currents:

$$Z_J\left(a_s(\mu_1),\mu_1\right) = Z_J\left(a_s(\mu_0),\mu_0\right) \exp\left(\int_{a_s(\mu_0)}^{a_s(\mu_1)} \mathrm{d}a\, \frac{\gamma_J(a)}{\beta(a)}\right). \tag{2.129}$$

Normally, to leading non-trivial order, the evolution factor is given by $(a_s(\mu_1)/a_s(\mu_0))^{\gamma_J^{(0)}/\beta_0}$, which diverges if one of the scales is sent to infinity. In our case, however, $\gamma^{(0)}_{A^0_\mu} = 0$, leading to a finite renormalization group evolution

$$\frac{Z_A^s\left(a_s(\mu_1),\mu_1\right)}{Z_A^s\left(a_s(\mu_0),\mu_0\right)} = \exp\left\{\frac{\gamma^{(1)}_{A^0_\mu}}{\beta_0}\left[\left(a_s(\mu_1) - a_s(\mu_0)\right)\right.\right. \\ \left.\left. + \frac{1}{2}\left(\frac{\gamma^{(2)}_{A^0_\mu}}{\gamma^{(1)}_{A^0_\mu}} - \frac{\beta_1}{\beta_0}\right)\left(a_s^2(\mu_1) - a_s^2(\mu_0)\right) + \ldots\right]\right\}. \tag{2.130}$$

This suggests a modified scheme (see, e.g., [127]), where the renormalization group running is absorbed into the renormalization constant:

$$\widehat{A}^{0\prime}_\mu = \left[Z_A^s(\infty)/Z_A^s(\mu)\right]\widehat{A}^0_\mu = Z_A^{s\prime} A^0_\mu, \tag{2.131}$$

and

$$Z_A^{s\prime} = Z_A^s(\mu=\infty) = \left[1 - \frac{\gamma^{(1)}_{A^0_\mu}}{\beta_0}a_s(\mu) + \frac{\gamma^{(1)}_{A^0_\mu}}{2\beta_0}\left(\frac{\gamma^{(1)}_{A^0_\mu}}{\beta_0} + \frac{\beta_1}{\beta_0} - \frac{\gamma^{(2)}_{A^0_\mu}}{\gamma^{(1)}_{A^0_\mu}}\right)a_s^2(\mu) + \cdots\right] Z_A^s(\mu). \tag{2.132}$$

Similar to the renormalization group invariant (RGI) scheme, in the above $\overline{\mathrm{MS}}'$ scheme the current is scale independent and the corresponding γ-function is trivial: $\gamma'_{A^0_\mu} = 0$. Renormalizing the singlet axialvector current in this way corresponds to the usual $\overline{\mathrm{MS}}$ scheme but setting $\mu = \infty$.

At present, the difference $Z_A^{s\prime} - Z_A$ has only been computed in perturbation theory, cf. fig. 2.6. Setting c_{SW} to its leading order value $c_{\mathrm{SW}} = 1$ within eq. (32) of [125], we obtain for our action

$$Z_A^s(\mu) = Z_A - a_s^2(a^{-1})\left[\gamma^{(1)}_{A^0_\mu}\ln(\mu^2a^2) + 2.834(11)\right], \tag{2.133}$$

where again we use the non-perturbative Z_A values of [123]. Note that we have replaced $g^2 \mapsto 4\pi^2 a_s(a^{-1})$, which is valid to this order in perturbation theory. Within the above conversion to the $\overline{\mathrm{MS}}$ scheme we vary the scale $\mu \in [\frac{1}{2}a^{-1}, 2a^{-1}]$ in order to estimate the systematics of omitting higher perturbative orders and take $\mu = a^{-1}$ for our central values. The results are then run via eq. (2.130) (not eq. (2.132)) to $\mu = \infty$ to obtain the scale independent $\overline{\mathrm{MS}}'$ result. This is carried out using the three-loop $\gamma_{A^0_\mu}$-function and, for the running of $a_s(\mu)$, starting from the value determined in [128], the five-loop β-function [129] (as implemented in version 3 of the RunDec package for Mathematica [130, 131]). For convenience, we also quote results in the more commonly used scale dependent prescription at the scales $\mu = 10\,\mathrm{GeV}$, $\mu = 2\,\mathrm{GeV}$ and $\mu = 1\,\mathrm{GeV}$ in QCD with $N_f = 3$ active quark flavours. The corresponding conversion factors are listed in tab. 2.1.

μ	$Z_A^s(\mu)/Z_A^{s\prime}$			
RG running	2-loop	2-loop	3-loop	3-loop
β-function	2-loop	5-loop	3-loop	5-loop
1 GeV	$1.0881(^{28}_{27})$	$1.0913(^{29}_{28})$	$1.1387(^{68}_{63})$	$1.1383(^{70}_{64})$
2 GeV	$1.0565(^{10}_{10})$	$1.0590(^{10}_{10})$	$1.0754(^{16}_{16})$	$1.0752(^{16}_{16})$
10 GeV	$1.0329(^{3}_{3})$	$1.0341(^{3}_{3})$	$1.0390(^{4}_{4})$	$1.0389(^{4}_{4})$

TABLE 2.1: Conversion factors $Z_A^s(\mu)/Z_A^{s\prime} = Z_A^s(\mu)/Z_A^s(\infty)$, computed according to eq. (2.130) for $N_f = 3$, combining different orders of the renormalization group running with different orders of the running of the coupling. The errors reflect the uncertainty in the Λ-parameter of $N_f = 3$ QCD [63].

3 Generation of $N_f = 2 + 1$ gauge ensembles with $\mathcal{O}(a)$ improved Wilson quarks

Since the precision of lattice calculations can always be improved by increasing statistics and simulating at finer lattice spacings or quark masses that are closer to the physical ones, simulations are always limited by the available compute power. A large fraction of the total computing time goes into the generation of the ensembles that can then be reused for a variety of observables in different contexts. For this reason it is common to jointly generate and share configurations that are then analyzed in smaller groups.

The ensembles that were analyzed in the course of this work are a subset of the ensembles that have been produced within the coordinated lattice simulations (CLS) initiative. The focus is put on large volume ensembles with $N_f = 2 + 1$ flavours of non-perturbatively improved Wilson-Clover fermions on a tree level improved Lüscher-Weisz gauge background. The general algorithmic ideas have already been discussed in sec. 2.2.4 and are described in detail in [132]. Many of the ensembles and the algorithmic setup used to generate them were introduced in [89, 133, 134], and here we give some details on the quark mass tuning to facilitate extrapolations to the physical point in sec. 3.1, before an overview over available ensembles in terms of their position in the quark mass plane and the lattice spacings is given in sec. 3.2. Some of the simulations employ very fine lattice spacings down to $a \approx 0.039$ fm for which open boundary conditions are necessary to prevent the freezing of the topology. We revisit autocorrelations in the simulations and introduce open boundary conditions in sec. 3.3. The scale has been set first in [135] via the intermediate Wilson flow scale and here we present an update and a global parametrization of the values of t_0^*/a^2 over all available lattice couplings in sec. 3.4 which will appear in [90]. There, also other important parameters like the critical hopping parameters are determined that we refer to at various places.

Sections 3.1 to 3.3 follow roughly the ideas of publications [89, 133, 134] that the author was involved in. The remainder on scale setting, sec. 3.4, will be part of a forthcoming publication [90] in similar or verbatim form.

3.1 Quark mass tuning at finite lattice spacing

Because with Wilson fermions chiral symmetry is broken explicitly by the doubler term, cf. the discussion in sec. 2.2.3, approaching the chiral limit or the point of physical quark masses is possible only with prior knowledge from previous simulations. In other words, the position of the physical point in terms of the physical input parameters of the simulation $(\beta, \kappa_\ell, \kappa_s)$ is unknown and related non-trivially through renormalization and lattice spacing effects.

Also, since large volume simulations at fine lattice spacings are computationally only feasible at quark masses larger than physical, it is necessary and more cost-effective to simulate a number of ensembles on trajectories that extrapolate to the physical point in a straight-forward manner. In practice, however, such extrapolations of physical observables cannot be carried out as simple linear fits: Possible mass mistuning needs to be taken into account, and higher orders of chiral perturbation theory are required to incorporate all data. üWell-defined quark mass trajectories that

lead to the physical point are still desirable for practical reasons and here we motivate the choices made for the CLS ensembles.

We start with defining the physical quark mass point of $N_f = 2 + 1$ flavour QCD as the point where the (charged) pion and kaon masses take their experimental values after subtracting electromagnetic and isospin breaking effects. Since we simulate pure QCD — neglecting electromagnetic corrections — with two degenerate light quarks, we are insensitive to such effects. This is described in detail in [136][1] and the following values are obtained:

$$M_\pi^{\text{ph}} = 134.8(3)\,\text{MeV}, \qquad M_K^{\text{ph}} = 494.2(3)\,\text{MeV}. \tag{3.1}$$

In the following we will use the superscript "ph" to refer to quantities at the physical quark mass point and in the continuum limit.

We can define the squared average and difference of the light pseudoscalar masses as

$$\overline{M}^2 \coloneqq \frac{2M_K^2 + M_\pi^2}{3}, \qquad \delta M^2 \coloneqq 2\left(M_K^2 - M_\pi^2\right). \tag{3.2}$$

At the above defined physical point, their values are

$$\overline{M}^{\text{ph}} = 410.9(5)\,\text{MeV}, \qquad \delta M^{\text{ph}} = 672.4(9)\,\text{MeV}. \tag{3.3}$$

To leading order these are related to the quark masses by the Gell-Mann–Oakes–Renner (GMOR) relations,

$$\overline{M}^2 \approx 2B_0\widehat{\overline{m}} = 2B_0\left(2\widehat{m}_\ell + \widehat{m}_s\right)/3, \qquad \delta M^2 \approx 2B_0\delta\widehat{m} = 2B_0\left(\widehat{m}_s - \widehat{m}_\ell\right), \tag{3.4}$$

where $B_0 = -\langle\overline{u}u\rangle/F^2$ is the ratio of the chiral condensate $\langle\overline{u}u\rangle < 0$ over the (squared) pion decay constant in the SU(3) chiral limit, F. and corrections to the above relations are $\mathcal{O}(\widehat{m}^2)$. Masses with a hat refer to the renormalized and $\mathcal{O}(a)$ improved quantities [118],

$$\widehat{m}_j = \frac{Z_A}{Z_P}\left\{1 + a\left[3\left(\tilde{b}_A - \tilde{b}_P\right)\overline{m} + \left(b_A - b_P\right)m_j\right]\right\}\widetilde{m}_j, \tag{3.5}$$

where Z_A and $Z_P(\mu)$ are renormalization factors of which the latter depends on the renormalization scale, $b_A, \tilde{b}_A, b_P$ and $\tilde{b}_P$ are improvement coefficients, see sec. 2.4.2, and $\widetilde{m}$ refers to the masses defined through the non-singlet partially conserved axialvector current (PCAC) relation of the corresponding flavour structures, e.g.,

$$2\widetilde{m}_\ell = \frac{\partial_0\langle\Omega|\tilde{A}_0^1|\mathcal{P}^1\rangle}{\langle\Omega|\tilde{P}^1|\mathcal{P}^1\rangle}, \qquad \widetilde{m}_\ell + \widetilde{m}_s = \frac{\partial_0\langle\Omega|\tilde{A}_0^4|\mathcal{P}^4\rangle}{\langle\Omega|\tilde{P}^4|\mathcal{P}^4\rangle}, \tag{3.6}$$

where the notation follows that of eqs. (2.79), (2.111) and (2.112). Since we employ Wilson fermions, we use these relations to define the chiral point where both the renormalized and the bare quark masses vanish, i.e., at this point $\kappa = \kappa_{\text{cr}}$ as defined in eq. (2.121).

One of our trajectories is defined along a line of constant average (bare) quark mass,

$$a\,\text{tr}\,M = 3a\,\overline{m} = \text{const}, \tag{3.7}$$

which is equivalent to keeping the sum of inverse hopping parameters constant at each value of β. In this particular setting, eq. (2.108) is constant along the complete trajectory since the lattice effects to the coupling depend only on the average unrenormalized mass and so all $\mathcal{O}(a)$ effects are constant along this trajectory [137]. Also the lattice spacing effects to the renormalized masses

[1] Note that in the more recent editions of the FLAG report no such numbers are given.

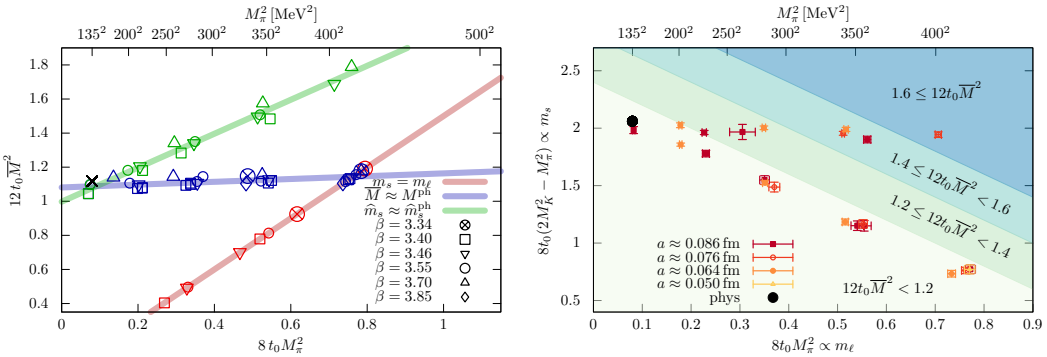


FIGURE 3.1: Simulated ensembles at several lattice spacings and along the simulated quark mass trajectories. **Left panel**: The horizontal axis is proportional to the light quark mass while the vertical axis corresponds to the average quark mass. The line of constant average quark mass (blue) meets with the line of physical strange quark mass (green) in the physical point (black cross) and with the $N_f = 3$ line at the flavour symmetric point that corresponds roughly to $\overline{M}^{\rm ph}$. **Right panel**: The subset of ensembles analysed in the following chapters for the η and η' results. In this case the axes correspond to combinations of pion and kaon masses that are approximately proportional to the light and strange quark masses via eq. (3.4). Again, the two trajectories lead to approximately the physical point (black dot) and the colour shades correspond to similar average quark masses upon which ensembles will be excluded when quantifying cut-off effects of the chiral expansion, cf. sec. 5.3.

$\widehat{\overline{m}}$ simplify and only the presumably small singlet contributions play a role [118]. This means that continuum and chiral extrapolations along this line can be performed more easily, because the leading lattice spacing effects do hardly depend on the sea quark masses. Using this choice, only one parameter requires tuning, namely tr M so that it corresponds to the physical point up to constant (but unknown) lattice spacing effects. As a consequence of this choice of the trajectory, the kaon mass is lighter than its physical value whenever $M_\pi > M_\pi^{\rm ph}$.

Another popular choice is to keep the renormalized strange quark mass close to its physical value, $\widehat{m}_s \approx \widehat{m}_s^{\rm ph}$. This leads to the aforementioned $\mathcal{O}(am)$ effects to both the coupling (or equivalently to the lattice spacing, see sec. 2.4.1) and the renormalized quark masses when varying the light quark mass. For this reason it is more difficult to hit the physical point and considerable mistuning can be observed in the literature, see, for example, the scatter of points in fig. 17 of [138]. In [89] we analysed the CLS ensembles that already existed along the tr M = const line and additional ensembles along $m_s = m_\ell$ to parametrize the relation between the renormalized quark masses and the hopping parameters. This allowed us to predict bare quark masses and the corresponding hopping parameters for the simulation, leading to a parametrization of $\kappa_s(\beta, \kappa_\ell)$ that leads to an almost constant renormalized strange quark mass. The right panel of fig. 3.1 shows this for the pseudoscalar mass combination that, according to eq. (3.4), corresponds to the strange quark masses of the ensembles analysed in the subsequent chapters. Ideally, this trajectory goes to the physical point and there meets with the tr M = const line.

In principle, there is also another interesting trajectory, both for determining renormalization factors and other quantities that are defined in the chiral limit and to study SU(3) flavour symmetry, namely that where the strange quark mass is set to the light quark mass, $m_s = m_\ell$ (which trivially also implies $\widehat{m}_s = \widehat{m}_\ell$) and, hence, $\delta M^2 = 0$. This quark mass configuration is also interesting for the analysis of η and η' physics, since there the η and η' states are pure octet and singlet states, respectively, see the discussion in secs. 2.1.6 and 2.3.5. This $N_f = 3$ trajectory, however, does not lead to the physical point. The line is fully parametrized by $\overline{M}^2$ and intersects with the tr M = const trajectory at the point $\overline{M} \approx \overline{M}^{\rm ph}$. Such simulations are only possible at comparably large quark masses and thus far away from the physical point, since the odd number of quarks requires to take the root of the quark determinant as discussed in sec. 2.2.4. Towards lighter quark masses the spectral gap cannot be guaranteed.

If there was no quark mass mistuning on the tr M = const or the $\widehat{m}_s = \widehat{m}_s^{\text{ph}}$ lines, extrapolations to the physical point on just one of the two are relatively straight-forward and require only one argument that parametrizes the "distance" to the physical point (for example δM along the tr M = const and M_π along the $\widehat{m}_s = \widehat{m}_s^{\text{ph}}$ trajectory). The continuum limit requires a parametrization of lattice spacing effects with additional terms. In the case of the tr M = const trajectory, these are relatively simple to $\mathcal{O}(a)$, but along the $\widehat{m}_s = \widehat{m}_s^{\text{ph}}$ line additional $a\overline{m}$ terms are required, see also sec. 2.4. Since the a posteriori determination of the physical point is likely to be more precise than the a priori estimate when choosing the simulation parameters, the physical point is usually not reached very precisely and possible quark mass mistuning must be compensated. One possibility is to shift observables slightly by expanding around their masses [135] or by quark mass reweighting [139, 140]. To incorporate more than one trajectory, the use of functional forms inspired by SU(3) ChPT is the obvious choice and takes care of possible quark mistuning automatically. Such continuum parametrizations for $N_f = 2+1$ flavours depend on two arguments that describe the complete quark mass plane, without making assumptions on the precision of the quark mass tuning, although higher order terms may be necessary to maintain good accuracy over the full range. For this reason the requirement that trajectories should lead to the physical point can be relaxed but the statistical and systematic errors can be reduced by simulating ensembles close to the physical point. This is the approach that we will pursue in ch. 5.

3.2 Ensemble overview

Along these three mass trajectories many ensembles have been generated to date, cf. fig. 3.2. Currently, these cover six lattice couplings $\beta \in \{3.34, 3.4, 3.46, 3.55, 3.7, 3.85\}$ that correspond to lattice spacings from $a \approx 0.098$ fm at $\beta = 3.34$ to $a \approx 0.039$ fm at $\beta = 3.85$, see tab. 3.3 for the individual values and sec. 3.4 for their determination. If all linear cut-off effects are under control, e.g., when using fully $\mathcal{O}(a)$ improved operators, the extrapolation can be performed in a^2. In this case the range of the extrapolation covers a factor of $a^2(\beta = 3.34)/a^2(\beta = 3.85) \approx 6.3$ and allows for precise continuum limit extrapolations, see the right panel of fig. 3.7 for an illustrative example. The smallest lattice constants are only possible to simulate with open boundary conditions that are introduced in the subsequent section, while for some other lattices at coarser lattice spacing (anti-)periodic boundary conditions were used.

The heaviest simulated pion masses on the two main quark mass trajectories, tr M = const and $\widehat{m}_s \approx \widehat{m}_s^{\text{ph}}$, are around $M_\pi \approx \overline{M}^{\text{ph}} \approx 411\,\text{MeV}$. The trajectory with constant average quark mass at this point meets with the $\widehat{m}_s = \widehat{m}_\ell$ line and correspondingly at this point also $M_K = M_\pi = \overline{M}$, see fig. 3.1. At $\beta = 3.4$ and $\beta = 3.55$, there exist simulation points that are very close to the physical point, but many more light ensembles in the range $200\,\text{MeV} < M_\pi < 300\,\text{MeV}$ are available that are particularly important to perform the chiral extrapolation of results in a controlled way.

The simulation volumes are chosen such that typically $L_s M_\pi > 4$ and $L_s > 2$ fm to ensure that finite volume effects are small [141]. For most quark mass points at the tr M = const line at $\beta = 3.4$, additional smaller volumes have been simulated to enable the study of these volume effects. The requirement of simulating at large $L_s M_\pi$ limits approaching the physical mass point at small lattice spacings: At the physical point at $\beta = 3.55$, E250, already a volume of $L_t \times L_s^3 = 96\,a^4$ is needed which is at the boundary of what can be achieved with present high performance computing systems, see [134] for some of the difficulties that arise with such simulations.

While the many available ensembles map out the physically relevant parameter space very well, the combination of very small lattice spacings and small pion masses is still very expensive to achieve and further algorithmic and hardware developments are needed. Nevertheless, the many available data points allow for an unprecedented accuracy in controlling the systematic errors.

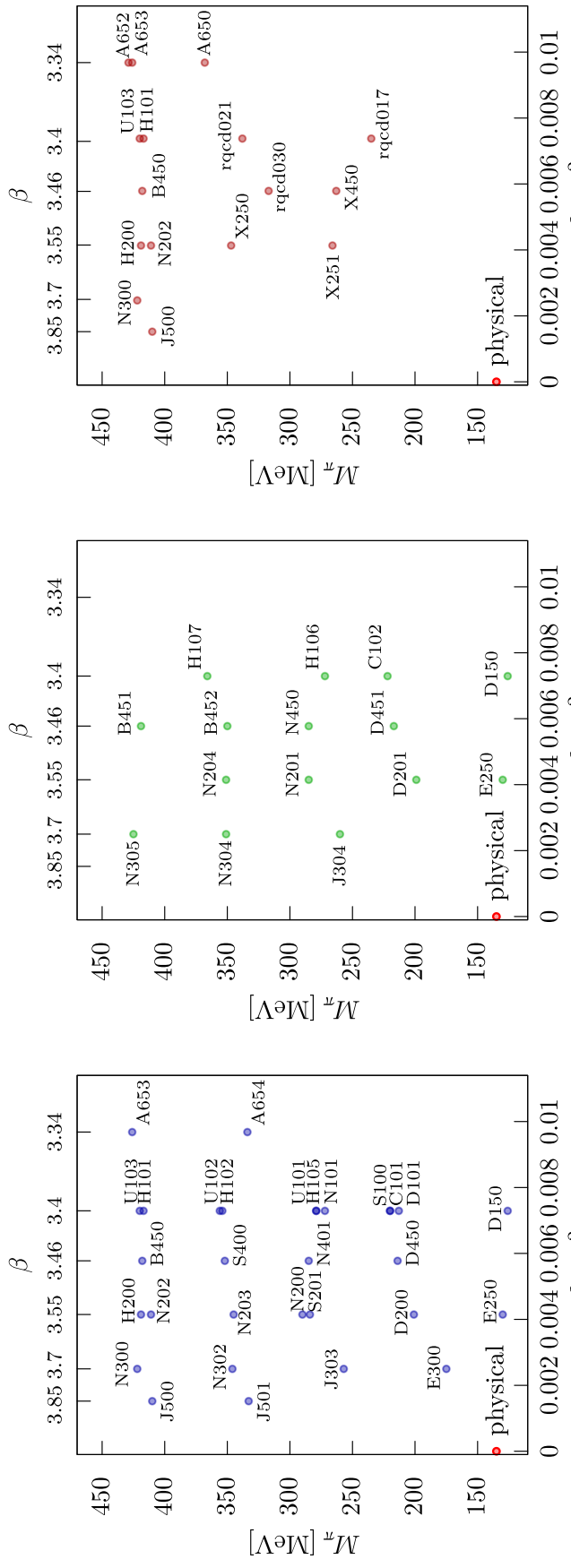


FIGURE 3.2: Ensembles on each of the three trajectories, tr M = const (left), $\hat{m}_s \approx \hat{m}_s^{\text{ph}}$ (center) and $m_s = m_\ell$ (right), and their pion masses (vertical axes) and lattice spacings (horizontal axes). Some points appear on more than one trajectory since these intersect (approximately) at the physical and the symmetric points, cf. sec. 3.1.

For the analysis of the η and η' mesons in this work, we mainly analyse 20 large volume ensembles at three different values of the lattice coupling, $\beta \in \{3.4, 3.46, 3.55\}$, complemented by a finer $N_f = 3$ ensemble at $\beta = 3.7$. We list all these ensembles, their pion and kaon masses, the Wilson flow scale and combinations thereof in tab. 3.1. Both the tr M = const as well as the $\widetilde{m}_s \approx \widetilde{m}_s^{\mathrm{ph}}$ trajectories are employed to improve the control over systematic errors that arise in the chiral extrapolation. At $\beta = 3.4$ we include the ensemble D150 which is very close to the physical point.

3.3 Autocorrelations and open boundary conditions

In (hybrid) Monte-Carlo simulations, each accepted trajectory is the starting point for the next. The result is that measurements are not independent of another and configurations are correlated in Monte-Carlo time. This needs to be taken into account when analyzing lattice data, cf. app. B. With finer and finer lattice spacings, however, autocorrelation times become increasingly long. For the RHMC algorithm, these autocorrelation times grow approximately with the inverse of the squared lattice spacing. With periodic boundary conditions, it has been observed that the topological charge eventually even freezes and the Monte-Carlo history is no longer ergodic, see, e.g. [142]. This phenomenon became known as "critical slowing down" and prevents the simulation of fine lattice spacings $a \lesssim 0.05$ fm.

It has been found that open boundary conditions in time [143] improve the situation by letting topological charge flow in and out through the temporal direction. Spatial directions remain periodic to allow for volume averaging and leaving the Fourier transform in these directions intact. In time direction, however, the boundaries are fixed and the temporal component of the field strength tensor is set to zero at the boundaries,

$$\widehat{F}_{0k}(0, \vec{x}) = \widehat{F}_{0k}(T, \vec{x}) = 0, \qquad k = 1, 2, 3, \tag{3.8}$$

where $T = a(N_t - 1)$ is the lattice extent in time direction and $\widehat{F}_{\mu\nu}$ is defined in eq. (2.63). Open boundaries can also be imposed by fixing the gauge fields pointing outside of the lattice at the last time slice to zero, while keeping the periodicity

$$U_0(T, \vec{x}) = U_0^\dagger(0, \vec{x}) = 0, \tag{3.9}$$

which automatically also leads to the correct boundary conditions of the fermion fields in the Dirac operator,

$$(1 + \gamma_0)\psi(0, \vec{x}) = (1 - \gamma_0)\psi(T, \vec{x}) = 0, \tag{3.10}$$

see, for example, the hopping term, eq. (2.65).

When computing observables, a distance to the boundary must be kept since the topological effects extend over a range of typically $b \lesssim 2$ fm [90, 133] in the vicinity of the boundaries. This effectively decreases the available volume and increases the computational cost to achieve the same error by a small linear factor compared to periodic boundaries. At small values of the lattice spacing, however, the simulation remains ergodic and this additional cost is compensated by decreased autocorrelation times.

Fig. 3.3 shows the action density computed on the central time slices of the gauge fields after applying the gradient flow [144] discussed in the next section,

$$a^4 E(t) = \frac{1}{(N_0 - 2b/a)\, N_s^3} \sum_{x_0 = b/a}^{N_0 - b/a - 1} \sum_{\vec{x}} \frac{1}{4} \widehat{F}^a_{\mu\nu}(x) \widehat{F}^a_{\mu\nu}(x), \tag{3.11}$$

β	id	BC	trajectory	$L_t \times L_s^3/a^4$	M_π/MeV	M_K/MeV	$L_s M_\pi$	$12t_0\overline{M}^2$	$8t_0\delta M^2$	t_0/a^2
3.4	H101	O	trM	96×32^3	420	420	5.85	1.158(16)	0	2.854(8)
	H102a	O	trM	96×32^3	352	442	4.90	1.116(16)	0.611(14)	2.884(9)
	H102b	O	trM	96×32^3	356	441	4.96	1.130(18)	0.595(11)	2.879(9)
	H105	O	trM	96×32^3	279	465	3.89	1.125(16)	1.199(12)	2.892(6)
	C101	O	trM	96×48^3	220	472	4.60	1.120(12)	1.548(14)	2.918(4)
	D150	P	trM/ms	128×64^3	126	478	3.51	1.074(15)	1.901(28)	2.948(3)
	H107	O	ms	96×32^3	366	545	5.10	1.511(12)	1.340(12)	2.719(8)
	H106	O	ms	96×32^3	272	516	3.79	1.289(20)	1.662(50)	2.823(7)
	C102	O	ms	96×48^3	222	500	4.64	1.208(8)	1.736(11)	2.868(5)
3.46	B450	P	trM	64×32^3	418	418	5.16	1.143(12)	0	3.663(11)
	S400	O	trM	128×32^3	352	442	4.35	1.136(16)	0.615(13)	3.692(7)
	N401	O	trM	128×48^3	285	461	5.28	1.114(18)	1.117(21)	3.684(5)
	B451	P	ms	64×32^3	418	572	5.16	1.678(10)	1.238(9)	3.426(7)
	B452	P	ms	64×32^3	350	544	4.32	1.489(8)	1.444(12)	3.529(7)
3.55	N202	O	trM	128×48^3	411	411	6.43	1.100(12)	0	5.165(14)
	N203	O	trM	128×48^3	345	442	5.40	1.108(11)	0.668(6)	5.146(6)
	N200	O	trM	128×48^3	284	462	4.44	1.114(7)	1.174(9)	5.160(7)
	D200	O	trM	128×64^3	201	480	4.19	1.107(6)	1.678(8)	5.179(4)
	N204	O	ms	128×48^3	351	544	5.49	1.512(8)	1.472(6)	4.947(8)
	N201	O	ms	128×48^3	285	523	4.46	1.351(7)	1.654(9)	5.043(8)
	D201	O	ms	128×64^3	199	500	4.15	1.191(10)	1.847(18)	5.138(7)
3.7	N300	O	trM	128×48^3	422	422	5.10	1.156(16)	0	8.576(21)

TABLE 3.1: Details of the main CLS ensembles analysed in chapters 4 to 6. For the corresponding information of the remaining ensembles, we refer to [89, 90, 133]. Both open ("O") and (anti-)periodic ("P") boundary conditions in time are employed. The pion and kaon masses as well as the gradient flow scale t_0/a^2 are taken from [90], while $12t_0\overline{M}^2$ and $8t_0\delta M^2$ have been determined on the same configurations as is employed in the remainder of this thesis. Ensembles H102a and H102b were generated with the same quark masses and lattice coupling but different simulation parameters and are therefore analysed separately.

where in this case t is the gradient flow time and x_0 the time coordinate. The temporal sum excludes the b/a timeslices closest to the boundaries so that $aN_0 - 2b \approx 1\,\text{fm}$. The components of the lattice field strength tensor are given by

$$\widehat{F}^a_{\mu\nu}(x) = 2i\,\text{tr}\left[\widehat{F}_{\mu\nu}(x)t^a\right]. \tag{3.12}$$

The action density $E(t = t_0)$ fluctuates slowly in Monte-Carlo time and shows long autocorrelation times for all simulated values of the lattice coupling. In fig. 3.3 all except the coarsest ensemble employ open boundary conditions and although the autocorrelation times clearly increase towards fine lattice spacings, the simulation is not stuck and clearly proceeds normally even at the finest lattice spacing. In fig. 3.4 the computed integrated autocorrelation times of [90] are plotted for the symmetric point ensembles and show the expected scaling.

3.4 Setting the scale using the Wilson flow time t_0

The gradient flow time t_0 is defined as the point where the dimensionless combination of the flowed action density, eq. (3.11), assumes the value

$$t^2E(t)|_{t=t_0} = 0.3, \tag{3.13}$$

see [144] for details of the flow time definition of gauge fields. A value for the ratio t_0/a^2 can be obtained on every ensemble by computing $t^2E(t)$ for a set of t/a^2 values and interpolating to the point defined in eq. (3.13). The advantage of the gradient flow scale t_0 is that it is free from renormalization [144], depends only mildly on sea quark masses [145] and can be computed with only moderate computational cost to a very high precision. These properties allow its use in scale setting, i.e., comparing the lattice spacings between several lattice simulations. The connection to experiment and the physical world requires additional input to determine a value for $t_0^{\text{ph}} = t_0(\overline{M}^{\text{ph}}, \delta M^{\text{ph}}, a = 0)$. Values for t_0^{ph} with $N_f = 2 + 1$ flavours have been determined in [135, 146–149]. These determinations are in rough agreement, although the newer determination on CLS ensembles [149] has shifted downwards by about one standard error in comparison to [135] where only a subset of ensembles was analysed. Both use the non-singlet light pseudoscalar decay constants in the isospin symmetric limit as physical inputs, subtracting electromagnetic effects similarly as for eq. (3.1). Systematic errors from these inputs and from the chiral extrapolation dominate over the statistical errors. When we need to convert our results to physical units in this work we use the value of [135]

$$\sqrt{8t_0^{\text{ph}}} = 0.415(4)\,\text{fm}, \qquad \text{corresponding to} \quad (8t_0^{\text{ph}})^{-1/2} = 475(6)\,\text{MeV}, \tag{3.14}$$

but give results also in dimensionless combinations with t_0^{ph}/a^2 so that results can be converted once a value with increased input precision from, e.g., the octet cascade baryon mass becomes available [90].

The quark mass dependence of t_0 has been computed in ChPT to next-to-next-to leading order (NNLO) in [145]. While at NNLO the continuum expression depends also on the difference of quark masses and the energy scale of the effective field theory, the NLO quark dependence is very simple,

$$t_0(\overline{M}, \delta M) = t_0^\chi\left(1 + k_1\frac{3\overline{M}^2}{(4\pi F)^2}\right) + \mathcal{O}(M^4). \tag{3.15}$$

This expression depends only on two constants, the flow scale in the chiral limit $t_0^\chi = t_0(\overline{M} = 0, \delta M = 0)$ and the slope parameter $3k_1/(4\pi F)^2$, where F is the non-singlet pseudoscalar decay constant in the chiral limit. We redefine the slope parameter to $\tilde{k}_1 = 3k_1/(8t_0(4\pi F)^2)$ and get

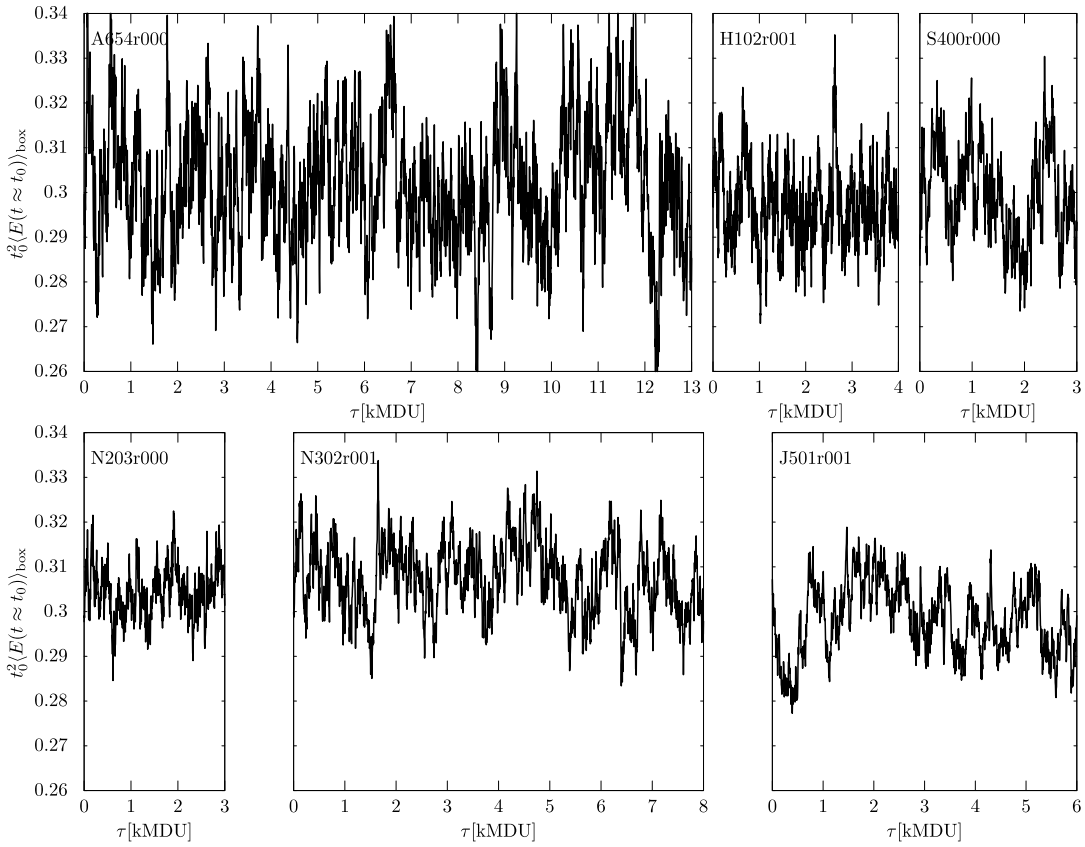


FIGURE 3.3: History of the Wilson flow action density, multiplied by t_0^2 [144], after a flow time close to t_0, inside a central sub-volume of approximately 1 fm $\cdot\ (aN_s)^3$, along a line of $M_\pi \approx 340$ MeV (for tr M = const) from coarse to fine lattice spacings. The amplitude of the fluctuation varies, e.g., due to somewhat different physical volumes. Autocorrelations increase from top left to bottom right, with the exception of A654 at $\beta = 3.34$ where we observe larger autocorrelation times than at $\beta = 3.4$. For the cases where more than one Monte-Carlo chain exists, only one replica is shown.

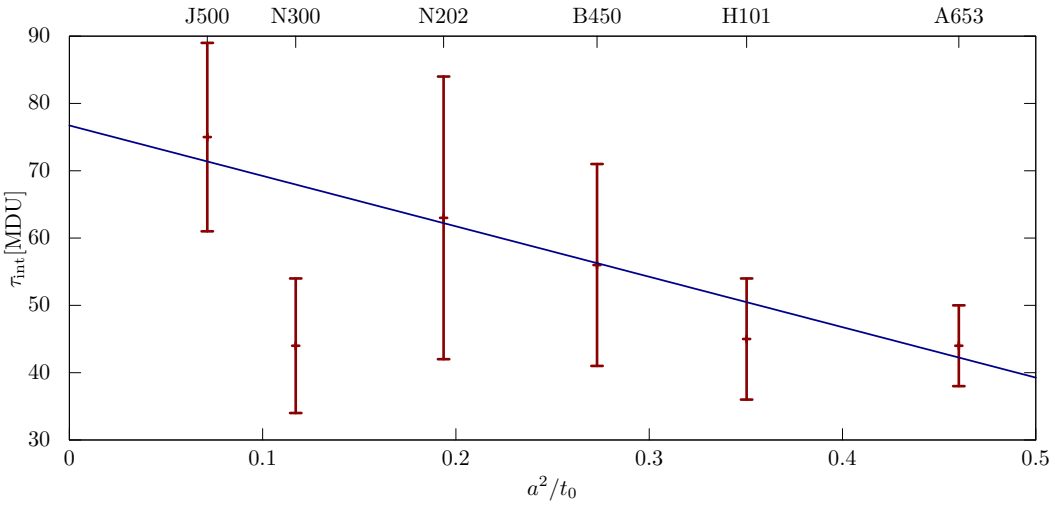


FIGURE 3.4: Integrated autocorrelation times τ_{int} of t_0/a^2 as a function of a^2/t_0 for the ensembles at the symmetric point. Data is taken from [90]. With the exception of N300, the autocorrelation times follow the expected quadratic scaling behaviour. The line represents a simple linear fit in a^2/t_0 to all points except N300. All except B450 and A643 employ open boundary conditions.

the NLO expression

$$t_0(\overline{M}, \delta M) = t_0^{\chi} \left(1 + \tilde{k}_1 \frac{2}{3}\phi_4\right), \tag{3.16}$$

which differs from eq. (3.15) only at NNLO. Here and in the following we conveniently use the dimensionless mass combinations

$$\phi_4 = 8t_0 \left(M_K^2 + \frac{1}{2}M_\pi^2\right) = 12 t_0 \overline{M}^2, \tag{3.17}$$

$$\phi_2 = 8t_0 M_\pi^2 = 8t_0 \left(\overline{M}^2 - \frac{1}{3}\delta M^2\right). \tag{3.18}$$

Since in eq. (3.15) t_0 depends only weakly on δM^2, an interesting limit is that of $N_f = 3$ mass degenerate flavours along which $M_K^2 = M_\pi^2$ and hence $\delta M^2 = 0$. Along this line and in the continuum limit we define

$$\phi_4^* = 8t_0^* \left(M_K^2 + \frac{M_\pi^2}{2}\right) = 12 t_0^* (\overline{M}^{\rm ph})^2 := 1.110, \tag{3.19}$$

which is a choice motivated by the previous determination [146] of $t_0^{\rm ph} \approx 0.413$ and the value of $\overline{M}^{\rm ph}$ as given in eq. (3.3): If δM^2 contributions to ϕ_4 are small in the continuum, then $\phi_4^* \approx \phi_4^{\rm ph}$. The values of ϕ_4 and t_0 in our symmetric point simulations at finite lattice spacings are close to the continuum ϕ_4^* but are typically larger than that value due to lattice spacing effects and mass mistuning and we refer to them as $\phi_4^{\rm symm}$ and $t_0^{\rm symm}$, respectively.

In the following, we explore the quark mass and lattice spacing dependence of t_0/a^2 on the simulated ensembles and extrapolate to the interesting continuum limit points $t_0^{\rm ph} = t_0(\overline{M}^{\rm ph}, \delta M^{\rm ph})$, $t_0^* = t_0(\overline{M}^{\rm ph}, 0)$ and $t_0^{\chi} = t_0(0,0)$. This allows us to assign a value of t_0^*/a^2 for each lattice coupling that we simulate and we can use this value to convert to physical units using external input for t_0^* (or $t_0^{\rm ph}$).

3.4.1 Strategy for continuum limit extrapolations

As described in sec. 2.4.1, the lattice spacing itself is subject to linear cut-off effects when keeping the coupling fixed. This induces linear cut-off effects on dimensionless combinations with the lattice spacing even if the observables themselves are otherwise subject to only quadratic lattice spacing effects. Here, we once again summarize the strategy for taking the continuum limit employing the Wilson flow scale as the auxilliary variable, cf. (2.110). When determining hadron masses, e.g., we expect to see cut-off effects

$$a\, M(\overline{M}, \delta M, a^2) = a_0 \left(1 + b_a a\overline{m}\right)\, M(\overline{M}, \delta M, 0) + \mathcal{O}(a^2), \tag{3.20}$$

see eq. (2.108) for the definitions of a_0 and b_a. Extrapolations of such combinations to the continuum limit would thus have to take this into account and parameterize these linear cut-off effects. The same is true for the determination of the gradient flow scale,

$$\left[\frac{t_0}{a^2}\right](\overline{M}, \delta M, a^2) = \frac{t_0(\overline{M}, \delta M, 0)}{a_0^2}\left(1 - 2b_a a\overline{m}\right) + \mathcal{O}(a^2), \tag{3.21}$$

which we will extrapolate to the continuum limit taking care of $a\overline{m}$ terms in sec. 3.4.3.

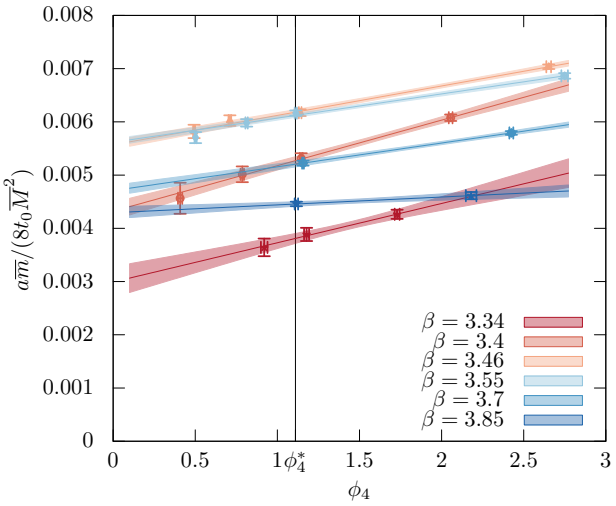


FIGURE 3.5: The ratio of the average lattice quark mass in lattice units over the square of the average pseudoscalar mass as fitted in eq. (3.23). The ratio varies only by 20%. This makes it impractical to discriminate between terms proportional to $a\overline{m}$ and terms $\propto 8\,t_0\overline{M}^2$.

In general, however, it is advantageous to take dimensionless combinations of the quantities of interest, for example masses or decay constants, with the gradient flow scale and extrapolate

$$\left[a\,M(\overline{M},\delta M,a^2)\right]\sqrt{t_0a^2(\overline{M},\delta M,a^2)/a^2} = M(\overline{M},\delta M,a^2)\sqrt{t_0(\overline{M},\delta M,a^2)} + \mathcal{O}(a^2), \tag{3.22}$$

instead, since the linear $a\overline{m}$ contribution cancels from such combinations. It is thus beneficial to use the Wilson flow scale t_0/a^2 from the same ensemble — at the same lattice spacing and the same quark masses — as the observable itself instead of using the interpolated values t_0^*/a^2 or a value for a in physical units. In this case the result would contain linear cut-off effects that are proportional to the difference of the average quark masses, $a\overline{m} - a\overline{m}^*$. Typically the chiral extrapolation is performed at the same time and in this case the chiral expansion of t_0, eq. (3.15) must be taken into account. Once the physical quark mass point in the continuum is reached, the scale can then be set using a value for t_0^{ph} in physical units, eq. (3.14).

3.4.2 The symmetric point parameters ϕ_4^{symm}, κ^{symm}, κ^*

The dimensionless combination ϕ_4^* at the flavour symmetric line is defined in the continuum and such that it corresponds to the physical average quark mass. At any finite lattice spacing, however, cut-off effects contribute. The knowledge of these effects is important for the quark mass tuning of the tr M = const trajectories, since in order to cross the physical point, the trajectory should be started at $\phi_4^{\text{symm}} > \phi_4^*$. Here we determine values of ϕ_4^{symm} and the corresponding hopping parameter κ^{symm} such that in the continuum they lead to $\phi_4 = \phi_4^* := 1.11$, see eq. (3.19). At every value of β we parametrize the dependence of ϕ_4 on the hopping parameter along the $m_s = m_\ell$ line using the κ_{crit} values of [90] as inputs

$$a\overline{m}(\kappa) = am(\kappa) = p_1\phi_4 + p_2\phi_4^2, \tag{3.23}$$

see eq. (2.121) for the relation of the hopping parameter to the bare quark masses. At the point $\phi_4 = \phi_4^*$ we obtain a value for am^*, defining κ^*. The data and the fits are displayed in fig. 3.5, and results are collected in table 3.2, where we also quote the critical hopping parameter values determined in [90].

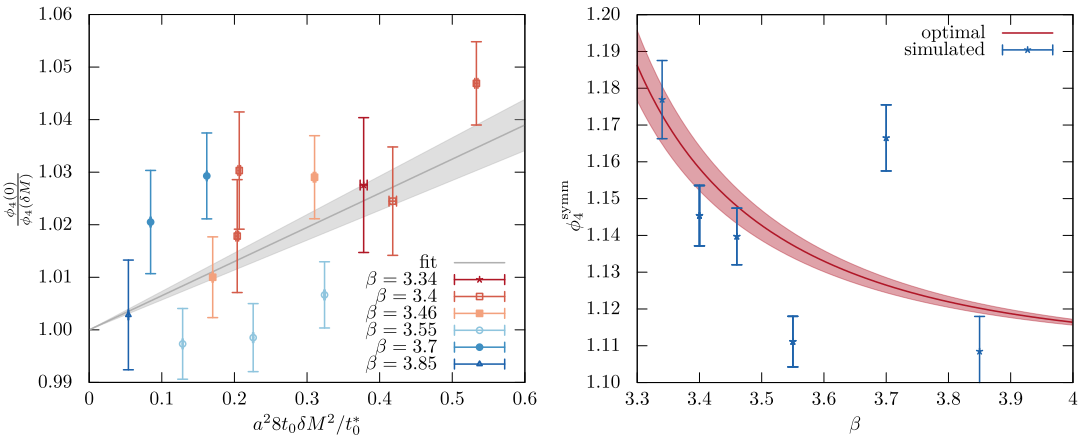


FIGURE 3.6: **Left panel**: fit to eq. (3.24) of the deviation of the ratio $\phi_4(0)/\phi_4(\delta M)$ from one, along the tr M = const trajectories. **Right panel**: optimal starting values of ϕ_4^{symm} at the symmetric point that should be selected in order to obtain $\phi_4\text{ph} = 1.11$ at the physical point, along with the actually simulated values of ϕ_4.

Starting from the symmetric point and approaching the physical point, the quark mass difference proportional to δM grows and there are associated lattice spacing effects that complicate the mass tuning. To resolve such δM dependent lattice spacing artifacts, we now take ratios of the symmetric point ensembles where $\delta M = 0$ and large volume ensembles along the tr M = const lines, where $\overline{M}^2$ = const and thus all quark mass dependent lattice spacing effects must be proportional to δM. To this order $\mathcal{O}(M^2)$ in ChPT, there are no such δM^2 terms in the continuum [145] and so we perform a simple linear fit to

$$\frac{\phi_4(0,a)}{\phi_4(\delta M,a)} = 1 + \delta c_\phi \frac{a^2}{t_0^*} 8t_0 \delta M^2. \tag{3.24}$$

All twelve ratios $\phi_4(0,a)/\phi_4(\delta M,a)$ that we have at our disposal for the large volume ensembles are roughly described by the parameter

$$\delta c_\phi = 0.065(13), \quad \chi^2/N_{\text{DF}} = 28/12. \tag{3.25}$$

This fit is shown in the left panel of fig. 3.6. While the fit describes the overall trend of the points as a function of δM^2 well, the poor value of χ^2 may be due to missing parametrizations of significant continuum limit terms that are proportional to δM^2 but of higher order. To NNLO, for example, $\overline{M}^2 \delta M^2$ contributes (next to other mass combinations and chiral logarithms). Within the tr M = const data, however, $\overline{M}^2$ is constant and the term can hardly be distinguished from $a^2 \delta M^2$. These terms would be separable along the $\widehat{m}_s = \widehat{m}_s^{\text{ph}}$ trajectory; however, there, the ratio gives no advantage and discretization terms like $a^2 \overline{M}^2$ need to be included in the parametrization. Given these complications, we here stick to this very simple one-parameter fit and keep in mind that the results will likely be only indicative of possible mistuning rather than precise predictions of necessary shifts in $\phi_4^{\text{symm}} - \phi_4$ for every lattice spacing.

Using eq. (3.24) and the physical point estimates $\phi_4(\delta M^{\text{ph}}) = \phi_4^{\text{ph}} = 1.110$ and $8t_0 \delta M^{\text{ph}\,2} = 1.981$, cf. eq. (3.1), we obtain predictions for $\phi_4^{\text{symm}} = \phi_4(\delta M = 0)$ at finite lattice spacing

$$\phi_4^{\text{symm}}(t_0^*/a^2) = 1.11 \left(1 + \delta c_\phi \frac{a^2}{t_0^*} 1.981\right), \tag{3.26}$$

where a value for $t_0^*/a^2(\beta)$ is required that we obtain in the following subsections. Using these

β	κ_{crit} [90]	κ^*	am^*	$\phi_{4,\text{symm}}$	κ_{symm}
3.34	0.1366953(51)	0.1365791(60)	0.00311(8)	1.175(8)	0.1365716(14)
3.4	0.1369160(12)	0.1367645(12)	0.00405(4)	1.159(6)	0.1367578(10)
3.46	0.1370616(10)	0.1368947(13)	0.00445(3)	1.149(5)	0.1368872(4)
3.55	0.1371714(19)	0.1370012(21)	0.00453(4)	1.138(3)	0.1369968(3)
3.7	0.1371530(9)	0.1370081(18)	0.00385(4)	1.127(2)	0.1370058(3)
3.85	0.1369771(26)	0.1368518(39)	0.00334(8)	1.120(1)	0.1368507(3)

TABLE 3.2: Results for the critical hopping parameter κ_{crit} and for the hopping parameter at the $N_f = 3$ symmetric point, where $\phi_4 = \phi_4^* = 1.11$, κ^*, together with the bare quark mass at this point, $am^* = (\kappa^{*-1} - \kappa_{\text{crit}}^{-1})/2$ and the corresponding AWI quark mass $a\widetilde{m}^*$. At the physical quark mass point $\phi_{4,\text{ph}} = 1.110$, however, lattice spacing effects shift this value when reducing δM. ϕ_4^{symm} denotes the $m_s = m_\ell$ starting point of the tr $M = \text{const}$ trajectory that will go through the physical point and κ_{symm} the corresponding value of the κ parameter. The actual simulations were performed at slightly different values, see the right panel of fig. 3.6 and [90, 133, 150].

values and its effective parametrization, eq. (3.35), the above equation can be converted into a function of β and we compare these postdicted values to the actually simulated symmetric ensembles in the right panel of fig. 3.6. Interpreting our fit with some caution as discussed above, at $\beta = 3.55$ we should ideally have started at a smaller κ_{symm} value while at $\beta = 3.7$ a larger value would have shifted our trajectory closer to the physical line. Since we employ two independent quark mass plane trajectories, we need to use a parametrization of the mass dependence over the full quark mass plane, and precision of the quark mass tuning is not too important and can be compensated in the extrapolation.

Combining the interpolating formula (3.23) for am with the above result, we can also predict the hopping parameter value that corresponds to this optimal symmetric point, κ_{symm}:

$$\kappa_{\text{symm}} = \left[2\left(p_1\phi_4^{\text{symm}}(t_0^*/a^2) + p_2\left[\phi_4^{\text{symm}}(t_0^*/a^2)\right]^2\right) + \frac{1}{\kappa_{\text{crit}}}\right]^{-1}. \tag{3.27}$$

The values of $\phi_{4,\text{symm}}$ and of κ_{symm} are shown in table 3.2.

3.4.3 Lattice spacing dependence of t_0/a^2

We start to explore the lattice spacing dependence of the scale parameter along the $m_s = m_\ell$ line (including ensembles away from $\overline{M} = (\overline{M}^{\text{ph}})$), keeping $\delta M = 0$ fixed. In this case we expect both linear lattice spacing effects from keeping the coupling fixed, parametrized by b_a as described in secs. 2.4.1 and 3.4.1, as well as quadratic and higher effects that will alter both the slope parameter of eq. (3.16) as well as the chiral limit,

$$\frac{t_0}{a^2} \sim \frac{t_0^\chi}{a_0^2}\left[1 + \tilde{k}(a)\, 8t_0(\overline{M}, 0)\overline{M}^2\right], \tag{3.28}$$

with a different set of parameters t_0^χ/a_0^2 and $\tilde{k}(a)$ for each value of the coupling. The lattice spacing in the chiral limit is denoted by $a_0 = a(g^2, a\overline{m} = 0)$, cf. eq. (2.108). We first perform these two-parameter fits to each value of β separately and they effectively describe our data well as demonstrated in fig. 3.7 (local fit). We can then use this parametrization and set $8t_0\overline{M}^2 = \frac{2}{3}\phi_4^* \approx 0.74$ to obtain a value for t_0^*/a^2 which we list in the first line of table 3.3.

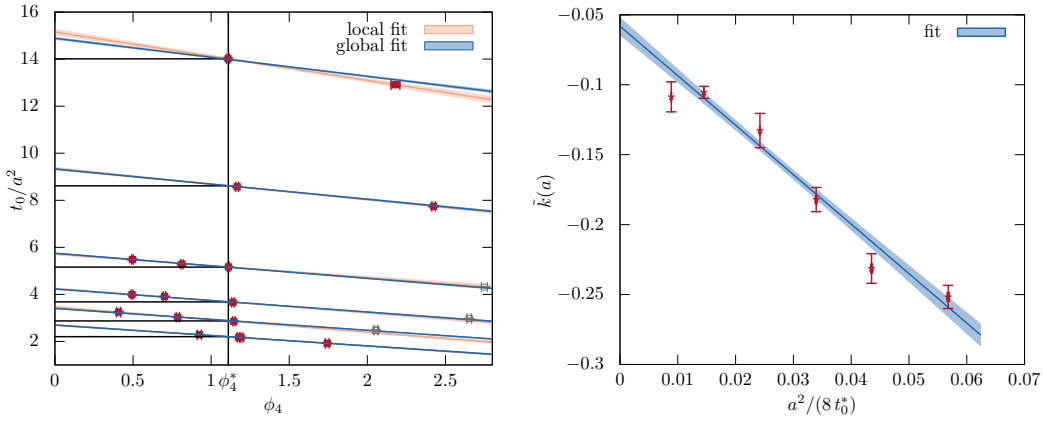


FIGURE 3.7: **Left panel**: Linear interpolation of t_0/a^2 for each value of the coupling, from $\beta = 3.34$ (bottom) to $\beta = 3.85$ (top), cf. eq. (3.28) (local fit). Also shown is the result of the global fit eqs. (3.35), (3.36), (3.39) and (3.40). The vertical line marks the value of the ϕ_4^* reference point and horizontal lines the resulting values of t_0^*/a^2. **Right panel**: The slope $\tilde{k}$, as a function of $a^2/(8t_0^*)$, together with a quadratic continuum limit extrapolation.

β	3.34	3.4	3.46	3.55	3.7	3.85
t_0^*/a^2, local fit	2.204(5)	2.872(10)	3.682(12)	5.162(16)	8.613(25)	14.011(39)
t_0^*/a^2, global fit	2.204(4)(4)	2.888(4)(7)	3.686(4)(10)	5.157(5)(14)	8.617(7)(21)	13.988(19)(28)
t_0^*/a^2 [63, 135]	—	2.862(5)	3.662(12)	5.166(15)	8.596(27)	13.880(220)
t_0^χ/a_0^2, local fit	2.703(20)	3.465(23)	4.253(20)	5.721(48)	9.348(53)	15.151(123)
t_0^χ/a_0^2, global fit	2.695(13)(2)	3.402(11)(1)	4.228(10)(5)	5.749(12)(1)	9.329(27)(4)	14.885(57)(14)
a/fm	0.09836(13)	0.08592(12)	0.07605(11)	0.06430(10)	0.04974(6)	0.03904(5)

TABLE 3.3: Results for t_0^* in lattice units from this work (lines 1 and 2), in comparison to the respective results obtained from reweighting [135] (updated in ref. [63], line 3). We consider the global fit result of the second line as the most reliable one. Also shown is the value of t_0 in the chiral limit, t_0^χ. The first errors are statistical, the second errors reflect the uncertainty of the improvement coefficient b_a (that is related to b_g). In the last line we also list the lattice spacing that results from the second line, t_0^*/a^2 using the physical value t_0^{ph} from eq. (3.14) [135].

Along this trajectory only relatively few data points are available and some are at large values of $\overline{M}^2$, potentially inducing both large higher order lattice spacing effects and sizable NNLO ChPT corrections. Hence, to minimize such possible biases we exclude the heaviest points from the fit where more than three well-separated data points were available, i.e., at $\beta = 3.4$, 3.46 and 3.55. Naturally, this fit is most trustworthy in the vicinity of the available data and can be used to interpolate to t_0^* where the close-by symmetric point at t_0^{symm} dominates the interpolation. The results compare well with the previous determination [135], updated in [63], that we show in the third row of table 3.3. At the same time, these fits are often constrained by data points at large quark masses and thus the validity of this extrapolation towards the chiral limit is questionable. Both the resulting values for t_0^χ/a_0^2 and the slopes $\tilde{k}(a)$ are listed in tab. 3.3 and we plot the latter in the right panel of fig. 3.7.

We then explore the lattice spacing effects to the continuum limit slope $\tilde{k}_1$ of eq. (3.16) and relate our fit results to the expectation making the linear cut-off effects of eq. (2.108) explicit,

$$\tilde{k}(a) = \tilde{k}_1 - \frac{2b_a a\overline{m}}{8t_0\overline{M}^2} + \mathcal{O}(a^2). \tag{3.29}$$

However, from our data shown in the right panel of fig. 3.7 we are unable to detect any linear contribution. Instead, a clearly quadratic dependence on a^2 is visible. Adding any higher power of a to the quadratic continuum limit extrapolation results in a coefficient that is compatible with zero. It should be noted that our failure to detect any linear dependence on a does not mean that the b_a coefficient itself is negligible but is due to the fact that within our range of lattice spacings the combination $a\overline{m}$ remains almost constant when keeping the average renormalized quark mass $\widehat{\overline{m}} = r_m Z_m m \propto \overline{M}^2$ fixed because the factor r_m decreases rapidly with β, see [90]. Therefore, within our errors, we cannot resolve any linear lattice spacing effect to t_0/a^2 and we will assume that b_a coincides with its one-loop value within a 100% error band: $b_a = 0.32(32)$. The naive quadratic continuum limit extrapolation gives

$$\tilde{k} = \tilde{k}(a = 0) = -0.0600(85). \tag{3.30}$$

To improve the extrapolation to the chiral limit and improve the data situation we now aim to incorporate data away from the symmetric $m_\ell = m_s$ line and combining all lattice couplings. To achieve this, we require an effective parametrization of lattice spacing effects also including terms $\mathcal{O}(a^2\delta M^2)$ and a function $\left[t_0^\chi/a_0^2\right](g^2)$ that describes the Wilson flow in the continuum and chiral limit as a function of the bare lattice coupling.

Interpolating formula for t_0^*/a^2

To build confidence in a combined global fit across several values of β, we start with parametrizing our existing ("local") results for $t_0^*/a^2(\beta)$.

For very small values of the coupling g^2 the dependence of t_0^*/a^2 on g^2 is controlled by the perturbative β function. Its three-loop coefficient β_2 is at present not known for our action. However, the ratio of Λ parameters,

$$\frac{\Lambda_{\mathrm{L}}}{\Lambda_{\overline{\mathrm{MS}}}} = 0.2887542, \tag{3.31}$$

was calculated in ref. [151]. Combining this with the recent determination [128]

$$\Lambda_{\overline{\mathrm{MS}}}\sqrt{8t_0^*} = 0.712(24), \tag{3.32}$$

we obtain for three quark flavours

$$t_0^*\Lambda_L^2 = 0.00528(36). \tag{3.33}$$

Integrating the β-function, we obtain for the running

$$h(g^2) := a(g^2)\Lambda_L = \exp\left[\frac{2\pi^2}{\beta_0 g^2} + \frac{2\beta_1}{\beta_0^2}\ln\frac{\beta_0 g^2}{4\pi^2}\right], \tag{3.34}$$

where the coefficients are given in sec. 2.1.3. Combining eqs. (3.33) and (3.34) and adding parametrizations of the leading lattice spacing corrections, we arrive at

$$\frac{t_0^*}{a^2(g^2)} = \frac{t_0^*}{a_0^2(g^2)}\left[1 - 2b_a(g^2)am^*\right] = \left[f(g^2) + c_{t_0} + d_{t_0}f^{-1/2}(g^2)\right]\left[1 - 2b_a(g^2)am^*\right], \tag{3.35}$$

where

$$f(g^2) = \frac{t_0^*\Lambda_{\mathrm{L}}^2}{h^2(g^2)} = 0.00528(36)\exp\left(-\frac{4\pi^2}{\beta_0 g^2} - \frac{\beta_1}{\beta_0^2}\ln\frac{-\beta_0 g^2}{4\pi^2} + b_{t_0}g^2 + \cdots\right). \tag{3.36}$$

The coefficient $b_{t_0} \approx (\beta_1^2 - \beta_0\beta_2)/(4\pi^2\beta_0^3)$ effectively parameterizes higher order perturbative contributions, c_{t_0} describes the leading $\mathcal{O}(a^2)$ lattice correction to eq. (3.32) and d_{t_0} a subleading $\mathcal{O}(a^3)$ correction. The resulting fit parameters read

$$c_{t0} = 0.18(12), \quad d_{t0} = -1.43(16), \quad b_{t_0} = -0.9293(46), \quad \chi^2/N_{\text{DF}} = 2.1/3. \tag{3.37}$$

Global interpolation of t_0/a^2

In a next step, we use the interpolating parametrization together with the continuum limit expectation, eq. (3.15), to incorporate all available ensembles in a global fit and obtain predictions of t_0/a^2 for the full range of couplings and quark masses that are available to us. This parametrization can then be used to predict values for $t_0^*/a^2(\beta)$ and its inverse can be used to predict lattice couplings for new simulations. We start from the continuum limit expectation eq. (3.16) and add effective parametrizations of the quark mass effects that are proportional to $a^2\overline{M}^2$ and $a^2\delta M^2$,

$$\frac{t_0}{a^2} \sim \frac{t_0^\chi}{a_0^2}(g^2)\left(1 + \tilde{k}\, 8t_0\overline{M}^2\right) + \bar{c}\, 8t_0\overline{M}^2 + \delta c\, 8t_0\delta M^2. \tag{3.38}$$

Shifting the chiral expansion around the symmetric point where we can use eq. (3.36) by substituting t_0^*/a^2 for t_0^χ/a_0^2 and $\tilde{k}_1$ for $\tilde{k}$ (eq. (3.29)), we obtain the following fit expression

$$\begin{aligned}\frac{t_0}{a^2}(\overline{M}^2, \delta M^2, a) &\sim \frac{t_0^*}{a^2}(g^2)\underbrace{\left[1 + \tilde{k}_1 A - 2b_a\left(a\overline{m} - am^*\right)\right]}_{\approx(1+\tilde{k}A)} + \bar{c}A + \delta c\, 8t_0\delta M^2 \\ &= \left[f(g^2) + c_{t_0} + d_{t_0} f^{-1/2}(g^2)\right]\left(1 + \tilde{k}_1 A - 2b_a a\overline{m}\right) + \bar{c}A + \delta c\, 8t_0\delta M^2,\end{aligned} \tag{3.39}$$

where

$$A = 8t_0\overline{M}^2 - \frac{2}{3}\phi_4^*, \tag{3.40}$$

and we used eqs. (3.35) and (3.36) to parameterize $t_0^*/a^2(g^2)$. The difference of the average quark mass, $a\overline{m}$, and the quark mass at the symmetric point of physical average pseudoscalar mass, am^*, is small. The parameters b_{t_0}, c_{t_0} and d_{t_0} are similar to those of eq. (3.35), and the total number of parameters for this combined fit across different values of g^2 is 6. Note that when inserting this parametrization the term that is proportional to $2b_a am^*$ cancels from the above equation but it resurfaces within the relation eq. (3.35) between t_0^*/a^2 and $f(g^2)$.

Again, the effect of b_a cannot be isolated within our range of lattice spacings (see the right panel of fig. 3.7) as $a\overline{m}$ is approximately proportional to $t_0\overline{M}^2$ and does not vary sufficiently across our data. This is shown in fig. 3.5, where the ratio as a function of ϕ_4 is plotted. The tr M = const ensembles can be found in the vicinity of the vertical $\phi_4 = \phi_4^*$ line. There is no detectable dependence of this ratio on δM. The slopes with respect to $\overline{M}^2 \propto \phi_4$ decrease with the lattice spacing, indicating that the dominant violations of the GMOR relation $\overline{m}/\overline{M}^2 = \text{const}$ are due to lattice artefacts. However, as already discussed above, the ratio itself does not decrease linearly with a. Instead, it first increases from $\beta = 3.34$ up to $\beta = 3.46$ and only then it slowly starts to decrease again.

Since the $a\overline{m}$ dependence is hard to distinguish from the $\overline{M}^2$ dependence, within the global fit we fix b_a to its one-loop value eq. (2.109), $b_a = b_a^{(1)} \approx 0.3158$. The relation (3.29) between $\tilde{k}_1$ and $\tilde{k}$ will enable a cross-check with the previous fit result eq. (3.30) that was obtained by extrapolating the individual slopes $\tilde{k}(a)$ of eq. (3.28) to the continuum limit. In addition to this central fit, we carry out a second fit, setting $b_a = 0$ and interpret the difference between the resulting parameters as a systematic uncertainty. This second fit also allows for a comparison not only with eq. (3.30) but also with the earlier results shown in eq. (3.37). We remark that we expect the actual t_0^*/a^2 values

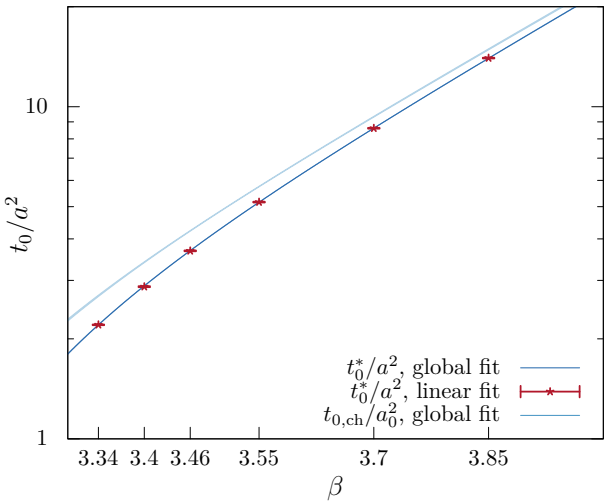


FIGURE 3.8: Global fit results (cf. eq. (3.39)) for t_0^*/a^2 and t_0^X/a_0^2, together with results for t_0^*/a^2 taken from separate linear fits (cf. eq. (3.28)) to the $m_s = m_\ell$ data.

to differ by even less between the two fits than the parameters b_{t_0}, c_{t_0} and d_{t_0} since a part of the difference can be accommodated for by the difference of the b_a values in eq. (3.35).

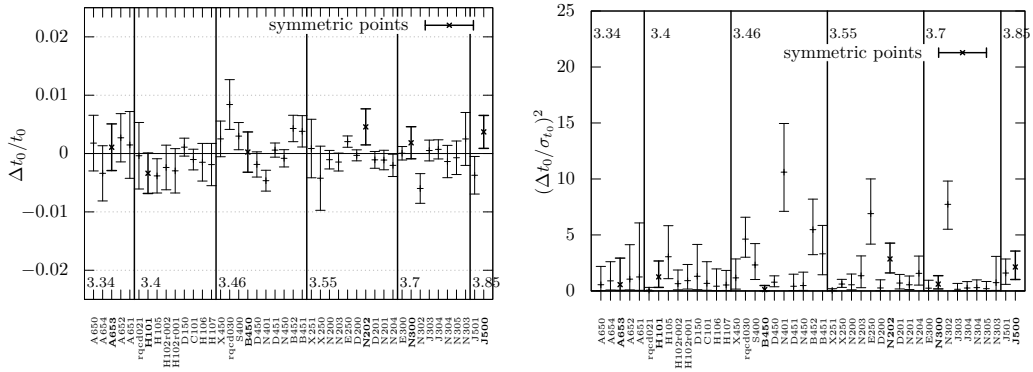


FIGURE 3.9: Deviations $\Delta t_0/a^2 = t_0/a^2(g^2, \overline{M}^2, \delta M^2) - t_0/a^2$ between the values postdicted by the fit eq. (3.39) and the data are typically below one per cent (left) or of the order of the statistical error (right), with only a few exceptions. The ensembles are sorted from left to right in terms of decreasing lattice spacing and increasing values of the average $\overline{M}^2$. Boldface ensemble names correspond to the "symmetric" points, closest to the position of ϕ_4^*.

In fig. 3.9 it is shown that the data are well described by the fit with $b_a = b_a^{(1)}$ and remark that the picture looks very similar for $b_a = 0$. The fit curve itself is shown for the $m_s = m_\ell$ points in

fig. 3.7. The respective fit parameters read

$$\begin{aligned}
\tilde{k}_1(b_a = b_a^{(1)}) &= -0.0466(62), & \tilde{k}_1(b_a = 0) = \tilde{k} &= -0.0506(63),\\
\bar{c}(b_a = b_a^{(1)}) &= -0.560(27), & \bar{c}(b_a = 0) &= -0.560(27),\\
\delta c(b_a = b_a^{(1)}) &= 0.0213(28), & \delta c(b_a = 0) &= 0.0210(28),\\
b_{t_0}(b_a = b_a^{(1)}) &= -0.9336(26), & b_{t_0}(b_a = 0) &= -0.9340(26),\\
c_{t_0}(b_a = b_a^{(1)}) &= 0.286(63), & c_{t_0}(b_a = 0) &= 0.254(63),\\
d_{t_0}(b_a = b_a^{(1)}) &= -1.567(84), & d_{t_0}(b_a = 0) &= -1.515(84),\\
\chi^2/N_{\mathrm{DF}}(b_a = b_a^{(1)}) &= 59.4/38, & \chi^2/N_{\mathrm{DF}}(b_a = 0) &= 58.7/38,
\end{aligned} \tag{3.41}$$

where the errors have been obtained from the bootstrap distributions of the parameters and scaled with $\sqrt{\chi^2/N_{\mathrm{DF}}}$. In fig. 3.8 we compare the global fit to the t_0^*/a^2 values obtained from the local interpolations above. In addition, we show the chiral limit of this ratio.

The differences between the central values of the two columns of eq. (3.41) constitute the systematic errors from varying b_a from zero to twice its one-loop value. As expected, the slope parameter $\tilde{k}_1$ is subject to the largest relative systematic uncertainty. The $b_a = 0$ result is in agreement with our previous estimate eq. (3.30). Moreover, the difference between the two values is indeed consistent with eq. (3.29). Combining the $b_a = 0$ result with the typical value $am^* = 0.004$ (see fig. 3.5 and tab. 3.2), we obtain $\tilde{k}_1 \approx \tilde{k} + 3b_a am^*/\phi_4^* \approx -0.0472(63)$, which indeed is very close to the result of the $b_a = b_a^{(1)}$ fit $\tilde{k}_1 = -0.0466(62)$. Moreover, the parameters b_{t_0}, c_{t_0} and d_{t_0} agree within the errors with the determination eq. (3.37), based on the local fit results.

We used the central value $\Lambda_{\overline{\mathrm{MS}}}\sqrt{8t_0^*} = 0.712$ of the determination [128] eq. (3.32) as an input. Instead, we could have included the normalization as a free fit parameter. Carrying out such a fit out of curiosity, we find $\Lambda_{\overline{\mathrm{MS}}}\sqrt{8t_0^*} = 0.68(17)$, in agreement with the more precise result that was obtained employing the step scaling function within the Schrödinger functional framework [128].

Results

Using this global fit, we compute values for t_0^*/a^2 for each value of the lattice coupling β and collect the results in the second line of tab. 3.3, where the first error is statistical and the second one reflects the impact of the uncertainty of b_a. Similarly, in the fourth line of the table we list t_0^χ/a_0^2, where naturally the statistical uncertainty is larger while the uncertainty of the b_a value has less of an impact. In general, the t_0^*/a^2 results from the global fit are in very good agreement with our local determination. In comparison to the (local) fit results of [63, 135], our values at the coarser lattice spacings $\beta = 3.4$ and $\beta = 3.46$ are significantly larger. This may be due to our ability to parametrize lattice spacing effects better by incorporating more data at varying quark masses as well as improved statistical accuracy by incorporating more data. This also increases the accuracy at the finest lattice spacing where only few data points are available. Therefore, we regard the results from the global interpolation where statistical fluctuations of individual ensembles to a large degree average out as more robust.

The influence of these determinations on actual results, however, is small since we extrapolate combinations of t_0/a^2 with the observables to the physical point and only then set the scale using eq. (3.14) as described in sec. 3.4. In some instances, however, a value of t_0^*/a^2 still is required, e.g., when matching the scale of the renormalization factors, $\mu = a^{-1}$, or for the parametrization of some lattice spacing effects, cf. 5.2.

Some of our results are defined in the chiral limit and therefore a relation of the value of t_0^{ph} at the physical point to the value in the chiral limit, t_0^χ is desirable. To this end, we compute the slope parameter [145] of eq. (3.15), where we use the globally fitted $\tilde{k}_1$ with the systematic error due to the

uncertainty of b_a included, reads

$$\tilde{k}_1 = -0.0466(62), \quad k_1 = \frac{8t_0^\chi (4\pi F_0)^2}{3}\tilde{k}_1 \approx -0.084(11), \tag{3.42}$$

assuming the $N_f = 2+1$ FLAG [152] values $F_\pi/F_0 = 1.062(7)$ and $F_\pi = F_{\pi^+} = 92.2\,\mathrm{MeV}$ based on the determinations in [153–157]. From the value of $\tilde{k}_1$ in the continuum limit and eq. (3.16), we can deduce the continuum limit ratio

$$\frac{t_0^\chi}{t_0^*} = \left(1 + \frac{2}{3}\tilde{k}_1\phi_4^*\right)^{-1} = 1.036(4). \tag{3.43}$$

This difference of about 4 % is larger than the typical statistical errors that we achieve and should be taken into account, when extracting, e.g., low energy constants in the chiral limit as we will do in sec. 5.4.1.

For convenience, we refit our predictions for the t_0^*/a^2 values, setting $b_{t_0} = 0$ but keeping the g^{-2} and $\ln(g^2)$ coefficients fixed, and find that they are well described by the interpolating formula

$$\begin{aligned}\frac{t_0^*}{a^2}(g^2) &= f_{\text{eff}}(g^2) + 0.285 - 1.566 f_{\text{eff}}^{-1/2}(g^2), \quad \text{where} \\ f_{\text{eff}}(g^2) &= \exp\left(17.54596\, g^{-2} - 7.507 + 0.790123\, \ln(g^2) - 0.9334 g^2\right).\end{aligned} \tag{3.44}$$

The relative errors are below 0.3 % over the entire fitted range $3.34 \le \beta \le 3.85$. Within the errors of the Λ parameter of ref. [128], this interpolation converges towards the two-loop running of the scale at small values of g^2, making this formula particularly useful for predictions of new simulation parameters. For values at already existing simulation points the second line of table 3.3 should be consulted.

4 Extraction of η and η' mass eigenstates

As discussed in sec. 2.1.6 the η and η' physical states are no flavour eigenstates, but related to the interpolating functions by flavour rotations. These depend on the combination and precise definition of the interpolators, including the smearing and other details of the lattice computation and need to be determined from the data.

In this chapter we give details on the measurements of the necessary correlation functions on the individual configurations and how we extract physical mass eigenstates from the ensembles of such measurements.

To this end we start in sec. 4.1 by introducing the stochastic methods necessary to calculate the disconnected contributions on each individual gauge field configuration. Thereby, we employ two noise reduction techniques: Time dilution, described in sec. 4.1.1 and the hopping parameter expansion, sec. 4.1.2. We explore the efficiency of these techniques for pseudoscalar and axialvector loops, and comment on the choice of parameters in sec. 4.1.3. In sec. 4.2.1 the determined loops are then used to construct matrices of correlators that we then take ensemble averages of and diagonalize in the remainder of sec. 4.2. This requires a careful assessment of excited states and statistical noise. This window problem is highlighted in sec. 4.2.2 and motivates our fit ansatz detailed in sec. 4.2.3: The matrix analogue of effective masses is introduced, and the statistical precision is improved by incorporating data at non-vanishing momentum. Details of the mass determination are given, and the results are compared to those obtained by employing the GEVP. Finally, in sec. 4.3, we discuss how the decay constants of the η and η' mesons are obtained from combined fits including correlation functions constructed with local currents at the sink.

Parts of this chapter have already been published in a similar form in [60]. This applies in particular to secs. 4.1.1, 4.1.2, 4.2.1, 4.2.3 and 4.3.

4.1 Stochastic measurements of disconnected loops

The basic building blocks of disconnected correlation functions, eq. (2.89), are quark loops of a particular flavour $f \in \{\ell, s\}$

$$L_s^{\Gamma,f}(\vec{p},t) = a^3 \sum_{\vec{x},\vec{y},\vec{z}} \mathrm{tr}\left(e^{-i\vec{p}\cdot\vec{x}}\phi^s(x,y)\Gamma D_f^{-1}(y,z)\phi^s(z,x)\right), \tag{4.1}$$

where Γ determines the pseudoscalar ($\Gamma = \gamma_5$) or axialvector ($\Gamma = \gamma_\mu\gamma_5$) Dirac structure. The trace is taken over the spin and colour components and the space-time positions x, y and z share the same time t, i.e. $x = (t,\vec{x}), y = (t,\vec{y})$ and $z = (t,\vec{z})$. The quark smearing functional ϕ^s is defined in sec. 2.3.2, where s is the number of smearing iterations. We implement three levels of smearing, $s \in S = \{0, s_1, s_2\}$, which allows us to analyse local matrix elements as well as to extend our basis of interpolators. The ratio s_2/s_1 is kept approximately constant on all ensembles and the number of smearing iterations is increased with decreasing pion mass and lattice spacing. We list the number of smearing iterations along with other parameters that will be discussed below in tab. 4.1.

Since D_f^{-1} is a very large matrix, only its matrix product with a source $|\rho\rangle$ can be computed with reasonable cost to obtain the solution vector

$$|\sigma_f\rangle = D_f^{-1}|\rho\rangle. \tag{4.2}$$

Hence, the full trace in eq. (4.1) cannot practically be computed exactly but must be estimated stochastically [158]. In this section, we will first introduce the basic concept of stochastic estimators for loops. In fact a number of algorithms have been suggested that take the structure of the Dirac operator into account and converge faster than the naive probing of the trace. These algorithms are typically based on partitioning, i.e., improved sources for the stochastic probing [159–162], low-mode deflation of the matrix [105, 161, 163–165] or inexact and hence cheaper matrix-vector products [166–169]. For our purposes it sufficed to concentrate on the hopping parameter expansion [170] and partitioning of the sources in time.

4.1.1 Time-partitioned sources and open boundary conditions

To suppress short-distance (near-diagonal) noise in the trace estimation, it is beneficial to construct sparse — diluted — sources [159]. It is possible to do this for spin and colour indices or space-time blocks. Here, we choose to set specific time slices to zero,

$$\rho_{\tau,i}(x,\alpha,a) = \begin{cases} \frac{r_i(x,\alpha,a)}{\sqrt{2}} & \mathrm{mod}(t/a, \Delta_t/a) = \tau \text{ and } b \leq t < L_t - b, \\ 0 & \text{otherwise} \end{cases}, \tag{4.3}$$

where $L_t = a\,N_0$ is the temporal lattice extent and $r_i(x,\alpha,a) \in \mathbb{Z}_2 \times i\mathbb{Z}_2$ are random numbers drawn independently for every site x, spin α and colour component a. Δ_t is the distance between timeslices on which the source has support. On lattices with open boundary conditions in time, we set $b > 0$ in order to suppress boundary effects. We comment on our parameter choices in sec. 4.1.3. These random sources span a space in the bulk of the lattice,

$$\begin{aligned} \sum_{\tau=0}^{\Delta_t/a-1} \sum_{i=0}^{N_{\mathrm{stoch}}-1} |\rho_{\tau,i}\rangle\langle\rho_{\tau,i}| =& N_{\mathrm{stoch}}\ \mathrm{diag}(\underbrace{0,\ldots,0}_{b/a},\underbrace{1,\ldots,1}_{L_t/a-2b/a},\underbrace{0,\ldots,0}_{b/a}) \otimes \mathbb{1}_{12V_3/a^3} \\ &+ \mathcal{O}(1/\sqrt{N_{\mathrm{stoch}}}), \end{aligned} \tag{4.4}$$

$$a^4\langle\rho_{\tau,i}|\rho_{\tau',j}\rangle = (L_t - 2b)V_3\,\delta_{i,j}\delta_{\tau,\tau'}, \tag{4.5}$$

where $V_3 = L_s^3$ and N_{stoch} such sources are created for every dilution index $\tau = 0,\ldots,\Delta_t/a-1$.

The lattice Dirac equation

$$D_f \sigma^f_{\tau,i} = \rho_{\tau,i} \tag{4.6}$$

is solved for each fermion flavour f and source $\rho_{\tau,i}$, labelled by a stochastic index i and time partition τ, to obtain the solution $\sigma^f_{\tau,i}$.

By summing over the dilution index τ we probe the loop computed on the i-th stochastic source,

$$L^{\Gamma,f}_{i,s}(\vec{p},t) = a^3 \sum_{\vec{x}} \sum_{\tau=0}^{\Delta_t/a-1} \mathrm{tr}\left(e^{-i\vec{p}\cdot\vec{x}}(\rho_{\tau,i}\phi^s)(x)\Gamma(\sigma^f_{\tau,i}\phi^s)(x)\right). \tag{4.7}$$

After averaging over these stochastic probes, we obtain an estimate of axialvector and pseudoscalar loops of a particular flavour and for a given gauge field configuration and smearing,

$$L^{\Gamma,f}_s(\vec{p},t) = \frac{1}{N_{\mathrm{stoch}}} \sum_{i=0}^{N_{\mathrm{stoch}}-1} L^{\Gamma,f}_{i,s}(\vec{p},t) + \mathcal{O}\left(\frac{1}{\sqrt{N_{\mathrm{stoch}}}}\right). \tag{4.8}$$

This requires $N_{\text{stoch}} \times \Delta_t/a$ inversions for each flavour. To extend our basis of interpolators, we compute loops with different levels of smearing, i.e. $s \in \{0, s_1, s_2\}$. Unlike in the connected case, this does not require any additional inversions: due to its hermiticity the smearing operator can be applied to the stochastic sources and solutions after the inversion, as indicated by the parentheses in eq. (4.7).

4.1.2 Hopping parameter expansion

The inverse of the Wilson Dirac operator[1] $D_f = \frac{1}{2\kappa_f}(\mathbb{1} - \kappa_f H)$ within the trace of eq. (4.1) can be expanded for small values of the hopping parameter κ_f. This yields a geometric series in terms of the nearest-neighbour hopping term H [69, 166, 170],

$$\text{tr}\left(\Gamma D_f^{-1}\right) = 2\kappa_f \sum_{i=0}^{\infty} \kappa_f^i \text{tr}\left(\Gamma H^i\right) = 2\kappa_f \sum_{i=0}^{n-1} \kappa_f^i \text{tr}(\Gamma H^i) + \kappa_f^n \text{tr}\left(\Gamma H^n D_f^{-1}\right). \tag{4.9}$$

Above we restricted ourselves to the case without smearing. On the right-hand side of the equation we have split the series into the first n terms for which $\text{tr}\,\Gamma H^i = 0$ and a remainder. Note that the value of n depends on Γ and the exact fermion action employed and can easily be determined numerically by applying the corresponding power $n(\Gamma)$ of the (Clover-improved) Dirac operator to a point source, multiplying with the Dirac structure and then contracting again with the source,

$$\epsilon = \sum_s \sum_c \langle \delta_{s,c,x} \mid \Gamma(\kappa D)^n \mid \delta_{s,c,x} \rangle.$$

This can be done for example on a small lattice and using a random gauge field. If ϵ is zero, then there is no contribution to that order. In our case, in the absence of smearing, $n(\gamma_5) = 2$ for pseudoscalar and $n(\gamma_\mu\gamma_5) = 4$ for axialvector loops.

In the stochastic estimation the first sum in eq. (4.9) only contributes to the noise coming from short-distance hops. Hence, we can obtain an improved estimate of the trace, by applying the Dirac operator n times to the solution, replacing

$$\Gamma\sigma^f_{\tau,i} \mapsto \Gamma(1 - 2\kappa_f D_f)^{n(\Gamma)}\sigma^f_{\tau,i} \tag{4.10}$$

in eq. (4.7) and thus cancelling out nearest-neighbour contributions that do not improve the signal.

4.1.3 Computational cost and choice of parameters

Any noise reduction technique comes at a computational cost that needs to be smaller than the time that would be required to simply increase the number of stochastic estimates. It is thus important to choose an efficient combination of such improvements, ensuring a good signal-to-cost ratio.

When increasing the stochastic estimates on each configuration, the error will slowly approach the gauge noise. The latter can only be reduced by prolonging the Monte-Carlo chain and generating more independent measurements. Due to the coupling of the pseudoscalar current to the topology of the vacuum, these loops vary only slowly and the gauge noise quickly dominates over the stochastic error from the estimation of the loops, see fig. 4.1 where the total error of the average loop is shown as a function of the number of stochastic estimates. For that reason, we found that the axialvector current needed for the computation of decay constants benefits the most from both a large number of stochastic sources and the employed noise reduction techniques whereas the error of the pseudoscalar loops quickly saturates to the gauge noise. This situation is greatly improved by including data at a non-vanishing momentum in the fits, see the bottom row of fig. 4.1 and the discussion in sec. 4.2.3.

[1] See eq. (2.65) for a definition of the hopping term neglecting the c_{SW} term.

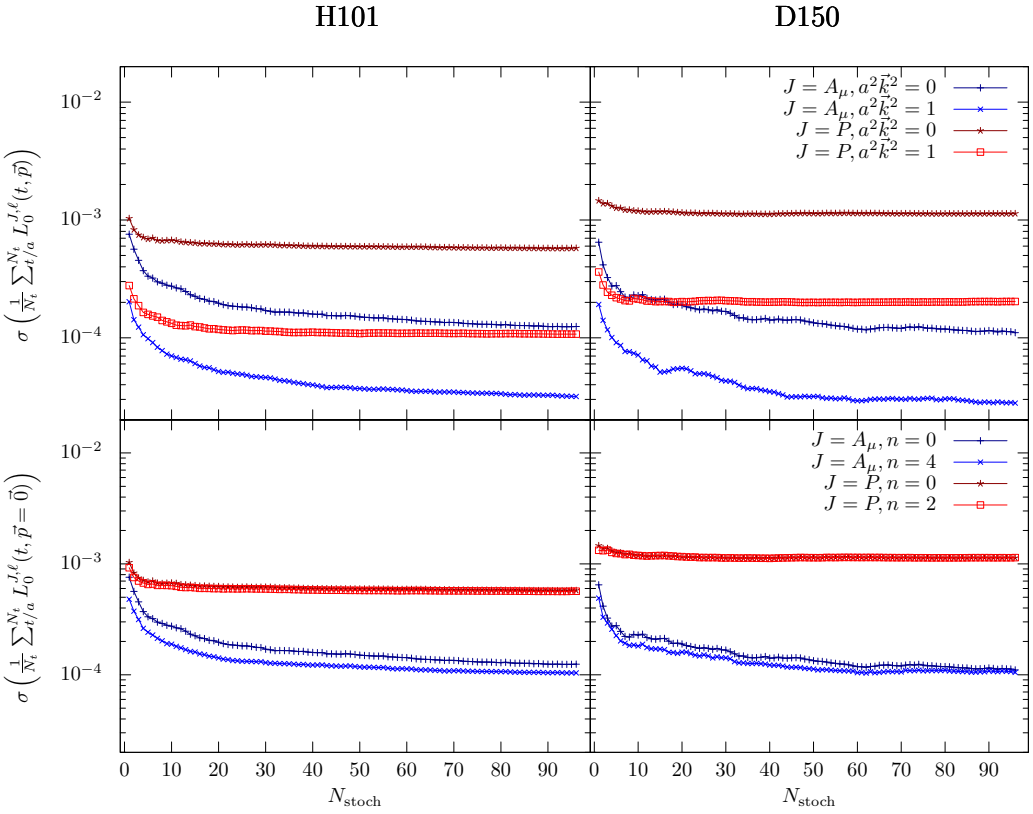


FIGURE 4.1: Errors of the light loop averaged over all N_t sourced time slices for the pseudoscalar ($J = P$, red) and axialvector ($J = A_\mu$, blue) local currents. The left panel shows H101, a $N_f = 3$ symmetric point ensemble; the quark masses of D150 displayed on the right are tuned to approximately their physical values. Colour shades in the top panel correspond to the zero (dark) and $a^2\vec{k}^2 = 1$ case (bright) and in the bottom panel to the number of applications of the hopping parameter expansion n, without HPE (dark) and using the maximum number of applications (bright). Finite momentum data has been averaged over all six realizations with $a^2\vec{k}^2 = 1$. Due to time dilution one stochastic source corresponds to solving the Dirac equation four times. The displayed error is computed over the full ensemble and approaches the gauge noise of the respective currents and ensembles. As demonstrated in the top panel momentum averaging yields a reduction in the error of the loop of almost an order of magnitude for both currents. The hopping parameter expansion is less effective but comes at almost no cost relative to the costs of solving the Dirac equation.

The hopping parameter expansion is — except for the maximum admissible value for its application n — parameter free and involves n application of the Dirac operator but no additional inversions. It thus is very cheap and cost-effective and can be applied to all unsmeared loops. Due to the smallness of the expansion parameter κ it is more effective at large quark masses and we found it to work better for the axialvector current where $n = 4$ and more stochastic probes are required to approach the gauge error.

Also time partitioning eliminates noise from neighbouring time slices, however at the expense of additional inversions, requiring a factor Δ_t/a more inversions. Based on initial experiments, we chose a time separation of $\Delta_t = 4a$ in the stochastic dilution, except for the finest lattice spacing where we used $\Delta_t = 6a$. These correspond to roughly $\Delta_t \approx 0.3\,\mathrm{fm}$. These choices are listed in tab. 4.1, as well as the number of configurations analysed and the distance between consecutive measurements in Hybrid Monte-Carlo molecular dynamics time units. We also list the smallest distance to the boundary b to control boundary effects. This parameter has been fixed from experience from the pion [90]: At a distance b boundary effects are smaller than the statistical error of the pion correlation function. This condition results in $b \gtrsim 1.9\,\mathrm{fm}$ and is a conservative choice, given the comparably large errors of the disconnected contributions.

We remark that due to the use of a highly efficient multigrid solver [171–173], we do not benefit from additional noise reduction techniques like, e.g., the truncated solver method (TSM) [166] within our setup. This is due to the fact that TSM reduces the average time needed for one solve by performing inversions only approximately and only correcting for the bias with exact solves. This requires in total more inversions but at a smaller total cost for all inversions. The computer time spent for contracting and smearing the source and solution vectors according to eq. (4.7), however, increases linearly with the number of estimates and soon dominates the total cost even though the implementation of the smearing kernel is highly optimized for the hardware available to us. In our implementation, where we only use exact solves, the computational cost for the smearing and the contractions still accounts for roughly a third of the total computing time.

Also low-mode averaging (LMA) [105, 161, 163, 164], even when reusing orthogonalized multigrid test vectors as studied in [165], is not expected to improve the signal-to-cost ratio due to a significant higher setup cost for the solver that scales approximately quadratically with the spatial volume. Also projecting out the high mode contribution and additional contractions become more expensive in particular at small pion masses and large spatial volumes.

In summary, we stick to a relatively simple setup employing only time partitioning and the hopping parameter expansion. The use of efficient solvers and hardware allows to compute the loops to a high precision at a modest cost. These parameters have to be chosen before actually carrying out the measurements and parameter tuning can only be done at relatively cheap ensembles and then conservative projections need to be made. In hindsight possibly a slightly smaller total number of stochastic estimates would have been sufficient even for the lighter ensembles, cf. fig. 4.1. The precision of the pseudoscalar loops in our case is mostly limited by the limited statistics of the Monte-Carlo chains.

4.2 Analysis of matrices of correlation functions

Having stochastic estimates of pseudoscalar and axialvector loops available at various smearing levels, we may now proceed to construct matrices of correlators from them. Using several smearings and interpolating functions using light, strange, octet and singlet combinations of loops together allows us to increase the overlap with the η and η' ground states. The basic idea and how this matrix is related to the physical eigenstates has already been discussed in sec. 2.3.4 and here we concentrate on the numerical task of constructing and diagonalizing these matrices, identifying the eigenstates with the physical states.

id	N_{conf}	ΔMDU	Δ_{bin}	S	b/a	Δ_t/a	N_{stoch}	$N_{\text{solves}}/10^3$
H101	963	8	4	$\{0, 55, 92\}$	30	4	96	370
H102a	490	8	4	$\{0, 63, 104\}$	30	4	96	376
H102b	491	8	4	$\{0, 63, 104\}$	30	4	96	377
H105	899	8	4	$\{0, 75, 125\}$	30	4	96	690
C101	504	8	4	$\{0, 88, 146\}$	30	4	96	387
D150	502	4	8	$\{0, 125, 208\}$	0	4	96	386
H107	778	8	4	$\{0, 63, 107\}$	30	4	96	598
H106	754	8	4	$\{0, 63, 104\}$	30	4	96	579
C102	729	8	4	$\{0, 88, 146\}$	30	4	96	560
B450	794	8	4	$\{0, 68, 113\}$	0	4	96	305
S400	796	8	4	$\{0, 78, 129\}$	30	4	96	611
N401	500	8	4	$\{0, 94, 156\}$	34	4	96	384
B451	1000	8	4	$\{0, 68, 113\}$	0	4	96	768
B452	962	8	4	$\{0, 83, 129\}$	0	4	96	739
N202	440	8	6	$\{0, 98, 163\}$	30	4	96	169
N203	563	8	6	$\{0, 111, 185\}$	30	4	96	432
N200	853	8	6	$\{0, 135, 225\}$	30	4	96	655
D200	582	8	8	$\{0, 165, 275\}$	30	4	96	447
N204	745	8	6	$\{0, 111, 185\}$	30	4	96	572
N201	757	8	6	$\{0, 135, 225\}$	30	4	96	581
D201	535	8	8	$\{0, 165, 275\}$	34	4	96	411
N300	754	8	10	$\{0, 165, 275\}$	49	6	96	434

TABLE 4.1: Parameters related to the measurement of the correlation functions: the number of analysed configurations N_{conf}, their separation in molecular dynamics units ΔMDU, the choice of binning Δ_{bin} to account for autocorrelation effects in the statistical analysis, the numbers of smearing iterations $s \in S$, the distance from the temporal boundaries b (in the case of open boundary conditions) and the time partitioning separation Δ_t. In the last column we display the total number of individual Dirac vector solves carried out on each ensemble to compute the disconnected correlation functions. The number of solves needed for the connected part is much smaller (72: 2 quark masses $\times$ 3 smearing levels $\times$ source spin-colour). Ensembles H102a and H102b were generated with the same quark masses and lattice coupling but different simulation parameters and are therefore analysed separately.

4.2.1 Construction of correlators

We define matrices of correlation functions using suitable momentum projected interpolators as those in eqs. (2.79),

$$C_{ij}(\vec{p},t) = \frac{1}{N_{t_{\text{in}}}} \sum_{t_{\text{in}}} \left\langle \Omega \left| \mathcal{B}_i(\vec{p}, t+t_{\text{in}}) \mathcal{B}_j^\dagger(-\vec{p}, t_{\text{in}}) \right| \Omega \right\rangle , \tag{4.11}$$

where $N_{t_{\text{in}}}$ denotes the number of source time slices that we average over. The Wick contractions are given in eqs. (2.93) to (2.94) and give both connected, $\tilde{C}$, and disconnected, $\tilde{D}$ correlation functions.

In order to estimate the disconnected two-point function, we correlate and average two of the loops, defined in eq. (4.7):

$$\tilde{D}^{\Gamma_1\Gamma_2}_{\substack{f_1,f_2\\ s_1,s_2}}(\vec{p},t) = \frac{1}{N_{\text{stoch}}(N_{\text{stoch}}-1)} \frac{a}{L_t} \sum_{\substack{i,j=0\\ i\neq j}}^{N_{\text{stoch}}-1} \sum_{t_{\text{in}}=0}^{L_t-a} \left\langle L^{\Gamma_1,f_1}_{i,s_1}(\vec{p}, t_{\text{in}}+t) L^{\Gamma_2,f_2}_{j,s_2}(-\vec{p}, t_{\text{in}}) \right\rangle . \tag{4.12}$$

Note that we are only allowed to sum over products of loops that have been obtained on different random sources, hence $i \neq j$. Equation (4.12) applies to periodic lattices, where $b = 0$ and the correlation functions wrap around the lattice (the periodicity of the loop is implicit, $L^{\Gamma,f}_{i,s}(\vec{p}, t+L_t) = L^{\Gamma,f}_{i,s}(\vec{p}, t)$). It is straightforward to adapt the above equation to lattices with open boundaries by restricting the sum over t_{in} such that both t_{in} and $t_{\text{in}} + t$ remain in the bulk of the lattice (defined to be a distance b away from the boundaries).

We implement forward-backward averaging for the disconnected two-point functions by simply symmetrizing with respect to the ordering of the source and sink operators:

$$\overline{D}^{\Gamma_1\Gamma_2}_{\substack{f_1,f_2\\ s_1,s_2}}(\vec{p},t) = \frac{1}{2} \left(\tilde{D}^{\Gamma_1\Gamma_2}_{\substack{f_1,f_2\\ s_1,s_2}}(\vec{p},t) + \text{sgn}(\Gamma_1,\Gamma_2) \tilde{D}^{\Gamma_2\Gamma_1}_{\substack{f_2,f_1\\ s_2,s_1}}(\vec{p},t) \right) , \tag{4.13}$$

where $\text{sgn}(\gamma_5,\gamma_5) = \text{sgn}(\gamma_\mu\gamma_5,\gamma_\mu\gamma_5) = 1$ and $\text{sgn}(\gamma_\mu\gamma_5,\gamma_5) = -1$. We use the same random sources for light and strange quark inversions, preserving the correlations between light-light, strange-strange as well as light-strange disconnected correlation functions. This is beneficial when computing differences of disconnected correlation functions which appear after the Wick contraction of some of the basis states.

The connected contributions are cheaply computed by exploiting γ_5-hermiticity, cf. eq. (2.91). Due to the reduced error compared to the disconnected correlation functions, it is sufficient to employ only a single source position per configuration at $x_{\text{in}} = (b, \vec{0})$, leaving $(L_s - 2b)/a$ timeslices for the extraction of the physical states. Unlike on lattices with periodic boundary conditions, in the case of open boundary conditions, in this case we do not carry out the forward-backward averaging of eq. (4.13).

4.2.2 Excited states and signal-to-noise problem

After combining connected and disconnected correlators to matrices of correlators, eq. (4.11), we will now focus on the extraction of the η and η' eigenstates from these. In this section we first focus on the difficulties that are involved in the extraction of these flavour diagonal states, in particular the large contribution from excited states at small times and the early onset of noise dominance over the signal.

To demonstrate these difficulties we first investigate the situation of the η' on H101, an ensemble where all three quarks are mass degenerate. In this case there is no mixing between the octet and the singlet currents and so all physical states are either pure octet or pure singlet states and $C(t)$ can

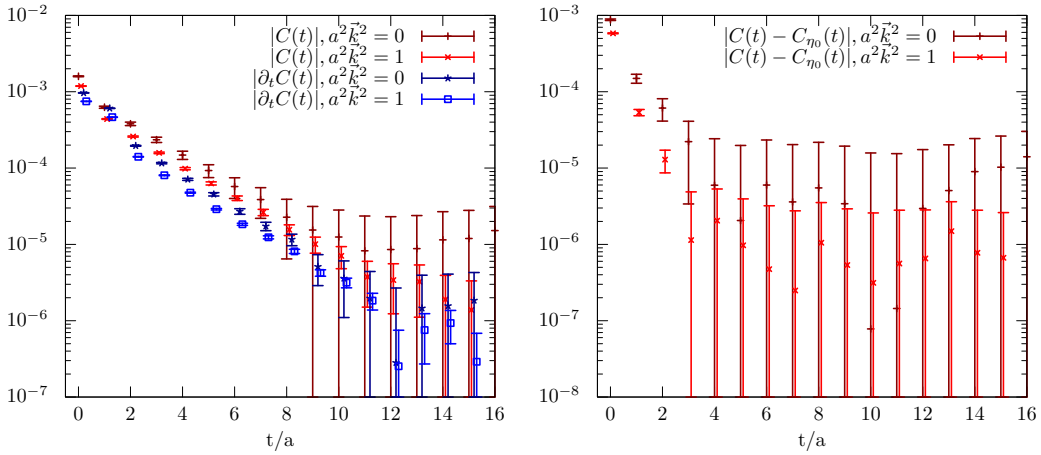


FIGURE 4.2: **Left panel**: Absolute values of the smeared-smeared, singlet-singlet pseudoscalar correlators ($\mathcal{B}_i = \mathcal{B}_j = \mathcal{P}^0_{\vec{p}}$ in eq. (4.11)) and their derivatives for the $N_f = 3$ symmetric point ensemble H101. There, the η' is a pure singlet and so no diagonalization is necessary. **Right panel**: The same data but after subtraction of the fitted η' ground state.

be chosen to be diagonal, e.g. by using $\mathcal{B}_i \in \{\mathcal{P}^8, \mathcal{P}^0\}$. The singlet interpolator will only create the η' and higher excitations, and we plot the singlet-singlet correlators at zero and the smallest available three-momentum $a^2\vec{k}^2 = 1$ on the left-hand side of fig. 4.2. The error of the $a^2k^2 = 0$ data increases rapidly with t/a and the signal is lost at $t/a \gtrapprox 9$, while the finite momentum data shows a signal over a few more timeslices and generally exhibits a better signal-to-noise ratio. In both cases the error is approximately constant over the range of t/a while the signal decays exponentially and eventually is lost in the noise, leaving only a few points for the extraction of the states. In addition these points are highly correlated and this effectively reduces the available information further. These correlations and potential shifts of the correlators away from zero at large times can be drastically reduced by extracting the states from the temporal derivative of the correlation functions [174, 175]. In deed, we observe reduced errors and an extended range suitable for the extraction of the states in this case, see fig. 4.2.

At small Euclidean times the correlators are more precise, but the correlator cannot be described as a single exponential since excited states contribute significantly. In nature, there are a multitude of resonances with the same quantum numbers lying just above the $\eta'(958)$: the $\eta(1295)$, $\eta(1405)$ and $\eta(1475)$ are all close-by and will, in general, contribute as excited states to $C(t)$. In the right panel of fig. 4.2 we plot the correlators after subtracting the lowest fitted η' state. In this case, the first excitation of the η' is easily detectable up to $t/a \approx 5$. It is therefore important to include at least a third state in our analysis to provide an effective parametrization of the contributions of these higher states. In principle, also strong decays of the η and η' should be taken into account. For the η', the dominant decay is $\eta' \to \eta\pi^+\pi^-$ (branching ratio 42.5 % [19]). This is kinematically only possible on ensemble D150, for our lightest quark mass. Its decay width, however, is about 80 keV, which would be very difficult to resolve considering the statistical precision we achieve. Other channels have even smaller decay rates and many, such as $\eta \to 3\pi^0$, are forbidden in the isospin limit of QCD that we simulate.

To summarize, our fitting procedure must address the following issues: At small Euclidean times, the correlators are affected by sizable excited states contributions. These need to be included in our parametrization. At the same time, high noise levels at large times set an upper limit for the usable time slices. In many cases fit ranges can only extend over about 0.5 fm to yield acceptable results in terms of their χ^2/N_{df}. These remaining time slices usable for the extraction are in general correlated and the range of time slices that is usable does not only depend on the quark mass and lattice spacing but also on the particular interpolators, i.e. the smearing and flavour combination

used and the signal for the octet combinations is better than that for the singlet ones. Therefore, it is necessary to choose fit ranges independently for the individual entries of the matrix C_{ij}.

4.2.3 Fits to matrices of correlation functions

Away from the flavour symmetric point, the matrix of correlators $C(t)$, eq. (4.11) is not diagonal and the masses of the η and η' are extracted from diagonalizing the full matrix. Usually, this is done by solving a generalized eigenvalue problem (GEVP) and fitting to the resulting eigenvalues [176, 177]. Here, we follow a different route and directly fit to the elements of $C(t)$ or its time-derivative, $\partial_t C(t)$. The latter reduces correlations in the Euclidean time t. This also allows us to adjust the fit ranges for the entries $C_{ij}(t)$ individually as motivated in the previous section.

Fit forms

In the limit of infinite statistics, C (eq. (4.11)) is a real symmetric, positive-definite $M \times M$ matrix, cf. eq. (2.96). The spectral decomposition gives

$$C_{ij}(t) = \sum_{n=0}^{\infty} \frac{1}{2E_n V_3} \left\langle \Omega \left| \mathcal{B}_i(t) \right| n \right\rangle \left\langle n \left| \mathcal{B}_j^\dagger(0) \right| \Omega \right\rangle , \tag{4.14}$$

$$= \sum_{n=0}^{\infty} \frac{1}{2E_n V_3} \exp(-E_n t) \left\langle \Omega \left| \mathcal{B}_i(0) \right| n \right\rangle \left\langle n \left| \mathcal{B}_j^\dagger(0) \right| \Omega \right\rangle , \tag{4.15}$$

where we suppress the momentum argument and only consider a single source at $t_{\text{in}} = 0$. The lowest energy states correspond to the ground states of the η/η' system, $|n = 0\rangle = |\eta\rangle$ and $|n = 1\rangle = |\eta'\rangle$. Equation (4.14) can be written as

$$C(t) = \widehat{Z}\widehat{D}(t)\widehat{Z}^\intercal, \tag{4.16}$$

where $\widehat{D}(t) = \text{diag}(\exp(-E_n t))$ for $n = 0, \ldots$ is time dependent, while

$$\widehat{Z}_{in} = \frac{1}{\sqrt{2E_n V_3}} \langle \Omega | \mathcal{B}_i(0) | n \rangle \tag{4.17}$$

are time independent amplitudes (that depend on the smearing and momentum). In practice, we truncate the infinite sum to determine only the lowest N states, hence,

$$C(t) = Z D(t) Z^\intercal + \mathcal{O}(\exp(-E_N t)), \tag{4.18}$$

where $D \in \mathbb{R}^{N \times N}$ and we assume phase conventions such that $Z \in \mathbb{R}^{M \times N}$ has positive entries on the diagonal.

Equation. (4.18) allows to perform a combined fit to $C(t)$, restricting the fit range to times large enough so that any contributions from higher excited states $n \geq N$ fall below the statistical precision. The $M \times N$ amplitudes Z_{in} and N masses are fitted simultaneously to the $(M + 1) \times M/2$ independent components of C_{ij}. The bases of interpolating operators used for each ensemble are detailed in tab. 4.2. We find fits to be most stable for $N = M = 3$. These involve 12 free parameters. We deviate from this choice for the $m_s = m_\ell$ ensembles, for which there is no mixing between the singlet and octet sectors. In this case, the matrix of correlation functions is block diagonal, and we choose $N = M = 4$, such that the problem decomposes into two independent singlet and octet $N = M = 2$ fits.

On the ensembles with open boundary conditions, we take boundary effects into account when computing the loops and connected correlation functions, using sources and sinks that only have

support in the bulk of the lattice, see sec. 4.2.1. This allows the simple ansatz

$$D(t) = \text{diag}(\exp(-E_n t)) \tag{4.19}$$

for the time dependent matrix in eq. (4.18). On lattices with (anti-)periodic boundary conditions in time, states can propagate across the boundary, and we modify D to take the backwards-propagating states into account:

$$D(t) = \text{diag}\left[2\exp\left(-E_n\frac{L_t}{2}\right)\cosh\left(-E_n\left(t-\frac{L_t}{2}\right)\right)\right]. \tag{4.20}$$

As motivated in sec. 4.2.2 and suggested in [174, 175], it is advantageous to reduce the correlations between time slices by fitting to the temporal derivative of the correlation functions. The fit form, eq. (4.16), is modified to

$$\partial_t C(t) \sim Z\left(\partial_t D(t)\right) Z^\intercal, \tag{4.21}$$

where $\partial_t C(t) = \left(C(t+a) - C(t-a)\right)/(2a)$ is the symmetric discretized derivative and

$$\partial_t D(t) = -\,\text{diag}\left[E_n \exp\left(-E_n t\right)\right] \qquad \text{(open boundaries)}, \tag{4.22}$$

$$\partial_t D(t) = -2\,\text{diag}\left[E_n \exp\left(-E_n\frac{L_t}{2}\right)\sinh\left(-E_n\left(t-\frac{L_t}{2}\right)\right)\right] \qquad \text{(periodic boundaries)}. \tag{4.23}$$

We find that this modification enables fits to discriminate between the η and η' contributions more easily, as long as higher excited states are either sufficiently well parameterized (by including them in the fit) or suppressed by the choice of the fit window. In addition, potential constant shifts in the correlation functions (arising from finite volume effects related to incomplete sampling of the topological sectors, see, e.g., [48, 52]) are automatically removed. Although we do not encounter any significant shifts within our data, we observe that utilizing eq. (4.21) leads to decreased correlations and more stable fit results, see also fig. 4.2 for an example.

Inclusion of data at finite momentum

We also include data with non-vanishing momentum in the fit, assuming the continuum dispersion relation

$$aE_n(\vec{p}) = \sqrt{a^2 M_n^2 + a^2\vec{p}^2}, \tag{4.24}$$

where $M_n = E_n(\vec{0})$ is the mass of the n-th eigenstate. On the lattice the momentum components are quantized: $p_j = 2\pi a k_j / L_s$ where k_j are integer multiples of a^{-1}. We average over the six smallest non-trivial lattice momenta ($a^2\vec{k}^2 = 1$) and carry out a combined fit with the $\vec{k} = \vec{0}$ data, assuming the same masses M_n. The error reduction on the level of the individual loops is demonstrated in fig. 4.1 and in fig. 4.3 we show examples of these fits to the full correlation functions (see also fig. 3 of [178] for the inclusion of higher momenta). In addition, we plot the naive lattice dispersion relation for a free scalar particle,

$$aE_n(\vec{p}) = \text{arccosh}\left(\cosh(aM_n) + \sum_j 2\sin^2(ap_j/2)\right). \tag{4.25}$$

Since within the relevant momentum range the differences between the two curves (4.24) and (4.25) are much smaller than the errors of the data, we conclude that assuming eq. (4.24) will not bias

our results. Moreover, we find all data to be well described by this ansatz. The combined analysis of zero and non-zero momentum data indeed reduces the statistical error, in particular, for the η' mass. This is in part due to the fact that the zero momentum data couple to the slowly fluctuating topological charge density and exhibit longer autocorrelations, see [48]. In total, we fit to $l \times M(M+1)/2$ correlation functions and the number of fit parameters is increased to $(l \times M + 1) \times N$, where l is the number of momenta. Specifically, for $M = N = 3$, by setting $l = 2$, we increase the number of correlation functions from 6 to 12 and the number of fit parameters from 12 to 21.

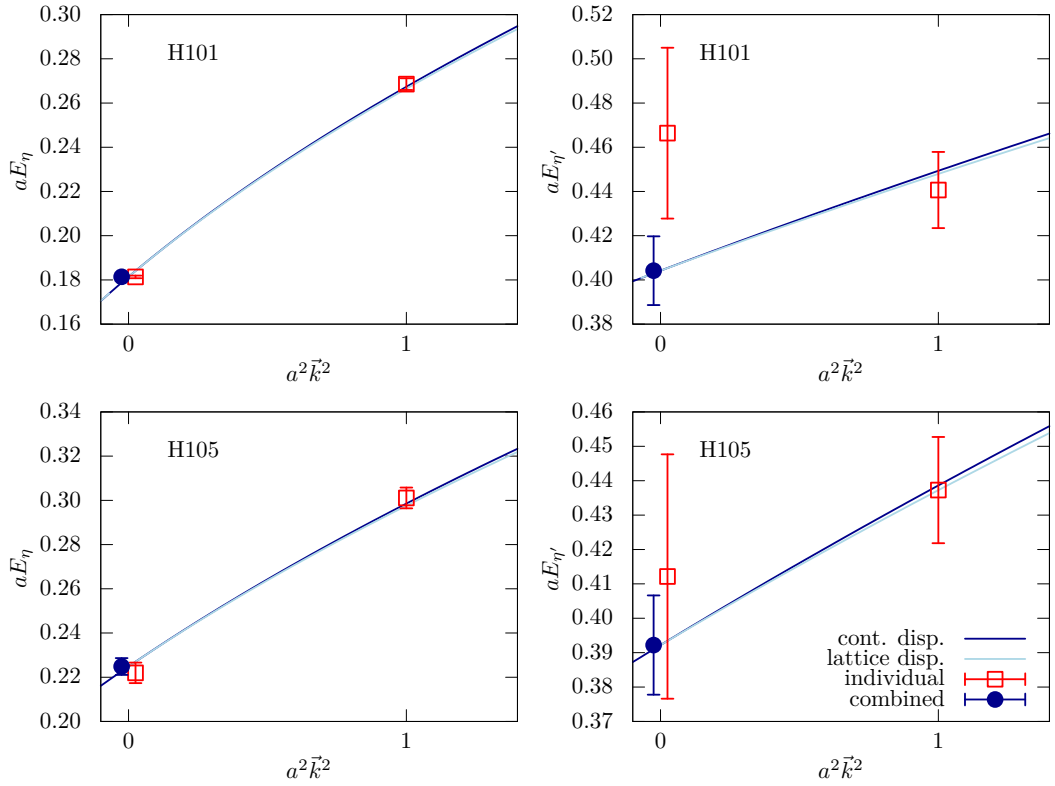


FIGURE 4.3: Energies of the η (left panels) and η' (right panels) mesons determined on ensembles H101 (top) and H105 (bottom). The red squares display the energies with $a^2\vec{k}^2 = 0$ and $a^2\vec{k}^2 = 1$ extracted from individual fits, while the blue filled symbols show the masses determined from a combined fit assuming the continuum dispersion relation (dark blue). The lattice dispersion relation (light blue) obtained using the masses extracted from the combined fit is also displayed. The data points at $a^2\vec{k}^2 = 0$ have been shifted slightly for better visibility. Note that $E_\eta = E_\pi$ on the symmetric ($m_s = m_\ell$) ensemble H101.

Generalized effective masses

The fit form involves a sum over N exponentials for each of the $M(M+1)/2$ independent entries of $C(t)$. As the number of states (N) included increases, the fits become more unstable and sensitive to the choice of the initial guesses. This motivates us to define a matrix analogue of the effective mass (for $N = M$),

$$\partial_t \log C(t) = (\partial_t C(t)) C^{-1}(t) \tag{4.26}$$

$$= \left(\widehat{Z}\partial_t \widehat{D}(t)\widehat{Z}^\intercal\right)\left(\widehat{Z}\widehat{D}(t)\widehat{Z}^\intercal\right)^{-1} \tag{4.27}$$

$$= -\,Z\ \mathrm{diag}_{n=0}^{N-1}(E_n)\, Z^{-1} + \mathcal{O}\left[\exp\left(-\left(E_N - E_{N-1}\right)t\right)\right], \tag{4.28}$$

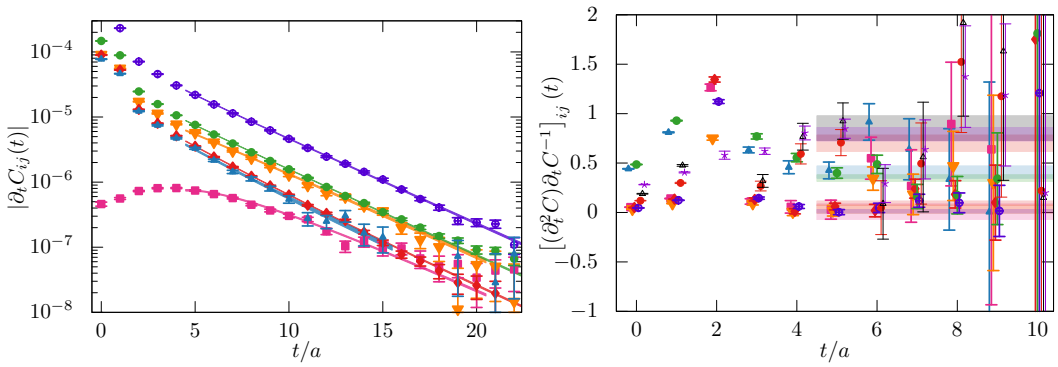


FIGURE 4.4: **Left panel**: The moduli of $C_{ij}(t)$ determined on ensemble H105 at momentum $a^2\vec{k}^2 = 1$. **Right panel**: The elements of the generalized effective mass matrix, eq. (4.29). The shaded regions in both panels correspond to the results of a simultaneous fit to eqs. (4.21) and (4.29). The widths of the regions indicate the fit ranges.

which is constant in time (up to excited states corrections and statistical noise).[2] Since $C(t)$ is non-singular, $(\partial_t C)C^{-1}$ is an unambiguous expression and can readily be computed. One can easily repeat this procedure and take the second derivative, leading to

$$(\partial_t^2 C)(\partial_t C)^{-1} = -Z\ \mathrm{diag}(E_n)\, Z^{-1} + \mathcal{O}\left[\exp\left(-\left(E_N - E_{N-1}\right) t\right)\right]. \tag{4.29}$$

This allows the use of the derivatives in the inverse and hence again reduces correlations. The contributions from higher excited states are somewhat altered by this but the large-time behaviour remains unaffected.

Note that $(\partial_t C)C^{-1}$ and $(\partial_t^2 C)(\partial_t C)$ are not symmetric and their M^2 elements converge to constant values at large times. In order to resolve N different states, $N(M+1)$ parameters (Z_{in} and E_n) need to be determined. The M^2 asymptotic values are not sufficient for this, unless $N \leq M^2/(M+1)$, which excludes the quadratic case $N = M$.[3] In this case, simultaneous fits are performed to eqs. (4.21) and (4.29), where the latter enables the fit to unambiguously resolve the spectrum of states. We plot an example for such a combined fit in fig. 4.4.

Results for the masses at finite lattice spacing and non-physical quark masses

For the individual ensembles we obtain results in the following way, summarizing our fitting strategy described in the previous paragraphs: we simultaneously fit the correlation functions with two momenta $a^2\vec{k}^2 = 0$ and $a^2\vec{k}^2 = 1$ to eqs. (4.21) and (4.29). Correlations between all entries of $(\partial_t C)C^{-1}$ and $\partial_t^2 C(\partial_t C)^{-1}$ at each time slice are taken into account, whereas correlations between time slices can be neglected due to fitting to derivatives of C (we have checked that this is indeed the case). A typical fit is shown in fig. 4.4. The resulting η and η' masses for all the ensembles are collected in tab. 4.2, along with the χ^2/N_{df} of the fits, where in most cases we achieve $\chi^2/N_{\mathrm{df}} \approx 1$.

[2] We remark that this construction is easily generalizable to the case $N \neq M$, employing a singular value decomposition of $ZD(t)Z^{\intercal}$. It should be noted, however, that the leading truncation errors then depend on the $\min(N, M)$ non-singular values.

[3] Setting $M = N + 1 = 4$ allows to determine all the parameters, however, this choice was found to result in larger errors than a combined fit to eqs. (4.21) and (4.29).

id	basis	$a\,M_\eta$	$\sqrt{8t_0}M_\eta$	$a\,M_{\eta'}$	$\sqrt{8t_0}M_{\eta'}$	χ^2/N_{df}
H101	$\{8^3,\ 8^2,\ 0^3,\ 0^2\}$	0.1814(6)	0.867(3)	0.404(16)	1.931(74)	1.09, 0.86
H102a	$\{\ell^3,\ s^3,\ 8^2\}$	0.1996(44)	0.958(21)	0.413(22)	1.981(107)	0.71
H102b	$\{\ell^3,\ s^3,\ 8^2\}$	0.1989(26)	0.955(13)	0.395(21)	1.896(99)	0.99
H105	$\{\ell^3,\ s^3,\ 8^2\}$	0.2249(38)	1.082(18)	0.392(14)	1.886(69)	1.21
C101	$\{\ell^3,\ s^3,\ 8^2\}$	0.2253(52)	1.089(25)	0.422(19)	2.038(92)	0.71
D150	$\{\ell^3,\ s^3,\ 8^2\}$	0.2280(183)	1.107(89)	0.382(31)	1.854(150)	1.04
H107	$\{\ell^3,\ s^3,\ 8^2\}$	0.2509(46)	1.170(21)	0.433(16)	2.019(73)	0.82
H106	$\{\ell^3,\ s^3,\ 8^2\}$	0.2511(49)	1.193(23)	0.391(14)	1.860(67)	0.85
C102	$\{\ell^3,\ s^3,\ 8^2\}$	0.2396(91)	1.148(44)	0.389(20)	1.864(97)	0.88
B450	$\{8^3,\ 8^2,\ 0^3,\ 0^2\}$	0.1611(17)	0.872(9)	0.357(12)	1.930(67)	1.84, 1.00
S400	$\{\ell^3,\ s^3,\ 8^2\}$	0.1837(28)	0.998(15)	0.335(10)	1.821(52)	1.46
N401	$\{\ell^3,\ s^3,\ 8^2\}$	0.1858(141)	1.009(77)	0.361(25)	1.959(134)	1.14
B451	$\{\ell^3,\ s^3,\ 8^2\}$	0.2386(29)	1.249(15)	0.370(11)	1.937(56)	1.19
B452	$\{\ell^3,\ s^3,\ 8^2\}$	0.2233(29)	1.186(15)	0.355(8)	1.887(42)	1.07
N202	$\{8^3,\ 8^2,\ 0^3,\ 0^2\}$	0.1313(17)	0.844(11)	0.331(20)	2.126(129)	0.87, 2.07
N203	$\{\ell^3,\ s^3,\ 8^2\}$	0.1567(29)	1.005(19)	0.282(22)	1.808(144)	0.68
N200	$\{\ell^3,\ s^3,\ 8^2\}$	0.1711(25)	1.099(16)	0.303(18)	1.948(117)	1.45
D200	$\{\ell^3,\ s^3,\ 8^2\}$	0.1768(22)	1.138(14)	0.330(13)	2.125(85)	1.43
N204	$\{\ell^3,\ s^3,\ s^2\}$	0.1970(35)	1.239(22)	0.315(16)	1.983(103)	1.07
N201	$\{\ell^3,\ s^3,\ 8^2\}$	0.1818(67)	1.154(43)	0.306(9)	1.944(57)	1.80
D201	$\{\ell^3,\ s^3,\ 8^2\}$	0.1874(90)	1.201(58)	0.327(21)	2.097(133)	1.03
N300	$\{8^3,\ 8^2,\ 0^3,\ 0^2\}$	0.1061(11)	0.878(9)	0.247(16)	2.046(130)	1.21, 1.87

TABLE 4.2: Masses of the η and η' mesons obtained from fits to eqs. (4.21) and (4.29) in lattice units and in units of the gradient flow scale, $\sqrt{8t_0}$ (determined on the same ensemble). See tab. 3.1 for the corresponding pion and kaon masses and tab. 3.3 for the lattice spacings. We also give the smearing bases used in the construction of the matrix of correlation functions, eq. (4.14), where $\ell, s, 8, 0$ refer to the light, strange, octet and singlet combinations of the pseudoscalar interpolating operators, respectively, and the superscript labels the smearing applied (element of S), see tab. 4.1. The resulting χ^2/N_{df} of the partially correlated fits are also given. For the ensembles with $m_s = m_\ell$, where we carry out two independent fits, we give both χ^2/N_{df} values for the octet (first) and singlet (second) cases. Ensembles H102a and H102b were generated with the same quark masses and lattice coupling but different simulation parameters and are therefore analysed separately.

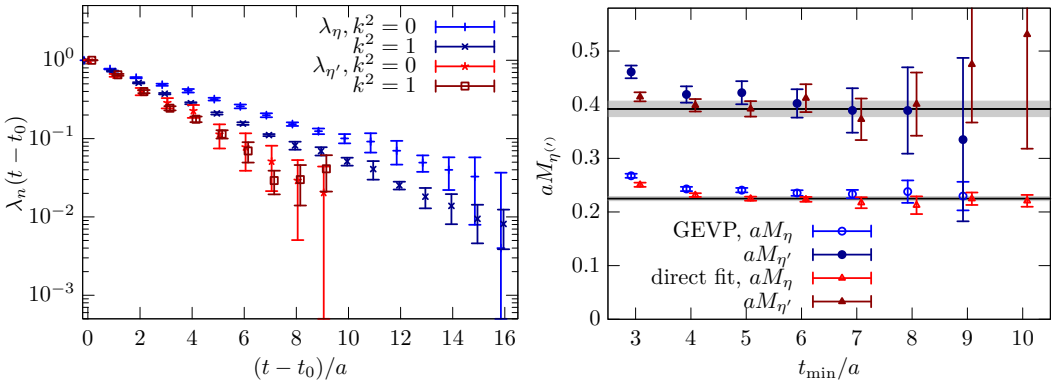


FIGURE 4.5: **Left panel**: The η and η' eigenvalue functions for two momenta, $a^2\vec{k}^2 = 0$ and $a^2\vec{k}^2 = 1$ determined on ensemble H105, obtained by solving the GEVP, with the reference time slice $t_0 = 5a$. For better visibility, some points are shifted slightly and the data for the η' are omitted for $(t-t_0)/a \geq 10$ due to the large errors. **Right panel**: Results for the masses determined by fitting to the lowest two eigenvalues of the GEVP and from direct fits (eqs. (4.21) and (4.29)) as functions of the starting point of the fit, $t_{\min}$. The analysis is similar in both cases, employing the same basis of interpolators and incorporating data at two momenta. The horizontal lines and grey error bands indicate the final results. These have been obtained from a slightly different fit, employing different $t_{\min}$ for different elements of the correlation matrix. For the GEVP results, $t_0/a = t_{\min}/a - 1$.

Comparison to the GEVP method

A standard way to extract the masses of the η and η' from the matrix of correlation functions is to solve the GEVP [176, 177],

$$C(t)V(t,t_0) = C(t_0)V(t,t_0)\Lambda(t,t_0), \tag{4.30}$$

where $\Lambda = \mathrm{diag}(\lambda_0, \ldots, \lambda_{M-1})$ is the diagonal matrix of the eigenvalues and V is the matrix of eigenvectors. One then fits to the eigenvalues $\lambda_n \propto e^{-E_n t}$ to extract the energies.

The reference timeslice t_0 needs to be chosen large enough [179] so that contributions from states with $n \geq M$ are sufficiently suppressed. In our setup, we found it hard to disentangle the excited state contributions from the lowest two eigenvalues, having only a limited number of timeslices $t > t_0$ available before the signal of the heavier η' vanishes in the statistical noise. In particular at larger times, it also becomes increasingly difficult to assign the correct physical states to the eigenvectors. We compare the GEVP with the results obtained from the fit strategy described in the previous section in fig. 4.5. While the two methods generally agree, the plateau regions start earlier when using the generalized effective mass fit method. This enables us to extract results with an increased statistical precision compared to using the GEVP. We remark that our fit method also allows us to extract amplitudes directly, in a straightforward manner, as is discussed below.

4.3 Local matrix elements and determination of decay constants

Having determined mass eigenstates of the η and η' in the previous section, given by the fitted amplitudes Z_{in} depending on the set of interpolating operators employed, we can now proceed to compute local matrix elements with a local current J that are needed to extract decay constants. We start from a vector of M correlation functions ($i = 1, \ldots, M$):

$$C_i^J(t) = \left\langle \Omega \left| J(t)\mathcal{B}_i^\dagger(0) \right| \Omega \right\rangle, \tag{4.31}$$

where $\mathcal{B}_i(0)$ is an interpolator with the momentum $\vec{p}$ inserted at the time $t_{\text{in}} = 0$ and is taken to be the same as that used to fit the eigenstates in the previous sections. For the local currents we employ those defined in eqs. (2.112) and (2.111), where we set the unknown mass-independent improvement coefficients to zero (i.e. we use $\check{P}^0$ and $\check{A}^0_\mu$ instead of $\tilde{P}^0$ and $\tilde{A}^0_\mu$, respectively).

For the connected contribution we utilize the translational invariance of the expectation value to move the momentum projection from the smeared point source to the local sink, as is common in this kind of calculation. For the disconnected two-point function, in order to increase the statistics, we replace $J(t,\vec{0}) \mapsto (a^3/V_3)\sum_{\vec{x}} e^{-i\vec{p}\cdot\vec{x}} J(t,\vec{x})$, again exploiting translational invariance. The two-point function is then constructed in analogy to eqs. (4.7) and (4.12), however, without smearing at the sink and with the additional normalization factor $1/V_3$.

We carry out a spectral decomposition similar to that of eq. (4.14):

$$\begin{aligned} C_i^J(t) &\approx \sum_{n=0}^{N-1} \frac{1}{2E_n V_3} \exp(-E_n t) \left\langle \Omega \left| J(0) \right| n \right\rangle \left\langle n \left| \mathcal{B}_i^\dagger(0) \right| \Omega \right\rangle \\ &= \sum_{n=0}^{N-1} Z_{in} D_{nn}(t) j_n, \end{aligned} \tag{4.32}$$

where

$$j_n = \frac{1}{\sqrt{2E_n V_3}} \langle \Omega | J | n \rangle, \tag{4.33}$$

and Z is the truncated overlap matrix, defined in eq. (4.17). One can also write this in terms of matrix multiplications, $C^J(t) \approx ZD(t)j$, where $C^J(t)$ and j are M- and N-dimensional vectors, respectively. Using the bootstrap samples of the previously obtained elements of Z and energies E_n, we carry out a fit to the above functional form, determining the matrix elements $\langle \Omega | J | n \rangle = \sqrt{2E_n V_3}\, j_n$.

Once the axialvector and pseudoscalar matrix elements are obtained, we can construct the partially improved, unrenormalized decay constants $\tilde{F}^8_n$ and $\check{F}^0_n$ for $n \in \{\eta, \eta'\}$ as well as the corresponding pseudoscalar matrix elements $\tilde{H}^8_n$ and $\check{H}^0_n$, defined in eqs. (2.113) and (2.115). The improvement of these has been discussed in sec. 2.4.2 and the remaining unknown improvement parameters $f_J, d_J, \tilde{d}_J, \delta c_J$ and c_J^s for $J \in \{A, P\}$ will be included in the corresponding continuum limit fits. We list these partially improved decay constants multiplied by the corresponding renormalization factor (see sec. 2.4.3) in tab. 4.3.

id	$\sqrt{8t_0}Z_A\tilde{F}^8_\eta$	$\sqrt{8t_0}Z_A\tilde{F}^8_{\eta'}$	$\sqrt{8t_0}Z^{s\prime}_A\check{F}^0_\eta$	$\sqrt{8t_0}Z^{s\prime}_A\check{F}^0_{\eta'}$
H101	0.2164(30)	0	0	0.1790(97)
H102a	0.2206(58)	−0.0263(104)	0.0142(100)	0.1701(314)
H102b	0.2138(32)	−0.0303(54)	0.0102(37)	0.1816(97)
H105	0.2155(27)	−0.0610(78)	0.0215(66)	0.1890(54)
C101	0.2098(55)	−0.0744(155)	0.0299(87)	0.2046(126)
D150	0.1954(193)	−0.1030(374)	0.0198(194)	0.1472(277)
H107	0.2230(42)	−0.0730(84)	0.0306(79)	0.1921(47)
H106	0.2159(46)	−0.0741(115)	0.0252(91)	0.1780(48)
C102	0.2092(83)	−0.0928(170)	0.0502(176)	0.1906(136)
B450	0.2184(30)	0	0	0.1947(68)
S400	0.2184(32)	−0.0384(46)	0.0178(80)	0.1923(74)
N401	0.2187(76)	−0.0643(151)	0.0177(158)	0.1686(197)
B451	0.2317(30)	−0.0668(64)	0.0298(82)	0.2132(134)
B452	0.2210(26)	−0.0870(57)	0.0438(51)	0.1960(37)
N202	0.2180(37)	0	0	0.1793(60)
N203	0.2236(32)	−0.0361(92)	0.0190(71)	0.2130(226)
N200	0.2238(29)	−0.0457(98)	0.0189(89)	0.1961(68)
D200	0.2243(30)	−0.0684(69)	0.0161(60)	0.1981(90)
N204	0.2297(53)	−0.0671(114)	0.0172(110)	0.2069(149)
N201	0.2198(34)	−0.0854(108)	0.0372(120)	0.1872(55)
D201	0.2299(102)	−0.0753(268)	0.0184(131)	0.1753(238)
N300	0.2124(32)	0	0	0.1781(171)

TABLE 4.3: Renormalized and partially improved octet and singlet decay constants of the η and η' mesons obtained from fits to eq. (4.32) in units of the gradient flow scale, $\sqrt{8t_0}$ (determined on the same ensemble). Ensembles H102a and H102b were generated with the same quark masses and lattice coupling but different simulation parameters and are therefore analysed separately.

5 Masses and decay constants of the η and η' mesons

The results at the individual ensembles determined in the previous chapter are determined at both a non-vanishing lattice spacing and mostly unphysical quark masses. In this chapter we combine our results of the masses and decay constants and extrapolate them to the continuum and to the physical quark mass point. The renormalization and linear improvement of our observables has been presented in sec. 2.4.

First, in sec. 5.1, we summarize literature results of large-N_c ChPT that will be used to parametrize the quark mass dependence of the η and η' masses and decay constants when performing the extrapolation to the physical point. We first present the general framework, before giving explicit expressions to LO and NLO in the following subsections 5.1.1 and 5.1.2, respectively. We then add parametrizations of lattice spacing effects in sec. 5.2, including the unknown linear improvement coefficients and higher quadratic terms. The choice of free parameters in that fit, as well as truncation of the ChPT fit forms and renormalization comprise the systematic errors that we attempt to quantify by doing a number of different fits in sec. 5.3. Of these fits we then single out one particular that we use to quote our central values and present results for the masses and decay constants in sec. 5.4, including results for the low energy constants (LECs) that parametrize our results in the continuum in sec. 5.4.1. Finally, we compare our results to the phenomenological and lattice literature available in sec. 5.5 and give an exemplary application for the asymptotic values of the transition form factors $\eta^{(\prime)} \to \gamma^*\gamma$ in sec. 5.5.4.

This chapter has already been published in similar or verbatim form in [60].

5.1 Large-N_c chiral perturbation theory

Conventional SU(3) ChPT entails expansions in the masses of the octet mesons (the pions, the kaons and the octet η_8) — the Goldstone bosons of the spontaneous breaking of $\mathrm{SU}(3)_A$ in the QCD vacuum, see sec. 2.1.6 for a more detailed discussion. To include the singlet η_0, one extends the symmetry group to U(3) and expands simultaneously around the limit $N_c \to \infty$ In this limit the axial anomaly vanishes but at finite N_c, the singlet state acquires its anomalous mass. Therefore, in U(3) large-N_c ChPT, the expansion is organized in powers of δ [180, 181], where the power counting is as follows:

$$p = \mathcal{O}(\sqrt{\delta}), \qquad m = \mathcal{O}(\delta), \qquad 1/N_c = \mathcal{O}(\delta) \tag{5.1}$$

with p being the momentum and m a quark mass, see tab. I of [181] for a complete list of the power counting rules.

The chiral Lagrangian at $\mathcal{O}(\delta^0)$ corresponds to massless QCD with an infinite number of colours. At LO, i.e. $\mathcal{O}(\delta^1)$, without the anomaly, the squared pseudoscalar masses μ_a^2 are related to the quark mass matrix in the adjoint representation via the GMOR relations, cf. eq. (3.4),

$$\mu_{ab}^2 = 2B_0 \,\mathrm{tr}[t^a \,\mathrm{diag}(m_\ell, m_\ell, m_s) t^b] = \delta^{ab}\mu_a^2. \tag{5.2}$$

However, at this order in δ one also has to add the $\mathcal{O}(1/N_c)$ Witten–Veneziano contribution [12, 14] to the singlet mass, $M_0^2 = 2N_f\tau_0/F^2$, where τ_0 denotes the quenched topological susceptibility. The mixing of the singlet ($a = 0$) and octet ($a = 8$) into physical states is in alignment with the discussion in sec. 2.1.6 and the corresponding non-diagonal part of the pseudoscalar mass matrix reads, see, e.g., [180, 181]:

$$\mu^2 = \begin{pmatrix} \mu_8^2 & \mu_{80}^2 \\ \mu_{80}^2 & \mu_0^2 \end{pmatrix}. \tag{5.3}$$

Its eigenvalues correspond to the (squared) η and η' masses:

$$R(\theta)\mu^2 R^{\intercal}(\theta) = \begin{pmatrix} M_\eta^2 & 0 \\ 0 & M_{\eta'}^2 \end{pmatrix}. \tag{5.4}$$

$R(\theta)$ is an orthogonal transformation as defined in eq. (2.45) which defines the so-called mass mixing angle θ. One can easily read off the relations

$$\mu_8^2 = M_\eta^2 \cos^2\theta + M_{\eta'}^2 \sin^2\theta, \tag{5.5}$$
$$\mu_0^2 = M_\eta^2 \sin^2\theta + M_{\eta'}^2 \cos^2\theta, \tag{5.6}$$
$$\mu_{80}^2 = (M_{\eta'}^2 - M_\eta^2)\sin\theta\cos\theta, \tag{5.7}$$
$$\theta = \frac{1}{2}\arcsin\left(\frac{2\mu_{80}^2}{\sqrt{(\mu_8^2-\mu_0^2)^2+4\mu_{80}^4}}\right) = \frac{1}{2}\arcsin\left(\frac{2\mu_{80}^2}{M_-^2}\right), \tag{5.8}$$

where

$$M_-^2 = M_{\eta'}^2 - M_\eta^2 = \sqrt{(\mu_8^2-\mu_0^2)^2+4\mu_{80}^4}, \tag{5.9}$$
$$M_+^2 = M_{\eta'}^2 + M_\eta^2 = \mu_8^2 + \mu_0^2. \tag{5.10}$$

The above relations apply to all orders in ChPT, however, the dependencies of the mass matrix parameters μ_8, μ_0 and μ_{80} on the masses of the η and η' mesons and the chiral anomaly vary with the order of the expansion. Also, the GMOR relations between these parameters and the quark masses are subject to NLO corrections. We will utilize the combinations

$$M_\eta = \sqrt{\frac{1}{2}\left(M_+^2 - M_-^2\right)} \quad \text{and} \quad M_{\eta'} = \sqrt{\frac{1}{2}\left(M_+^2 + M_-^2\right)} \tag{5.11}$$

when performing the extrapolation of the η and η' masses to the physical point.

The functions μ_8, μ_0 and μ_{80} depend on low energy parameters and quark masses, with the latter typically being replaced by combinations of the pion and kaon masses via the GMOR relations. Along one of the two trajectories that we consider here, the average quark mass is constant and therefore, a more convenient parametrization is in terms of the average and difference of the squared pion and kaon masses, $\overline{M^2}$ and δM^2, cf. eqs. (3.2) and (3.4).

The computation of the decay constants is more involved. A common parametrization is that of the two-angle mixing scheme, where the four physical decay constants are expressed in terms of two angles θ_0 and θ_8 and two constants F^0 and F^8 [115, 181] (recalling eq. (2.100)),

$$\begin{pmatrix} F_\eta^8 & F_\eta^0 \\ F_{\eta'}^8 & F_{\eta'}^0 \end{pmatrix} = \begin{pmatrix} F^8\cos\theta_8 & -F^0\sin\theta_0 \\ F^8\sin\theta_8 & F^0\cos\theta_0 \end{pmatrix}, \tag{5.12}$$

leading to

$$F^8 = \sqrt{(F^8_\eta)^2 + (F^8_{\eta'})^2}, \qquad F^0 = \sqrt{(F^0_\eta)^2 + (F^0_{\eta'})^2}, \tag{5.13}$$

$$\tan\theta_8 = \frac{F^8_{\eta'}}{F^8_\eta}, \qquad \tan\theta_0 = -\frac{F^0_\eta}{F^0_{\eta'}}. \tag{5.14}$$

The decay constants in the flavour basis can be expressed in the same way,

$$F^s = \sqrt{(F^s_\eta)^2 + (F^s_{\eta'})^2}, \qquad F^\ell = \sqrt{(F^\ell_\eta)^2 + (F^\ell_{\eta'})^2}, \tag{5.15}$$

$$\tan\phi_s = -\frac{F^s_{\eta'}}{F^s_\eta}, \qquad \tan\phi_\ell = \frac{F^\ell_\eta}{F^\ell_{\eta'}}. \tag{5.16}$$

The latter is a popular choice in phenomenological studies due to the fact that $\phi_\ell \approx \phi_s$ at the physical point, which allows one to express all four decay constants in terms of only three parameters [115].

5.1.1 Fit forms for LO

As explained above, at leading order the elements of the pseudoscalar mass matrix are linear in the quark masses and can be related to combinations of the non-singlet pseudoscalar meson masses via the LO GMOR relations

$$(\mu_8^{\mathrm{LO}})^2 = \frac{2}{3}B_0\,(m_\ell + 2m_s) = \overline{M}^2 + \frac{1}{3}\delta M^2, \tag{5.17}$$

$$(\mu_0^{\mathrm{LO}})^2 = \frac{2}{3}B_0\,(2m_\ell + m_s) + M_0^2 = \overline{M}^2 + M_0^2, \tag{5.18}$$

$$(\mu_{80}^{\mathrm{LO}})^2 = -\frac{2\sqrt{2}}{3}B_0\,(m_s - m_\ell) = -\frac{\sqrt{2}}{3}\delta M^2, \tag{5.19}$$

where the anomalous contribution $M_0^2 = 6\tau_0/F^2$ is proportional to the quenched topological susceptibility τ_0 [12, 14] and contributes at $\mathcal{O}(N_c^{-1})$ to μ_0^2, while $\overline{M}^2$ is the $\mathcal{O}(m)$ value of the squared singlet mass.

To this order, all singlet and octet decay constants can be expressed in terms of the pion decay constant F (in the chiral limit) and the angle θ, defined in eq. (5.8):

$$F^8_\eta = F^0_{\eta'} = F\cos\theta, \qquad -F^0_\eta = F^8_{\eta'} = F\sin\,\theta, \tag{5.20}$$

i.e. $F^8 = F^0 = F$ and $\theta_8 = \theta_0 = \theta$. Note that to this order θ only depends on M_0^2 and δM^2. A single mixing angle in the octet/singlet basis is not consistent with phenomenological investigations [115, 181] and also the results of the present study clearly show $F^8_\eta \neq F^0_{\eta'}$ and $F^0_\eta \neq -F^0_{\eta'}$.

5.1.2 Fit forms for NLO

The large-N_c ChPT expansion for the masses and decay constants has been worked out to NNLO in [182]. Here, we use the results of [181] and truncate these at NLO. To this order, only four additional LECs, L_5, L_8, $\Lambda_1(\mu)$ and $\Lambda_2(\mu)$ appear. The elements of the squared mass matrix

are given by

$$(\mu_8^{\mathrm{NLO}})^2 = (\mu_8^{\mathrm{LO}})^2 + \frac{8}{3F^2}\,(2L_8 - L_5)\,\delta M^4, \tag{5.21}$$

$$(\mu_0^{\mathrm{NLO}})^2 = (\mu_0^{\mathrm{LO}})^2 + \frac{4}{3F^2}\,(2L_8 - L_5)\,\delta M^4 - \frac{8}{F^2}L_5\overline{M}^2 M_0^2 - \tilde{\Lambda}\overline{M}^2 - \Lambda_1 M_0^2, \tag{5.22}$$

$$(\mu_{80}^{\mathrm{NLO}})^2 = (\mu_{80}^{\mathrm{LO}})^2 - \frac{4\sqrt{2}}{3F^2}\,(2L_8 - L_5)\,\delta M^4 + \frac{4\sqrt{2}}{3F^2}L_5 M_0^2 \delta M^2 + \frac{\sqrt{2}}{6}\tilde{\Lambda}\delta M^2, \tag{5.23}$$

where we substituted $\tilde{\Lambda} = \Lambda_1 - 2\Lambda_2$. To NLO of the chiral expansion the latter combination does not depend on the QCD renormalization scale μ [181]. In general the LECs can depend both on the QCD scale, due to the anomalous dimension of the singlet decay constants, and the ChPT renormalization scale, due to loop corrections. However, in large-N_c ChPT loop corrections are suppressed by a factor of δ^2 and, hence, the LECs are independent of the ChPT scale at NLO.

Substituting the leading order pseudoscalar masses and decay constant by the NLO expressions [181] and expressing everything in terms of $\overline{M}$ and δM, the decay constants are given by

$$F_\eta^8 = F\left[\cos\theta + \frac{4L_5}{3F^2}\left(3\cos\theta\overline{M}^2 + (\sqrt{2}\sin\theta + \cos\theta)\delta M^2\right)\right], \tag{5.24}$$

$$F_{\eta'}^8 = F\left[\sin\theta + \frac{4L_5}{3F^2}\left(3\sin\theta\overline{M}^2 + (\sin\theta - \sqrt{2}\cos\theta)\delta M^2\right)\right], \tag{5.25}$$

$$F_\eta^0 = -F\left[\sin\theta\left(1 + \frac{\Lambda_1}{2}\right) + \frac{4L_5}{3F^2}\left(3\sin\theta\overline{M}^2 + \sqrt{2}\cos\theta\delta M^2\right)\right], \tag{5.26}$$

$$F_{\eta'}^0 = F\left[\cos\theta\left(1 + \frac{\Lambda_1}{2}\right) + \frac{4L_5}{3F^2}\left(3\cos\theta\overline{M}^2 - \sqrt{2}\sin\theta\delta M^2\right)\right], \tag{5.27}$$

where θ is the mass mixing angle defined in eq. (5.8), evaluated with the entries of the NLO mass matrix, eqs. (5.21)–(5.23). Note that in the standard $\overline{\mathrm{MS}}$ scheme Λ_1 as well as F_η^0 and $F_{\eta'}^0$ depend on μ. In general, $\theta_8 \neq \theta_0 \neq \theta$ to this order.

In sec. 3.4.1 we motivate our general strategy for continuum limit extrapolation, taking dimensionless combination with the gradient flow scale $8t_0/a^2$. Accordingly, we multiply the entries of the mass matrix eqs. (5.21) to (5.23) by $8t_0$ and the decay constants eqs. (5.24) to (5.27) by $8\sqrt{t_0}$. The values of $8t_0/a^2$ were determined on the same ensemble, see tab. 3.1 for their values. To NLO in the continuum the gradient flow scale also depends slightly on the average pseudoscalar mass $\overline{M}$, see eq. (3.16), and to be consistent we need to take these corrections into account. The dimensionful LECs are then defined directly in the chiral limit, $8t_0^\chi M_0^2$ and $\sqrt{8t_0^\chi}F$, and we relate the expression to that point by replacing $8t_0^\chi M_0^2 = 8t_0 M_0^2(1 - \tilde{k}_1\, 8t_0\overline{M}^2)$, and $\sqrt{8t_0^\chi}F = \sqrt{8t_0}F(1 - \frac{\tilde{k}_1}{2}\, 8t_0\overline{M}^2)$, respectively. Therefore, we have to add a term $-\tilde{k}_1 M_0^2\overline{M}^2 8t_0$ to eq. (5.22). Regarding the decay constants, the term $-(k/2)F\cos\theta\overline{M}^2 8t_0$ needs to be added to eqs. (5.24) and (5.27) while the term $-(k/2)F\sin\theta\overline{M}^2 8t_0$ has to be added to eq. (5.25) and subtracted from eq. (5.26).

In eq. (3.42) we determined $\tilde{k}_1 = -0.0466(62)$, and therefore the effect of including the expansion of t_0^χ is a small effect. Regarding the LECs, the biggest effect is on F, L_5 and L_8, which decrease by $2.9(4)$ MeV, by $7.7(1.0)\cdot 10^{-5}$ and by $6.2(8)\cdot 10^{-5}$, respectively, which is well below the total errors that we find for these parameters: 4.8 MeV, $2.1\cdot 10^{-4}$ and $1.4\cdot 10^{-4}$.

5.2 Continuum limit parametrization

The lattice data do not only depend on the quark masses but also on the lattice spacing. Here we outline our continuum limit extrapolation procedure. We shall label the ChPT functional forms given above as $f_O^{\mathrm{cont}}(\overline{M}^2, \delta M^2 | \ldots)$ where the ellipses represent the fit parameters (i.e. the LECs)

and O can be either of the two masses or four decay constants. We remind the reader that for the decay constants not all the $\mathcal{O}(a)$ improvement coefficients are known, see sec. 2.4.2. Therefore, we start from the following ansatz

$$f_O(a,\overline{M}^2,\delta M^2) = f_O^{\text{cont}}(\overline{M}^2,\delta M^2|\ldots)h_O^{(1)}(a,am_\ell,am_s|\ldots)h_O^{(2)}(a^2,a^2\overline{M}^2,a^2\delta M^2|\ldots), \tag{5.28}$$

where $h_O^{(1)}$ contains the linear lattice spacing effects with known or unknown coefficients and $h_O^{(2)}$ is a quadratic function of a. The input data for the fits to the decay constants are the partially improved combinations $\tilde{F}_n^8$ and $\check{F}_n^0$ (see eqs. (2.113) and (2.112)) with $n = \eta,\eta'$. In terms of the linear lattice spacing effects, for $O = M_\eta, M_{\eta'}$, $h_O^{(1)} = 1$, while for the octet decay constants the linear improvement has been discussed in sec. 2.4.2. We recall that the functions $h_{F_n^8}^{(1)}$ contain the known parameters b_A, $\tilde{b}_A$ and the free parameter f_A while in the singlet case within $h_{F_n^0}^{(1)}$ the unknown parameters d_A, $\tilde{d}_A$ and δc_A appear (see eqs. (2.124) and (2.119)). For $h_O^{(2)}$ we make a generic quadratic ansatz. Explicit formulae will be given below.

Our input data are transformed into dimensionless units: $M_n \mapsto \sqrt{8t_0}M_n$, $\check{F}_n^0 \mapsto \sqrt{8t_0}\check{F}_n^0$ and $\tilde{F}_n^8 \mapsto \sqrt{8t_0}\tilde{F}_n^8$, where the scale t_0 is obtained on the same ensemble. Moreover, the parametrizations for the unrenormalized decay constants need to be divided by Z_A and Z_A^s, respectively. The lattice spacing is given in units of t_0^*: $a \mapsto a/\sqrt{8t_0^*}$ (see sec. 3.4 and tab. 3.3). The six parametrizations share the LECs and some of the improvement coefficients. Hence, we carry out simultaneous fits to all these data. Results at the physical point can be obtained by evaluating the continuum limit functions at that point, eq. (3.1). For the linear lattice effects on the octet and singlet decay constants, we combine the results of eqs. (2.120), (2.124) and (2.119), to obtain the functions

$$\begin{aligned} h_{F_n^8}^{(1)}(a|f_A^l) = 1 - 3a\tilde{b}_A\overline{m} - ab_A\frac{Z_A}{\sqrt{3}}\frac{m_\ell\check{F}_n^\ell - \sqrt{2}m_s\check{F}_n^s}{f_{F_n^8}^{\text{cont}}(\overline{M}^2,\delta M^2)} \\ + \sqrt{2}af_AZ_A\frac{\delta m\check{F}_n^0}{f_{F_n^8}^{\text{cont}}(\overline{M}^2,\delta M^2)}, \end{aligned} \tag{5.29}$$

$$\begin{aligned} h_{F_n^0}^{(1)}(a|d_A^l,\tilde{d}_A^l,\delta c_A^l) = 1 - 3a\tilde{d}_A\overline{m} - ad_A\frac{Z_A^s}{\sqrt{3}}\frac{\sqrt{2}m_\ell\check{F}_n^\ell + m_s\check{F}_n^s}{f_{F_n^0}^{\text{cont}}(\overline{M}^2,\delta M^2)} \\ - a\delta c_AZ_A^s\frac{\check{H}_n^0}{f_{F_n^0}^{\text{cont}}(\overline{M}^2,\delta M^2)}, \end{aligned} \tag{5.30}$$

where $n = \eta,\eta'$ and the flavour decay constants are constructed from the singlet and octet ones via the rotation eq. (2.103). We have substituted the data on the decay constants $\tilde{F}_n^8$ and $\check{F}_n^0$ by the fitted continuum limit parametrizations $f_{F_n^a}^{\text{cont}}(\overline{M}^2,\delta M^2)$, which enables us to include data points where the denominator is small and hence carries a large relative error. This replacement is admissible since the difference is of $\mathcal{O}(a^2)$. Note that in ansatz (5.28) $h_O^{(1)}$ is multiplied by $f_{F_n^a}^{\text{cont}}(\overline{M}^2,\delta M^2)$. Above, we suppressed the dependence of the improvement coefficients on g^2. The only unknown functions are $f_A(g^2)$, $d_A(g^2)$, $\tilde{d}_A(g^2)$ and $\delta c_A(g^2)$ and we parametrize these as follows

$$f_A(g^2) = f_A^lg^6, \quad d_A(g^2) = b_A(g^2) + d_A^lg^4, \quad \tilde{d}_A(g^2) = \tilde{d}_A^lg^4, \quad \delta c_A(g^2) = \delta c_A^lg^4, \tag{5.31}$$

such that only f_A^l, d_A^l, $\tilde{d}_A^l$ and c_A^l appear as free parameters on the left hand sides of eqs. (5.29) and (5.30). The above powers of g^2 correspond to the first non-trivial orders of the perturbative expansions, see the discussion in sec 2.4.2.

Turning to the quadratic lattice effects and the functions $h_O^{(2)}$, we allow for three more fit parameters per observable O:

$$h_O^{(2)}(a^2, t_0\overline{M}^2, t_0\delta M^2 | l_O, m_O, n_O) = 1 + a^2\left(l_O + m_O\overline{M}^2 + n_O\delta M^2\right). \tag{5.32}$$

The terms multiplied by l_O, m_O and n_O correspond to lattice spacing effects proportional to $a^2\Lambda^2$, $a^2\Lambda(2m_\ell + m_s)$ and $a^2\Lambda(m_s - m_\ell)$, respectively, where $\Lambda \gg m_s \geq m_\ell$ is the QCD scale. Due to this hierarchy of scales, other quadratic lattice spacing effects depending solely on the quark masses like, for example, $a^2 m_\ell^2 \approx a^2 M_\pi^4/(4\,B_0^2)$, are neglected. We remark that for the non-singlet pseudoscalar decay constants significant $\mathcal{O}(a^2)$ effects have been reported in lattice results determined using our action [135].

In summary, in the simultaneous fits of the two masses and four decay constants a total of four parameters are needed to account for the linear cut-off effects and $6 \times 3 = 18$ more coefficients to parametrize the a^2-effects. These are in addition to the LECs M_0, F_0, L_5, L_8, Λ_1 and $\tilde{\Lambda}$ that appear in the continuum expressions. As will be discussed in the next subsection, most of the lattice spacing terms cannot be resolved in our data and the corresponding coefficients will be set to zero in the fits that we use to determine the final results.

5.3 Estimation of systematic errors

We now use the rather complicated fit forms (5.28) to carry out a series of fits, identifying and fixing irrelevant parameters to zero. We now describe this procedure, determining which fit parameters are most relevant and how we estimate the systematic uncertainty associated with the chosen set of fit forms. We also repeat fits varying the included ensembles to quantify ChPT cut-off effects and for varying the scale where we match the renormalization scale of the singlet to perturbation theory.

Each fit is performed simultaneously to the six observables determined on ensembles which lie on two trajectories in the quark mass plane and span four lattice spacings. Correlations between the η and η' masses and the decay constants as well as the arguments of the fit function ($8t_0\overline{M}^2$, $8t_0\delta M^2$) on each ensemble are taken into account, see app. B.3. The fits are performed on the ensemble averages of the data and the statistical uncertainties in the fit parameters are obtained by repeating the fit on 500 bootstrap samples. The statistical uncertainty is taken to be the interval that contains the central 68.3 % of the 500 bootstrap values of each parameter.

Since the simulations are performed in a finite box, there is the possibility of finite volume effects. Such correction to M_η have been computed in ChPT [141] and are suppressed with respect to M_π by M_π^2/M_η^2. This means that at the symmetric points where very large volumes are available we expect the same size of volume effects as for the pion while towards the physical point this decreases to only about 7 % of that of the pion. Since for the latter on all the analyzed ensembles only very moderate finite size effects were observed even for the physical point ensemble D150 [90] and spatial extents of $L_s M_\pi \gtrsim 4$ are realized for most of the analyzed ensembles, finite volume effects can safely be neglected within our statistical precision.

In terms of the lattice spacing effects, in a first step we establish which terms in the fit forms presented in the previous subsection can be resolved. We start with fits to all data employing the NLO large-N_c ChPT continuum limit parametrization and only include $O(a)$ terms with non-perturbatively determined coefficients, i.e. those involving b_A, $\tilde{b}_A$ and c_A. All $O(a^2)$ coefficients are omitted. For this reference fit we obtain $\chi^2/N_{\text{df}} \approx 220/126 \approx 1.75$. Additional discretization terms are subsequently included and those fits for which the coefficients can be resolved with reasonable precision are given in tab. 5.1. The reference fit has the id "1" in the table. The LECs extracted from these fits are collected in tab. 5.2 and the results for the masses and decay constants at the physical point are detailed in tab. 5.3. The coefficients of the discretization terms are provided in app. C.

id	f_A^l	d_A^l	$\tilde{d}_A$	$\delta^l c_A$	$l_{F_\eta^8}$	$n_{F_\eta^8}$	$l_{F_{\eta'}^8}$	$m_{F_{\eta'}^8}$	$n_{F_{\eta'}^8}$	$l_{F_\eta^0}$	$m_{F_\eta^0}$	$n_{F_\eta^0}$	$l_{F_{\eta'}^0}$	$m_{F_{\eta'}^0}$	$n_{F_{\eta'}^0}$
1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
2	×	×	×	×	—	—	—	—	—	—	—	—	—	—	—
3	×	×	×	—	—	—	—	—	—	—	—	—	—	—	—
4	×	×	—	—	—	—	—	—	—	—	—	—	—	—	—
5	×	×	—	—	×	—	×	—	—	—	—	—	—	—	—
6	×	×	—	—	—	—	—	×	—	—	—	—	—	—	—
7	×	×	—	—	—	×	—	—	×	—	—	—	—	—	—
8	×	×	—	—	×	×	—	—	—	—	—	—	—	—	—
9	×	×	—	—	—	—	×	—	×	—	—	—	—	—	—
10	×	×	—	—	—	—	—	—	—	×	×	—	—	—	—
11	×	×	—	—	—	—	—	—	—	—	—	×	—	×	×
12	×	×	—	—	—	—	—	—	—	—	—	—	×	—	—
13	×	×	—	—	×	×	—	—	×	—	—	—	—	—	—
14	×	×	—	—	×	×	—	—	×	—	—	—	—	—	×
15	×	×	—	—	×	—	—	—	×	—	—	—	—	—	×
16	×	×	—	—	—	×	—	—	×	—	—	—	—	—	×
17	×	×	—	—	×	—	—	—	×	—	—	—	—	—	—

TABLE 5.1: Fit forms employed to estimate the systematic uncertainty associated with performing the continuum limit extrapolation. A cross indicates the corresponding term is included in the fit form (see sec. 5.2) and the coefficient is reasonably well determined. f_A, d_A, $\tilde{d}_A$ and δc_A are the unknown $\mathcal{O}$ improvement coefficients and the coefficients of $\mathcal{O}(a^2)$ corrections l_O, m_O and n_O are defined in eq. (5.32). The values of these coefficients can be found in app. C. In all the cases, the NLO large-N_c expressions are used for the continuum part of the fit function.

id	$\sqrt{8t_0^\chi}F$	$8t_0^\chi M_0^2$	$L_5 \cdot 10^3$	$L_8 \cdot 10^3$	Λ_1	$\tilde{\Lambda}$
1	$0.1909\,(^{14}_{13})$	$2.52\,(^{6}_{9})$	$1.426\,(^{25}_{32})$	$1.047\,(^{23}_{37})$	$-0.28\,(^{1}_{2})$	$-0.16\,(^{3}_{5})$
2	$0.1913\,(^{20}_{25})$	$2.81\,(^{10}_{18})$	$1.362\,(^{87}_{54})$	$0.797\,(^{133}_{50})$	$-0.23\,(^{5}_{5})$	$-0.01\,(^{5}_{14})$
3	$0.1918\,(^{25}_{23})$	$2.72\,(^{12}_{21})$	$1.345\,(^{73}_{48})$	$0.798\,(^{137}_{52})$	$-0.27\,(^{3}_{11})$	$0.03\,(^{5}_{11})$
4	$0.1922\,(^{17}_{25})$	$2.80\,(^{5}_{17})$	$1.336\,(^{80}_{49})$	$0.826\,(^{122}_{52})$	$-0.22\,(^{1}_{4})$	$-0.01\,(^{5}_{11})$
5	$0.1936\,(^{32}_{39})$	$2.74\,(^{4}_{18})$	$1.461\,(^{107}_{79})$	$0.884\,(^{110}_{55})$	$-0.25\,(^{1}_{5})$	$-0.06\,(^{10}_{15})$
6	$0.1898\,(^{27}_{30})$	$2.77\,(^{6}_{19})$	$1.462\,(^{117}_{115})$	$0.900\,(^{133}_{67})$	$-0.22\,(^{1}_{4})$	$-0.14\,(^{10}_{14})$
7	$0.1890\,(^{23}_{31})$	$2.79\,(^{6}_{17})$	$1.576\,(^{139}_{59})$	$0.955\,(^{127}_{46})$	$-0.22\,(^{1}_{4})$	$-0.20\,(^{4}_{13})$
8	$0.1961\,(^{32}_{30})$	$2.73\,(^{6}_{17})$	$1.436\,(^{114}_{83})$	$0.845\,(^{109}_{64})$	$-0.26\,(^{2}_{4})$	$0.00\,(^{6}_{11})$
9	$0.1884\,(^{14}_{35})$	$2.80\,(^{5}_{17})$	$1.531\,(^{134}_{58})$	$0.947\,(^{136}_{42})$	$-0.21\,(^{1}_{4})$	$-0.19\,(^{5}_{14})$
10	$0.1914\,(^{19}_{26})$	$2.77\,(^{6}_{16})$	$1.358\,(^{83}_{51})$	$0.802\,(^{124}_{57})$	$-0.24\,(^{1}_{4})$	$0.00\,(^{5}_{10})$
11	$0.1918\,(^{31}_{27})$	$2.85\,(^{5}_{20})$	$1.353\,(^{72}_{57})$	$0.812\,(^{127}_{52})$	$-0.21\,(^{3}_{7})$	$-0.02\,(^{9}_{12})$
12	$0.1922\,(^{27}_{24})$	$2.92\,(^{8}_{28})$	$1.337\,(^{75}_{47})$	$0.825\,(^{120}_{51})$	$-0.17\,(^{3}_{9})$	$-0.05\,(^{9}_{10})$
13	$0.1920\,(^{27}_{44})$	$2.74\,(^{9}_{17})$	$1.548\,(^{147}_{68})$	$0.943\,(^{123}_{57})$	$-0.24\,(^{1}_{4})$	$-0.13\,(^{5}_{16})$
14	$0.1919\,(^{31}_{45})$	$2.73\,(^{8}_{17})$	$1.548\,(^{148}_{71})$	$0.942\,(^{123}_{54})$	$-0.24\,(^{2}_{5})$	$-0.13\,(^{5}_{17})$
15	$0.1925\,(^{29}_{47})$	$2.72\,(^{10}_{16})$	$1.523\,(^{111}_{70})$	$0.935\,(^{115}_{53})$	$-0.25\,(^{1}_{4})$	$-0.11\,(^{6}_{17})$
16	$0.1890\,(^{19}_{32})$	$2.77\,(^{8}_{16})$	$1.574\,(^{144}_{51})$	$0.954\,(^{127}_{33})$	$-0.22\,(^{1}_{4})$	$-0.19\,(^{4}_{14})$
17	$0.1927\,(^{27}_{52})$	$2.73\,(^{10}_{17})$	$1.523\,(^{115}_{71})$	$0.936\,(^{133}_{57})$	$-0.24\,(^{1}_{4})$	$-0.11\,(^{6}_{17})$

TABLE 5.2: Results for the LECs obtained when employing the fit forms detailed in tab. 5.1. The dimensionful LECs are given in units of the gradient flow scale in the chiral limit.

id	χ^2/N_{df}	$\sqrt{8t_0^{\mathrm{ph}}}M_\eta$	$\sqrt{8t_0^{\mathrm{ph}}}M_{\eta'}$	$\sqrt{8t_0^{\mathrm{ph}}}F_\eta^8$	$\sqrt{8t_0^{\mathrm{ph}}}F_{\eta'}^8$	$\sqrt{8t_0^{\mathrm{ph}}}F_\eta^0$	$\sqrt{8t_0^{\mathrm{ph}}}F_{\eta'}^0$
1	1.75	$1.162\,(^{8}_{10})$	$1.982\,(^{17}_{13})$	$0.2149\,(^{10}_{14})$	$-0.1015\,(^{25}_{25})$	$0.0337\,(^{22}_{18})$	$0.1884\,(^{7}_{20})$
2	1.63	$1.169\,(^{8}_{12})$	$1.950\,(^{21}_{11})$	$0.2202\,(^{12}_{27})$	$-0.0847\,(^{38}_{86})$	$0.0230\,(^{62}_{33})$	$0.1943\,(^{44}_{53})$
3	1.64	$1.170\,(^{17}_{13})$	$1.952\,(^{53}_{11})$	$0.2200\,(^{17}_{28})$	$-0.0849\,(^{57}_{93})$	$0.0231\,(^{64}_{45})$	$0.1907\,(^{24}_{93})$
4	1.64	$1.169\,(^{8}_{12})$	$1.958\,(^{22}_{11})$	$0.2195\,(^{9}_{25})$	$-0.0864\,(^{34}_{84})$	$0.0260\,(^{44}_{24})$	$0.1949\,(^{7}_{42})$
5	1.52	$1.167\,(^{7}_{11})$	$1.956\,(^{21}_{12})$	$0.2234\,(^{21}_{43})$	$-0.0901\,(^{31}_{80})$	$0.0242\,(^{49}_{16})$	$0.1960\,(^{9}_{42})$
6	1.58	$1.164\,(^{9}_{10})$	$1.959\,(^{21}_{13})$	$0.2196\,(^{8}_{28})$	$-0.0918\,(^{41}_{82})$	$0.0252\,(^{47}_{16})$	$0.1957\,(^{6}_{44})$
7	1.47	$1.168\,(^{7}_{12})$	$1.958\,(^{22}_{10})$	$0.2219\,(^{15}_{31})$	$-0.0939\,(^{23}_{83})$	$0.0224\,(^{44}_{25})$	$0.1974\,(^{12}_{40})$
8	1.56	$1.168\,(^{6}_{12})$	$1.952\,(^{21}_{12})$	$0.2256\,(^{16}_{42})$	$-0.0877\,(^{29}_{76})$	$0.0233\,(^{46}_{26})$	$0.1962\,(^{13}_{37})$
9	1.49	$1.168\,(^{9}_{9})$	$1.960\,(^{23}_{9})$	$0.2201\,(^{7}_{28})$	$-0.0935\,(^{21}_{90})$	$0.0238\,(^{42}_{17})$	$0.1968\,(^{12}_{39})$
10	1.63	$1.169\,(^{6}_{12})$	$1.952\,(^{21}_{12})$	$0.2201\,(^{10}_{27})$	$-0.0851\,(^{42}_{84})$	$0.0232\,(^{53}_{38})$	$0.1930\,(^{12}_{37})$
11	1.65	$1.169\,(^{9}_{11})$	$1.955\,(^{23}_{13})$	$0.2200\,(^{13}_{25})$	$-0.0855\,(^{34}_{78})$	$0.0247\,(^{62}_{36})$	$0.1965\,(^{24}_{58})$
12	1.64	$1.168\,(^{9}_{10})$	$1.957\,(^{23}_{17})$	$0.2194\,(^{16}_{25})$	$-0.0866\,(^{31}_{82})$	$0.0274\,(^{41}_{32})$	$0.1998\,(^{22}_{86})$
13	1.47	$1.168\,(^{9}_{10})$	$1.957\,(^{23}_{10})$	$0.2237\,(^{17}_{36})$	$-0.0931\,(^{24}_{87})$	$0.0232\,(^{45}_{29})$	$0.1969\,(^{12}_{38})$
14	1.48	$1.168\,(^{8}_{9})$	$1.957\,(^{23}_{10})$	$0.2237\,(^{18}_{35})$	$-0.0930\,(^{21}_{88})$	$0.0230\,(^{47}_{26})$	$0.1966\,(^{15}_{42})$
15	1.47	$1.169\,(^{11}_{10})$	$1.957\,(^{23}_{9})$	$0.2235\,(^{17}_{35})$	$-0.0927\,(^{20}_{85})$	$0.0237\,(^{47}_{32})$	$0.1961\,(^{14}_{38})$
16	1.48	$1.168\,(^{11}_{9})$	$1.958\,(^{22}_{9})$	$0.2219\,(^{13}_{28})$	$-0.0938\,(^{23}_{84})$	$0.0222\,(^{41}_{24})$	$0.1968\,(^{17}_{42})$
17	1.46	$1.169\,(^{10}_{10})$	$1.958\,(^{23}_{9})$	$0.2236\,(^{17}_{38})$	$-0.0927\,(^{22}_{86})$	$0.0240\,(^{42}_{29})$	$0.1965\,(^{10}_{39})$

TABLE 5.3: Results for the masses and decay constants of the η and η' at the physical point in units of the gradient flow scale obtained when employing the fit forms detailed in tab. 5.1.

Among the linear improvement terms (see eqs. (5.29) and (5.30)) those involving f_A^l and d_A^l have the largest effect, shifting both the singlet and octet decay constants considerably when they are included. We find the difference between the octet and the singlet quark mass independent improvement coefficients, δc_A^l, is zero within errors and $\tilde{d}_A^l$ is very small and only weakly constrained by the data. We were unable to resolve discretization effects on the masses. All fits require the f_A^l and the d_A^l terms. The fits with the ids 7, 9 and 13–17 in tab. 5.1 have the lowest and very similar χ^2 values, see tab. 5.3. All these fits have in common that an $\mathcal{O}(a^2)$ effect proportional to δM^2 was added to the octet decay constant of the η' meson ($n_{F_{\eta'}^8} \neq 0$ in eq. (5.32)). In what follows we take fit 7 with $\chi^2/N_{\mathrm{df}} \approx 179/122 \approx 1.47$ as our main fit. This was selected from the fits with $1.46 \leq \chi^2/N_{\mathrm{df}} \leq 1.49$ since the resulting parameter values are in the centre of the scatter between the different fit forms, see tabs. 5.4–5.6. We remark again that all correlations between observables determined on the same ensemble are taken into account in the fits. Performing an uncorrelated fit with fit form 7 leads to $\chi^2/N_{\mathrm{df}} \approx 155/122 \approx 1.27$. The systematic uncertainty associated with the continuum extrapolation is assigned to be the 68.3 % interval of the scatter of the central values of the continuum limit fits performed with fit forms 2 to 17. Fit 1 is excluded as important $\mathcal{O}(a)$ terms in the parametrization of the octet and singlet decay constants were omitted in this case.

The χ^2/N_{df} for our best fits are somewhat larger than one: either have we underestimated the errors of our masses and decay constants by about 20 % on average or the functional forms employed do not describe the data sufficiently well. Since the lattice spacing effects seem to be relatively mild, this suggests that NLO large-N_c ChPT does not perfectly describe the data over the range of quark masses available and higher order contributions in the chiral expansion have to be taken into account. The main parameter that determines the convergence of the chiral expansion is the average pseudoscalar meson mass. To investigate the systematics of the chiral extrapolation, we restrict the mass ranges of the data entering the fit, introducing the cut-offs $12t_0\overline{M}^2 < c$, where $c = 1.6, 1.4, 1.2$. These values correspond to $\overline{M} \approx 493\,\mathrm{MeV}$, $462\,\mathrm{MeV}$ and $427\,\mathrm{MeV}$, respectively. Note that $12t_0\overline{M}^2 = 1.11$ corresponds to the physical point and our data cover the range $1.07 \lesssim 12t_0\overline{M}^2 \lesssim 1.68$, see tab. 3.1 and the right panel of fig. 3.1 indicating the cuts graphically. Applying the cuts (successively decreasing c) leads to data points being removed along the trajectory where the strange quark mass is kept constant. For $c = 1.2$ only one ensemble (D201) remains

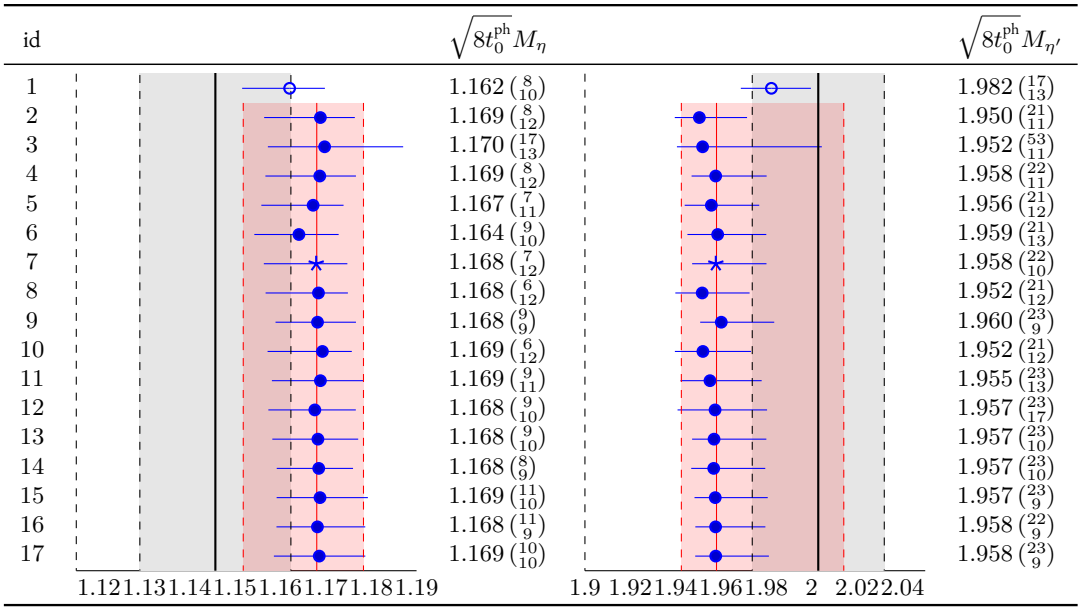


TABLE 5.4: Results for the masses of the η and η' at the physical point in units of the gradient flow scale. The black line marks the experimental result converted using $(8t_0^{\mathrm{ph}})^{-1/2} = 475(6)\,\mathrm{MeV}$ [135] and the grey shaded region marks the uncertainty due to the error on t_0^{ph}. The red line indicates the central values predicted by fit 7 (see tab. 5.1). The red shaded region represents the total uncertainty of our final results where all errors are added in quadrature (see sec. 5.5). The first fit does not sufficiently parametrize the lattice spacing effects and is not included in the determination of the associated systematic error.

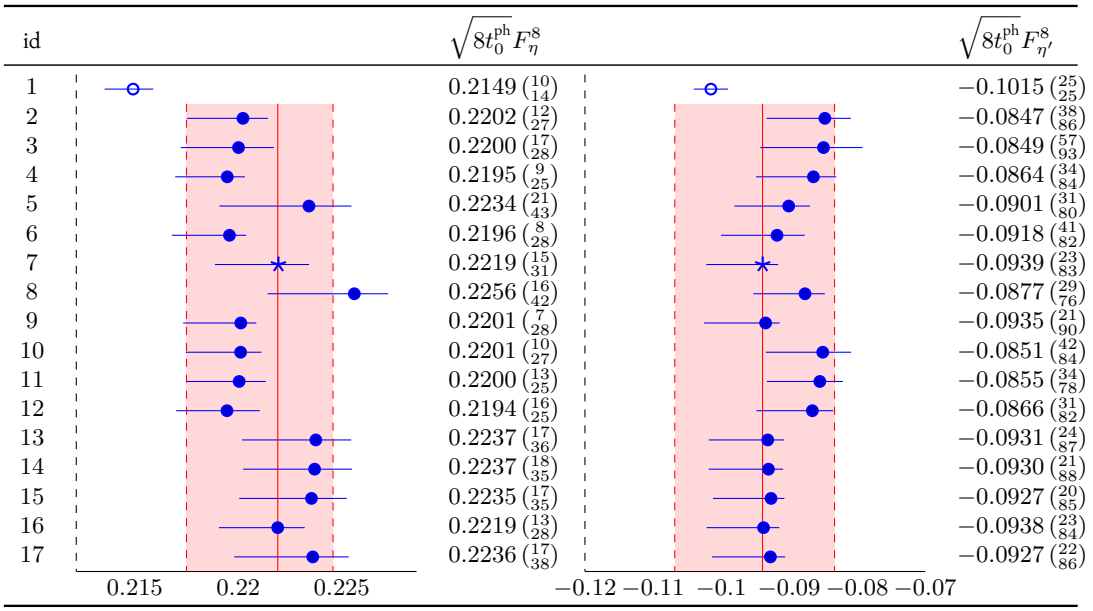


TABLE 5.5: Octet decay constants of the η and η' at the physical point in units of the gradient flow scale, displayed as in tab. 5.4.

id	$\sqrt{8t_0^{\rm ph}}F_\eta^0$	$\sqrt{8t_0^{\rm ph}}F_{\eta'}^0$
1	$0.0337\,(^{22}_{18})$	$0.1884\,(^{7}_{20})$
2	$0.0230\,(^{62}_{33})$	$0.1943\,(^{44}_{53})$
3	$0.0231\,(^{64}_{45})$	$0.1907\,(^{24}_{93})$
4	$0.0260\,(^{44}_{24})$	$0.1949\,(^{7}_{42})$
5	$0.0242\,(^{49}_{16})$	$0.1960\,(^{9}_{42})$
6	$0.0252\,(^{47}_{16})$	$0.1957\,(^{6}_{44})$
7	$0.0224\,(^{44}_{25})$	$0.1974\,(^{12}_{40})$
8	$0.0233\,(^{46}_{26})$	$0.1962\,(^{13}_{37})$
9	$0.0238\,(^{42}_{17})$	$0.1968\,(^{12}_{39})$
10	$0.0232\,(^{53}_{38})$	$0.1930\,(^{12}_{37})$
11	$0.0247\,(^{62}_{36})$	$0.1965\,(^{24}_{58})$
12	$0.0274\,(^{41}_{32})$	$0.1998\,(^{22}_{86})$
13	$0.0232\,(^{45}_{29})$	$0.1969\,(^{12}_{38})$
14	$0.0230\,(^{47}_{26})$	$0.1966\,(^{15}_{42})$
15	$0.0237\,(^{47}_{32})$	$0.1961\,(^{14}_{38})$
16	$0.0222\,(^{41}_{24})$	$0.1968\,(^{17}_{42})$
17	$0.0240\,(^{42}_{29})$	$0.1965\,(^{10}_{39})$

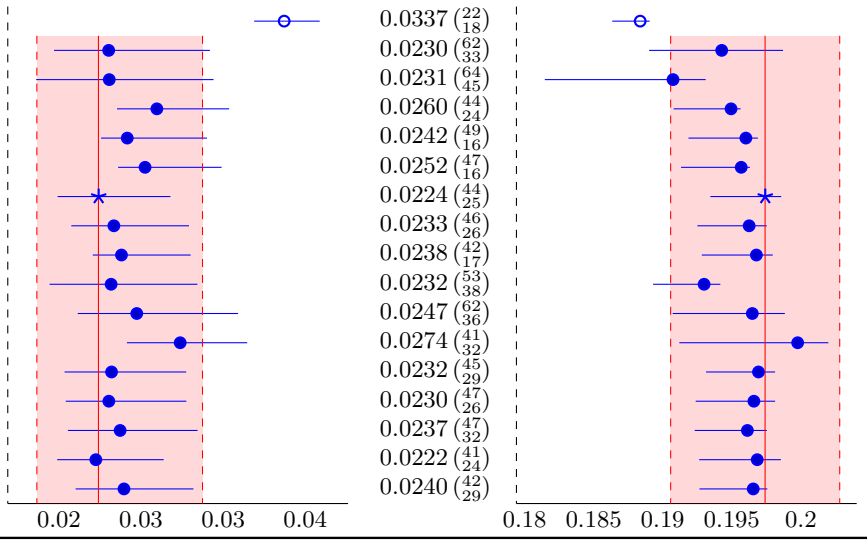


TABLE 5.6: Singlet decay constants of the η and η' at the physical point in units of the gradient flow scale, displayed as in tab. 5.4.

c	$\chi^2/N_{\rm df}$	$\sqrt{8t_0^{\rm ph}}M_\eta$	$\sqrt{8t_0^{\rm ph}}M_{\eta'}$	$\sqrt{8t_0^{\rm ph}}F_\eta^8$	$\sqrt{8t_0^{\rm ph}}F_{\eta'}^8$	$\sqrt{8t_0^{\rm ph}}F_\eta^0$	$\sqrt{8t_0^{\rm ph}}F_{\eta'}^0$
—	1.47	$1.168\,(^{7}_{12})$	$1.958\,(^{22}_{10})$	$0.2219\,(^{15}_{31})$	$-0.0939\,(^{23}_{83})$	$0.0224\,(^{44}_{25})$	$0.1974\,(^{12}_{40})$
1.6	1.49	$1.163\,(^{7}_{11})$	$1.954\,(^{20}_{11})$	$0.2217\,(^{15}_{31})$	$-0.0937\,(^{18}_{82})$	$0.0229\,(^{42}_{21})$	$0.1967\,(^{10}_{38})$
1.4	1.38	$1.173\,(^{8}_{12})$	$1.961\,(^{25}_{14})$	$0.2230\,(^{19}_{30})$	$-0.0881\,(^{36}_{82})$	$0.0206\,(^{41}_{31})$	$0.1947\,(^{11}_{34})$
1.2	1.25	$1.162\,(^{10}_{12})$	$2.006\,(^{23}_{20})$	$0.2225\,(^{24}_{35})$	$-0.1021\,(^{75}_{99})$	$0.0203\,(^{42}_{37})$	$0.1978\,(^{18}_{31})$

TABLE 5.7: Results for masses and decay constants of the η and η' at the physical point in units of the gradient flow scale obtained when employing fit 7 of tab. 5.1 and imposing cut-offs $12t_0\overline{M}^2 < c$, as well as including all the data (first row).

on this trajectory. We perform a fit for each cut using fit form 7. The results are listed in tabs. 5.7 and 5.8. The $\chi^2/N_{\rm df}$ of these fits decrease down to a value of 1.25 as the data are restricted to smaller values of $\overline{M}^2$. This trend suggests that higher order effects should be considered. However, the results are all fairly independent of the cut-off. Only the central value of $M_{\eta'}$ moves upwards and $\tilde{\Lambda}$ downwards by two statistical standard deviations.

In principle, large-N_c ChPT expressions for the masses and decay constants to NNLO are available [181, 182], however, the large number of additional LECs cannot be resolved when fitting our data. Instead, we perform a partial NNLO fit, only including the loop terms which appear at this order. These do not involve any additional LECs, see app. D for details on the parametrization and the resulting LECs. However, fits to this functional form did not improve the description of the

c	$\sqrt{8t_0^\chi}F$	$8t_0^\chi M_0^2$	$L_5\cdot 10^3$	$L_8\cdot 10^3$	Λ_1	$\tilde{\Lambda}$
—	$0.1890\,(^{23}_{31})$	$2.79\,(^{6}_{17})$	$1.576\,(^{139}_{59})$	$0.955\,(^{127}_{46})$	$-0.22\,(^{1}_{4})$	$-0.20\,(^{4}_{13})$
1.6	$0.1894\,(^{20}_{33})$	$2.75\,(^{5}_{16})$	$1.559\,(^{131}_{62})$	$0.930\,(^{126}_{35})$	$-0.23\,(^{1}_{4})$	$-0.19\,(^{4}_{14})$
1.4	$0.1911\,(^{20}_{32})$	$2.78\,(^{6}_{15})$	$1.487\,(^{130}_{74})$	$0.882\,(^{125}_{67})$	$-0.25\,(^{2}_{3})$	$-0.08\,(^{6}_{16})$
1.2	$0.1852\,(^{37}_{45})$	$2.79\,(^{8}_{13})$	$1.777\,(^{132}_{150})$	$1.111\,(^{141}_{131})$	$-0.23\,(^{2}_{3})$	$-0.49\,(^{17}_{24})$

TABLE 5.8: LECs obtained for the fits detailed in tab 5.7. The dimensionful LECs are given in units of the gradient flow scale.

data and our best fit gives a $\chi^2/N_{\rm df} = 2.56$, indicating that a consistent full NNLO parametrization is required.

Utilizing the available data we cannot resolve additional NNLO LECs. The impact on our results from imposing different cut-offs on $\overline{M}^2$ was marginal and hence, we take as our central values the results of fit 7 to all our ensembles, where $\chi^2/N_{\rm df} \approx 179/122 \approx 1.47$. To account for the somewhat inferior quality of this fit, we inflate our statistical errors by the factor $\sqrt{\chi^2/N_{\rm df}} = 1.21$. We also add the NLO truncation error of large-N_c ChPT as a further systematic error (with subscript χ). This corresponds to the range of central values resulting from the fits with different cut-offs.

5.4 Results at the physical point and discussion

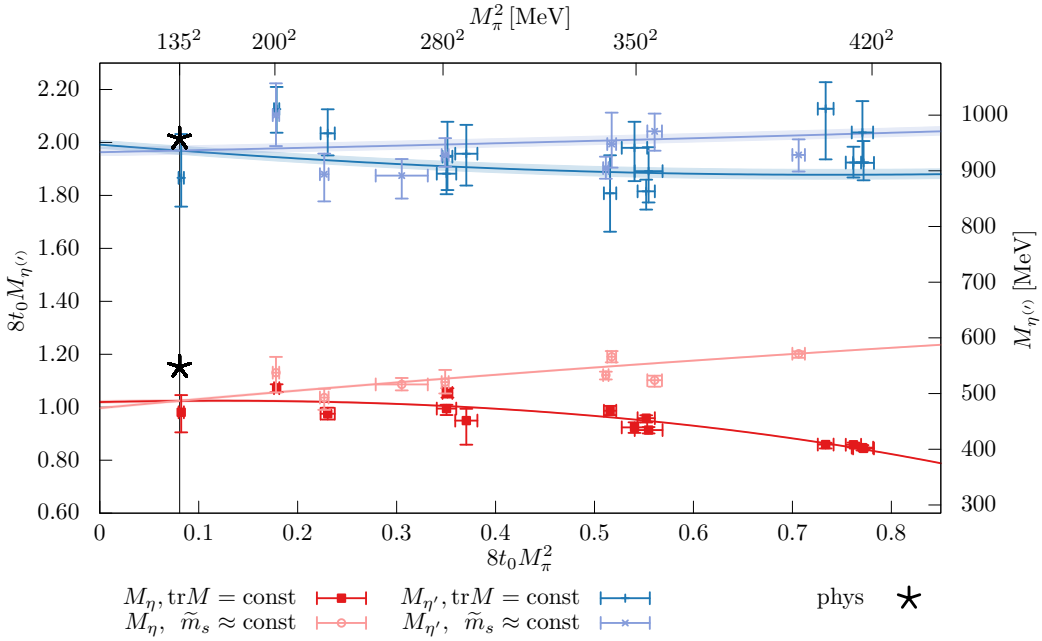


FIGURE 5.1: LO parametrization of the masses of the η and the η' mesons for our two trajectories in the quark mass plane. The data are corrected for lattice spacing effects according to the fit.

LO For completeness, we perform a fit to the η and η' masses employing the LO large-N_c ChPT expressions (see sec. 5.1.1). The decay constants are not included in the analysis as our data clearly contradict the LO ChPT expectation that, e.g., $F_\eta^0 = -F_{\eta'}^8$. The parametrization of the lattice spacing effects was explored in a similar way to the procedure described in the previous subsection. Our best fit gives $\chi^2/N_{\rm df} \approx 91/41 \approx 2.35$ and includes the two quark mass dependent discretization terms n_{M_η} and $n_{M_{\eta'}}$ that are proportional to the difference of the quark masses, cf. eq. (5.32). This fit is displayed in fig. 5.1 and the points are shifted to compensate for the parametrized lattice spacing effects. Since in our NLO fits described below no lattice spacing dependent terms had to be added for the masses, we suspect that in the LO case the $a^2\delta M^2$ terms mostly compensate for a shortcoming of the continuum parametrization.

The masses extracted at the physical point read: $M_\eta = 1.024(8t_0^{\rm ph})^{-1/2} = 487\,\text{MeV}$ and $M_{\eta'} = 1.970(8t_0^{\rm ph})^{-1/2} = 936\,\text{MeV}$, where we do not quote any errors since the fit does not describe the data sufficiently well. The above numbers, in particular the one for the η meson, are significantly lower than the corresponding experimental masses, $M_\eta \approx 548\,\text{MeV}$ and $M_{\eta'} \approx 958\,\text{MeV}$. To this order, the continuum parametrization depends only on one LEC, the (squared) anomalous mass contribution in the chiral limit: we find $M_0^2 = 2.787(8t_0^\chi)^{-1} = (785\,\text{MeV})^2$.

μ_0	$\chi^2/N_{\rm df}$	$8t_0^\chi M_0^2$	Λ_1	$\tilde{\Lambda}$	$\sqrt{8t_0^{\rm ph}}F_\eta^0$	$\sqrt{8t_0^{\rm ph}}F_{\eta'}^0$
$a^{-1}/2$	1.51	$2.68\,(^{6}_{15})$	$-0.27\,(^{1}_{3})$	$-0.15\,(^{3}_{14})$	$0.0204\,(^{43}_{24})$	$0.1922\,(^{13}_{35})$
a^{-1}	1.47	$2.79\,(^{6}_{17})$	$-0.22\,(^{1}_{4})$	$-0.20\,(^{4}_{13})$	$0.0224\,(^{44}_{25})$	$0.1974\,(^{12}_{40})$
$2\,a^{-1}$	1.46	$2.84\,(^{6}_{18})$	$-0.19\,(^{1}_{4})$	$-0.22\,(^{4}_{18})$	$0.0232\,(^{44}_{24})$	$0.2000\,(^{10}_{43})$

TABLE 5.9: Results that depend on Z_A^s, varying the scale at which we match to perturbation theory.

Since the LO fit does not describe our data well, this value of M_0 should also be treated with caution.

NLO Our final results are obtained employing the NLO continuum limit parametrization within simultaneous fits to all data on the two masses and four decay constants. This involves a total of six LECs. Lattice spacing effects are also accounted for as discussed in sec. 5.3. The central values are taken from the results of fit 7, which gave $\chi^2/N_{\rm df} \approx 179/122 \approx 1.47$. This fit is displayed in fig. 5.2. We obtain for the masses at the physical point, in the continuum limit

$$\sqrt{8t_0^{\rm ph}}M_\eta = 1.168\,(^{8}_{14})_{\rm stat}\,(^{1}_{0})_a\,(^{5}_{6})_\chi \quad \text{and} \quad \sqrt{8t_0^{\rm ph}}M_{\eta'} = 1.958\,(^{27}_{13})_{\rm stat}\,(^{0}_{6})_a\,(^{48}_{3})_\chi\,, \tag{5.33}$$

where the first error is statistical (inflated by $\sqrt{1.47}$), and the rest are systematic errors: the second error is taken from the spread of results when varying the parametrization of lattice spacing effects and the third represents the uncertainty due to the (continuum) quark mass dependence, see sec. 5.3 for details. The results are converted to physical units in sec. 5.5. In this section we keep all the results in units of $8t_0^{\rm ph}$.

The final results for the octet and singlet decay constants read

$$\begin{aligned}
\sqrt{8t_0^{\rm ph}}F_\eta^8 &= \quad 0.2219\,(^{18}_{37})_{\rm stat}\,(^{17}_{24})_a\,(^{10}_{2})_\chi\,,\\
\sqrt{8t_0^{\rm ph}}F_{\eta'}^8 &= -0.0939\,(^{28}_{100})_{\rm stat}\,(^{84}_{0})_a\,(^{58}_{82})_\chi\,,\\
\sqrt{8t_0^{\rm ph}}F_\eta^0(\mu=\infty) &= \quad 0.0224\,(^{53}_{30})_{\rm stat}\,(^{28}_{0})_a\,(^{5}_{21})_\chi\,(^{20}_{8})_{\rm renorm}\,,\\
\sqrt{8t_0^{\rm ph}}F_{\eta'}^0(\mu=\infty) &= \quad 0.1974\,(^{14}_{48})_{\rm stat}\,(^{0}_{31})_a\,(^{4}_{27})_\chi\,(^{52}_{26})_{\rm renorm}\,.
\end{aligned} \tag{5.34}$$

The singlet decay constants depend on the QCD scale. As detailed in sec. 2.4.3, prior to the fits we run our results from a scale $\mu_0 = a^{-1}$ up to $\mu = \infty$. To quantify the systematic error from the matching to the $\overline{\rm MS}$ scheme, we repeat the fits, setting $\mu_0 = a^{-1}/2$ and $\mu_0 = 2a^{-1}$, and add the range of results (see tab. 5.9) as an additional systematic error.[1] As one may expect, this error is dominated by the fit where we set $\mu = a^{-1}/2$, see tab. 5.9. The results for the decay constants can also be converted to the strange/light flavour basis (eq. (2.103)) and/or given in terms of two angles and two dimensionful decay constants, see eqs. (2.100) and (2.102). All the results in the different conventions and for the additional QCD scales $\mu \in \{1\,\text{GeV}, 2\,\text{GeV}, 10\,\text{GeV}\}$ are collected in tab. E.1 in app. E. We discuss the results and their scale dependence in detail in sec. 5.5.2.

[1] We remark that slightly different results are also obtained for the scale independent quantities. However, the differences are well below any other systematic error, with the exception of those for $\tilde{\Lambda}$.

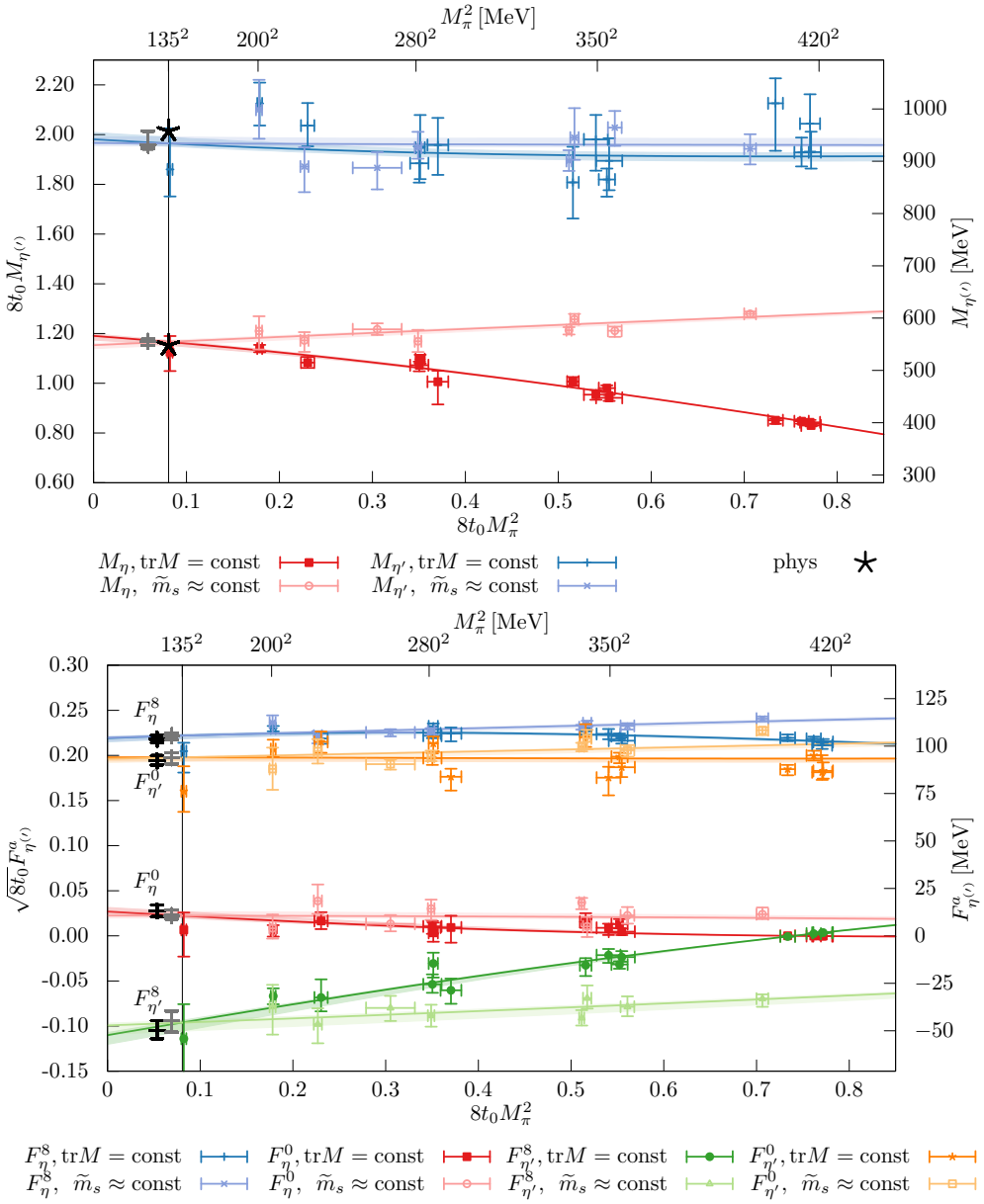


FIGURE 5.2: Simultaneous fit to the masses (top) and four decay constants (bottom) of the η and η'. The fit form incorporates the NLO large-N_c ChPT expressions and the discretization terms corresponding to fit 7 in tab. 5.1. The points have been shifted to compensate for lattice spacing effects and lie along two trajectories leading to the physical point. The continuum fit functions are indicated by the lines and shaded regions (statistical errors only), where the darker and lighter colours correspond to the trajectories where the flavour average quark mass is held constant, and the strange quark mass is kept constant, respectively. The grey and black error bars (shifted to the left of the physical pion mass for better visibility) are our final results, without priors (grey) and including the experimental η and η' masses (black stars) as priors (black, see sec. 5.4.1). All errors are added in quadrature.

5.4.1 Results for the large-N_c low energy constants

Our results from the fits detailed above for the large-N_c LECs read

$$\begin{aligned}
L_5 &= 1.58\,\binom{17}{7}_{\text{stat}}\binom{0}{22}_{a}\binom{20}{9}_{\chi}\cdot 10^{-3},\\
L_8 &= 0.96\,\binom{15}{6}_{\text{stat}}\binom{0}{14}_{a}\binom{16}{7}_{\chi}\cdot 10^{-3},\\
M_0(\mu=\infty) &= 1.67\,\binom{2}{6}_{\text{stat}}\binom{1}{2}_{a}\binom{0}{1}_{\chi}\binom{3}{2}_{\text{renorm}}\left(8t_0^{\chi}\right)^{-1/2},\\
F &= 0.1890\,\binom{27}{37}_{\text{stat}}\binom{36}{0}_{a}\binom{21}{38}_{\chi}\left(8t_0^{\chi}\right)^{-1/2},\\
\Lambda_1(\mu=\infty) &= -0.22\,\binom{1}{5}_{\text{stat}}\binom{0}{3}_{a}\binom{0}{3}_{\chi}\binom{6}{3}_{\text{renorm}},\\
\tilde{\Lambda} &= -0.20\,\binom{5}{16}_{\text{stat}}\binom{19}{0}_{a}\binom{12}{29}_{\chi}\binom{3}{5}_{\text{renorm}}.
\end{aligned} \tag{5.35}$$

The combination $\tilde{\Lambda}=\Lambda_1-2\Lambda_2$ is scale invariant to this order in ChPT [33, 181], however, since its central value varies when changing μ_0, see tab. 5.9, we also assign a renormalization error in this case. The above results give

$$\Lambda_2(\mu=\infty) = -0.1\,\binom{8}{4}_{\text{stat}}\binom{0}{10}_{a}\binom{14}{8}_{\chi}\binom{5}{3}_{\text{renorm}}. \tag{5.36}$$

The fits on which these results are based give η and η' masses that are compatible, within errors, with experiment, see above and sec. 5.5.1. Nevertheless, incorporating prior knowledge of the experimental masses helps to constrain the fit and reduces the errors on the LECs. To this end, we modify our χ^2 function to penalize fits that give masses, that are incompatible with experiment:

$$\begin{aligned}
\chi^2_{\text{priors}} = \chi^2 &+ \left[\frac{\sqrt{8t_0^{\text{ph}}}M_{\eta}^{\text{ph}} - f_{M_\eta}(a=0,12t_0\overline{M}^2,8t_0\delta M^2)}{\sigma\left(\sqrt{8t_0^{\text{ph}}}M_{\eta}^{\text{ph}}\right)}\right]^2\\
&+ \left[\frac{\sqrt{8t_0^{\text{ph}}}M_{\eta'}^{\text{ph}} - f_{M_{\eta'}}(a=0,12t_0\overline{M}^2,8t_0\delta M^2)}{\sigma\left(\sqrt{8t_0^{\text{ph}}}M_{\eta'}^{\text{ph}}\right)}\right]^2,
\end{aligned} \tag{5.37}$$

where we use the physical values from the Particle Data Group (PDG) [19] for M_η and $M_{\eta'}$, see eq. (5.41). These are converted to dimensionless numbers, using $(8t_0^{\text{ph}})^{-1/2}=475(6)$ MeV [135]. Note that the errors are dominated by the scale and are thus highly correlated. This is taken into account by sampling Gaussian distributed values for $(8t_0^{\text{ph}})^{1/2}$, rather than independently sampling the two dimensionless combinations $(8t_0^{\text{ph}})^{1/2}M_{\eta}^{\text{ph}}$ and $(8t_0^{\text{ph}})^{1/2}M_{\eta'}^{\text{ph}}$. If a more precise determination of t_0^{ph} became available, the priors could be further constrained and the uncertainties reduced.

Repeating the whole fitting analysis, now including the priors, we obtain results that are very similar to those of eqs. (5.35) and (5.36):

$$\begin{aligned}
L_5 &= 1.66\,\binom{12}{9}_{\text{stat}}\binom{0}{26}_{a}\binom{13}{8}_{\chi}\cdot 10^{-3},\\
L_8 &= 1.08\,\binom{11}{6}_{\text{stat}}\binom{0}{12}_{a}\binom{3}{10}_{\chi}\cdot 10^{-3},\\
M_0(\mu=\infty) &= 1.62\,\binom{2}{4}_{\text{stat}}\binom{3}{1}_{a}\binom{3}{0}_{\chi}\binom{2}{1}_{\text{renorm}}\left(8t_0^{\chi}\right)^{-1/2},\\
F &= 0.1866\,\binom{26}{29}_{\text{stat}}\binom{54}{0}_{a}\binom{19}{16}_{\chi}\left(8t_0^{\chi}\right)^{-1/2},\\
\Lambda_1(\mu=\infty) &= -0.25\,\binom{1}{4}_{\text{stat}}\binom{3}{1}_{a}\binom{1}{1}_{\chi}\binom{5}{2}_{\text{renorm}},\\
\tilde{\Lambda} &= -0.46\,\binom{8}{10}_{\text{stat}}\binom{21}{0}_{a}\binom{9}{10}_{\chi}\binom{1}{2}_{\text{renorm}},\\
\Lambda_2(\mu=\infty) &= 0.11\,\binom{5}{5}_{\text{stat}}\binom{0}{9}_{a}\binom{6}{5}_{\chi}\binom{3}{2}_{\text{renorm}}.
\end{aligned} \tag{5.38}$$

In sec. 5.5.3 we will convert the above results into physical units and discuss them.

In general, the large-N_c LECs will differ from their SU(3) ChPT equivalents, see also the discussion in [181] and in sec. 5.5.3. In particular, the above LECs do not depend on the ChPT scale since chiral logarithms only appear starting at NNLO in large-N_c ChPT. As mentioned above, we checked whether such contributions improved the description of the data by adding the NNLO loop terms to the NLO parametrization. However, this decreased the quality of the fits, with the best fit giving $\chi^2/N_{\text{df}} \approx 312/122 \approx 2.56$. The functional form and the resulting LECs are detailed in app. D.

5.5 Discussion of results and comparison with literature

In this section we summarize our results on the masses, decay constants and the low energy constants parametrizing them in the continuum. These are compared to other determinations from lattice or phenomenological studies. In addition, we study the implications of our findings on the photoproduction transition form factors of the η and η' mesons.

The results are converted into physical units using $(8t_0^{\text{ph}})^{-1/2} = 475(6)$ MeV [135]. The dimensionful low energy constants are defined in the chiral limit and so for these we convert this value of t_0^{ph} to t_0^{χ} using eq. (3.43) and obtain $(8t_0^{\chi})^{-1/2} = 470(7)$ MeV. For some of our results the uncertainty of this scale significantly contributes to the total error. Since improved determinations may become available in the future, we quote this uncertainty separately to the statistical and other systematic errors.

5.5.1 The η and η' meson masses

Our final results for the masses of the η and η' mesons are (see eq. (5.33))

$$M_\eta = 554.7\,\binom{4.0}{6.6}_{\text{stat}}\binom{2.4}{2.7}_{\text{syst}}(7.0)_{t_0},\ \text{MeV} \quad \text{and} \tag{5.39}$$

$$M_{\eta'} = 929.9\,\binom{12.9}{6.0}_{\text{stat}}\binom{22.9}{3.3}_{\text{syst}}(11.7)_{t_0}\ \text{MeV}, \tag{5.40}$$

where we added the systematic errors associated with the continuum and physical quark mass point extrapolations in quadrature. We find reasonably good agreement when comparing these results of $N_f = 2+1$ QCD with the known experimental masses,

$$\text{PDG [19]}: \qquad M_\eta^{\text{ph}} = 547.862(17)\,\text{MeV} \quad \text{and} \quad M_{\eta'}^{\text{ph}} = 957.78(6)\,\text{MeV}. \tag{5.41}$$

The masses are 0.7 standard errors above and one standard error below the experimental values for the η and η', respectively. For M_η, the combined relative error is 1.7 % with the statistical and scale setting uncertainties forming the biggest contributions. Our value for $M_{\eta'}$ has a total uncertainty

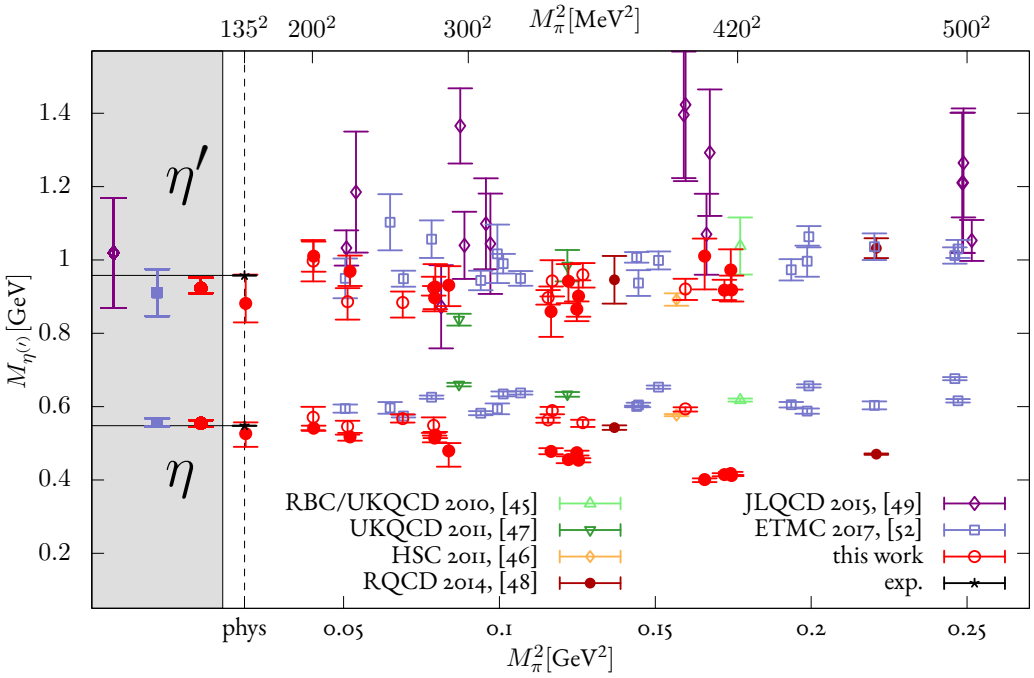


FIGURE 5.3: Recent $N_f = 2 + 1(+1)$ lattice results for the masses of the η and η' mesons. Most points have been simulated at approximately physical strange quark masses (open symbols), whereas in this work we also include an additional trajectory along which the average of the quark masses is kept constant (filled symbols). The three sets of points in the shaded regions left of the physical point (dashed line) correspond to the continuum and chirally extrapolated results of JLQCD [49] (who do not give an estimate of M_η), ETMC [52] and this work.

of 2.3 %, where in this case the statistical error and the uncertainty from the quark mass extrapolation dominate. In both cases, lattice spacing effects are less significant. This reflects the fact that we are not able to resolve any such effects in the masses, see sec. 5.3. For $M_{\eta'}$ this is not so surprising considering the relatively large statistical error obtained on the individual ensembles. We remark that the precision of the final results was achieved by utilizing NLO large-N_c ChPT to simultaneously fit the two masses and four decay constants (summarized in the next subsection) determined on twenty-one ensembles lying along two trajectories in the quark mass plane and comprising four lattice spacings.

Our results at unphysical quark masses as well as in the physical limit are displayed in fig. 5.3, together with $N_f = 2 + 1(+1)$ results of other groups that we are aware of. The η mass is sensitive to the masses of the light and strange quarks and the data points clearly fall along two lines which converge at the physical point, reflecting the two sets of ensembles employed: for one set the physical strange quark mass is kept approximately constant while for the other the flavour average of the strange and light quark masses is held fixed. The singlet contribution to the mass of the η' is significant and no clear quark mass dependence is observed.

Overall, the results for the η' are consistent across different collaborations and actions (also at larger quark masses), whereas for the η some scatter is visible. The latter may be due to mistuning of the strange quark mass and/or lattice spacing effects. In particular, a previous exploratory study of our group [48] is affected by mistuning. The ETMC [52] and JLQCD [49] collaborations employ pion masses reaching down to approximately 220 MeV. In this work, we obtain results close to the physical point for the first time. While the errors are relatively large for our $M_\pi \approx 126$ MeV ensemble, the results are in good agreement with the quark mass extrapolation.

To our knowledge the only other studies which attempt a physical limit extrapolation are those of ETMC [50–52] and JLQCD [49]. The latter $N_f = 2 + 1$ work utilizes gluonic correlation

functions to determine the η' mass. A simple linear extrapolation is performed which is justified in view of the large statistical errors. ETMC [52] employ the twisted mass fermion formulation and simulate $N_f = 2 + 1 + 1$ QCD. The physical point is approached keeping the strange quark mass approximately equal to its physical value, although some mistuning is visible in the results for M_η. This is compensated for by including terms proportional to m_s in the quark mass extrapolation, in addition to terms proportional to m_ℓ and a^2. This leading order ansatz yields an effective parametrization of the data, however, the η and η' masses are assumed to be independent of each other and are fitted separately, ignoring potential correlations in the data. The final errors for M_η and $M_{\eta'}$ at the physical point are larger than ours, in particular for the latter, although the uncertainties on the individual ensembles are similar and at the percent and few-percent level, for the η and η', respectively. We achieve smaller final errors by simultaneously fitting the quark mass and lattice spacing dependence of six observables (two masses and four decay constants), which have been determined on ensembles following two trajectories to the physical point. This, together with including ensembles with small quark masses, enables the quark mass extrapolation to be tightly constrained. The results from ETMC at the physical point are in agreement with our estimates and the experimental values within the quoted errors.

5.5.2 Decay constants

We carry out two sets of fits to extract the four decay constants, one where we simultaneously fit to our lattice results for the masses and decay constants from which the values of the masses at the physical point (presented in the previous subsection) are taken and another set where we constrain the masses to reproduce the physical values by adding prior terms to the χ^2 function. The latter fits enable the LECs to be better constrained, see sec. 5.4.1. The two sets of results, detailed in tabs. E.1 and E.2 in app E, are consistent within errors. In the following, we will only discuss the second set of results. Since the singlet axialvector current has an anomalous dimension in the $\overline{\text{MS}}$ scheme, some of the results depend on the QCD scale, see sec. 2.4.3. Although the fits were performed setting $\mu = \infty$ in $N_f = 3$ QCD, in this section we will mostly quote results at $\mu = 2\,\text{GeV}$. This simplifies a comparison to literature values, as discussed below. The results obtained for a range of scales are also listed in tabs. E.1 and E.2.

Summary of the results. The decay constants, converted to the angle representation of the octet/singlet basis, read at $\mu = 2\,\text{GeV}$ in $N_f = 3$ QCD

$$F^8 = \quad 115.0\,(^{1.1}_{1.2})_{\text{stat}}\,(^{1.6}_{2.4})_{\text{syst}}\,(1.5)_{t_0}\,\text{MeV}, \tag{5.42}$$

$$\theta_8 = -25.8\,(^{1.2}_{2.1})_{\text{stat}}\,(^{2.2}_{0.3})^{\circ}_{\text{syst}}\,, \tag{5.43}$$

$$F^0(\mu = 2\,\text{GeV}) = \quad 100.1\,(^{7}_{1.9})_{\text{stat}}\,(^{2.0}_{2.7})_{\text{syst}}\,(1.3)_{t_0}\,\text{MeV}, \tag{5.44}$$

$$\theta_0 = -8.1\,(^{1.0}_{1.1})_{\text{stat}}\,(^{1.5}_{1.5})^{\circ}_{\text{syst}}\,, \tag{5.45}$$

where we added the systematic errors arising from the continuum and chiral extrapolation in quadrature. This representation has the advantage that only F^0 depends on the scale, however, often the flavour basis in the angle representation is employed in the literature. We find

$$F^\ell(\mu = 2\,\text{GeV}) = 88.28\,(^{1.20}_{2.02})_{\text{stat}}\,(^{3.00}_{1.74})_{\text{syst}}\,(1.12)_{t_0}\,\text{MeV}, \tag{5.46}$$

$$\phi_\ell(\mu = 2\,\text{GeV}) = 36.2\,(^{1.1}_{2.0})_{\text{stat}}\,(^{1.3}_{0.4})^{\circ}_{\text{syst}}\,, \tag{5.47}$$

$$F^s(\mu = 2\,\text{GeV}) = 124.3\,(^{1.7}_{1.6})_{\text{stat}}\,(^{2.7}_{4.3})_{\text{syst}}\,(1.6)_{t_0}\,\text{MeV}, \tag{5.48}$$

$$\phi_s(\mu = 2\,\text{GeV}) = 37.9\,(^{1.0}_{1.3})_{\text{stat}}\,(^{1.4}_{0.8})^{\circ}_{\text{syst}}\,, \tag{5.49}$$

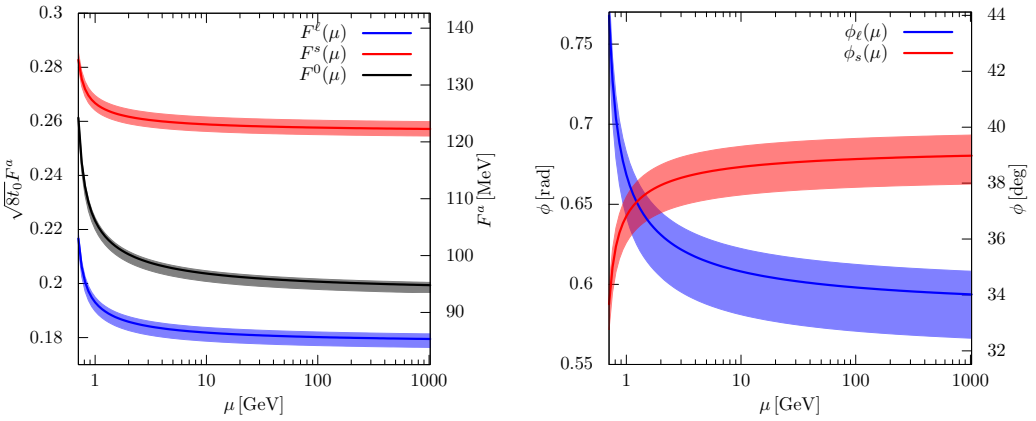


FIGURE 5.4: Scale dependence of the decay constants and their mixing angles in the flavour basis. The approximation $\phi_\ell \approx \phi_s \approx \phi$ is only valid for μ inbetween about 1 GeV and 2 GeV. In the same region the decay constants vary considerably with the scale. The (asymmetric) errors indicated by the coloured bands are statistical only.

where all quantities depend non-trivially on the scale. The popularity of the flavour representation is due to the similarity of the two angles which suggests that the four (independent) decay constants can be described by only three parameters, setting $\phi_\ell = \phi_s = \phi$. This approximation is made in the Feldmann-Kroll-Stech (FKS) scheme [115, 183, 184] and to NLO in large-N_c ChPT it is equivalent to neglecting Okubo-Zweig-Iizuka (OZI) rule violating terms, specifically those involving Λ_1. At NLO the latter parameter is related to the angles and decay constants via [32, 115]

$$\frac{\sqrt{2}}{3}F_\pi^2\Lambda_1 = F^\ell F^s \sin(\phi_\ell - \phi_s). \tag{5.50}$$

Thus, if Λ_1 is set to zero, then within this approximation $\phi_\ell = \phi_s$.

Dependence on the QCD scale. In effect, the assumption $\Lambda_1 = 0$ renders the singlet decay constant independent of the scale since [32, 185]

$$\mu\frac{\mathrm{d}}{\mathrm{d}\mu}\frac{F_0(\mu)}{\sqrt{1+\Lambda_1(\mu)}} = 0. \tag{5.51}$$

The results in eqs. (5.46)–(5.49) show that at $\mu = 2\,\mathrm{GeV}$ the angles almost agree within errors. However, our estimate for $\Lambda_1 = -0.25(5)$ (see sec. 5.4.1) determined at $\mu = \infty$, suggests that this approximation cannot hold at high scales. We display the scale dependent decay constants and angles as a function of μ in fig. 5.4. The two angles are significantly different at large scales where the combination $2(\phi_s - \phi_\ell)/(\phi_s + \phi_\ell)$ approaches 16 %. However, this difference decreases towards lower μ and in the range $0.9\,\mathrm{GeV} \lesssim \mu \lesssim 2\,\mathrm{GeV}$ then $\phi_\ell \approx \phi_s$. This is due to $\Lambda_1(\mu)$ crossing zero around 1 GeV as shown in fig. 5.5, where we display both OZI violating LECs [186],

$$\Lambda_1(\mu) = \left(\frac{Z_A^s(\mu)}{Z_A^{s\prime}}\right)^2 (1+\Lambda_1(\mu=\infty))-1, \quad \Lambda_2(\mu) = \frac{Z_A^s(\mu)}{Z_A^{s\prime}}(1+\Lambda_2(\mu=\infty))-1. \tag{5.52}$$

The LEC Λ_2, which mostly impacts on the masses, becomes small at high scales but should not be neglected at $\mu < 4\,\mathrm{GeV}$. This provides an explanation for the observation of some ChPT studies that Λ_2 plays a more important role than Λ_1 in terms of reproducing the physical η and η' masses [182, 187].

The scale dependence of some observables complicates direct comparisons to phenomenology in many studies employing ChPT, where the relevant QCD scale depends on the processes that are

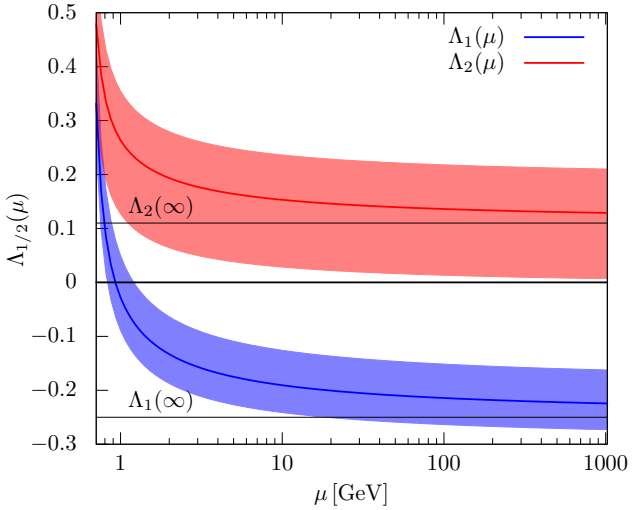


FIGURE 5.5: Scale dependence of the large-N_c ChPT LECs Λ_1 and Λ_2.

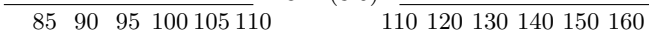
considered to fix the LECs and on the order of ChPT. Typically, the LECs are determined using experimental input from, e.g., the η and η' masses and the widths of radiative decays. This implies a low QCD renormalization scale. Λ_1 varies rapidly in this region which means that if this LEC is determined using physical processes dominated by different physical scales, the predictions for, e.g., F^0 will be affected. Clearly, results obtained using the FKS scheme should be compared at the scale where Λ_1 vanishes. One possibility to mitigate this problem is to compare results in the octet/singlet basis (where only F^0 depends on μ) and form the scale independent combination $F^0/\sqrt{1+\Lambda_1}$ [32]. Our result for the latter reads

$$F^0/\sqrt{1+\Lambda_1} = 107.3\,(^{1.5}_{7})_{\text{stat}}\,(^{1.3}_{1.3})_{\text{syst}}\,(1.4)_{t_0}\,\text{MeV}. \tag{5.53}$$

Comparison with phenomenological results. A comparison with a variety of results for the decay constants in the light/strange flavour and octet/singlet bases is shown in tabs. 5.10 and 5.11, respectively. Most of the results rely on large-N_c ChPT using experimental input to fix the LECs. One of the first such computations was undertaken at NLO by Leutwyler [32], using predominantly pseudoscalar meson masses and non-singlet decay constants to fix the LECs. Only scale independent combinations are quoted and a result for the singlet decay constant is not given. Feldmann [184] then employed the FKS scheme discussed above to give values also for the scale dependent decay constants and the single flavour mixing angle (in this scheme). This approximate scheme is also used on the lattice by ETMC [51, 52] to relate the pseudoscalar matrix elements to the (axial) decay constants, which will be discussed further below. In the first NNLO large-N_c ChPT calculation, Guo et al. [182] take lattice input for M_η, $M_{\eta'}$ and the non-singlet pseudoscalar masses and decay constants at unphysical quark masses from the literature. This allows them to constrain the LECs to NLO, but further assumptions are needed for the many NNLO coefficients. Subsequently, Gu et al. [194] extended the analysis by also utilizing the decay constant results from ETMC [52]. However, additional constraints on the parameters still seem to be necessary in order to obtain stable NNLO results. Bickert et al. [181] also perform an NNLO analysis, in this case combining LECs obtained from the literature and derived from experimental input for the masses and non-singlet decay constants. Again only QCD scale independent combinations are given.

There exist a number of other studies, some of which are based on models, for instance, Benayoun [188, 195] employs vector meson dominance, while others involve more phenomenologically driven extractions, for example, Escribano et al. [185, 189, 190, 192], use experimental data on, e.g., the transition form factors $\gamma\gamma^* \to \eta$ and $\gamma\gamma^* \to \eta'$. However, a connection to NLO large-N_c

ref	F^ℓ/MeV	F^s/MeV
Benayoun et al. [188]	—	—
Escribano and Frere [189]	100.4(2.8)	152.8(5.5)
Escribano et al. [190]	97.6(1.0)	144(20)
Chen et al. [191]	105.9(3.7)	143.6(5.5)
Escribano et al. [192]	98.7(9)	128(13)
Escribano et al. [185]	94.9(3.7)	125.3(3.7)
Leutwyler [32]	—	—
Feldmann [184]	98.5(1.8)	123.4(5.5)
Guo et al. [182] NLO-A	94.1(2.5)	122.3(5.2)
Guo et al. [182] NNLO-B	92.7(1.4)	139(15)
Bickert et al. [181] NLO-I	—	—
[181] NNLO w/o Ci ($\mu_{\text{EFT}} = 1$GeV)	—	—
[181] NNLO w/ Ci ($\mu_{\text{EFT}} = 1$GeV)	—	—
Ding et al. [193]	101	138
ETMC [52]	88.4(5.4)	125.6(2.6)
Gu et al. [194] NNLO-A9p(F_π)	89.8(4.7)	126.3(2.9)
this work ($\mu = 1$GeV)	91.7(3.3)	126.7(4.3)
this work ($\mu = 2$ GeV)	88.3(3.2)	124.3(4.2)
this work ($\mu = \infty$)	84.4(3.0)	121.5(4.1)



ref	ϕ_ℓ	ϕ_s
Benayoun et al. [188]	—	—
Escribano and Frere [189]	39.9(1.3)°	41.4(1.4)°
Escribano et al. [190]	40.3(1.8)°	40.3(1.8)°
Chen et al. [191]	34.5(1.8)°	36.0(1.4)°
Escribano et al. [192]	39.3(1.2)°	39.2(1.2)°
Escribano et al. [185]	39.6(2.3)°	40.8(1.8)°
Leutwyler [32]	—	—
Feldmann [184]	39.3(1.0)°	39.3(1.0)°
Guo et al. [182] NLO-A	40.4(4.8)°	39.9(2.8)°
Guo et al. [182] NNLO-B	35.8(1.2)°	37.1(1.4)°
Bickert et al. [181] NLO-I	—	—
[181] NNLO w/o Ci ($\mu_{\text{EFT}} = 1$GeV)	—	—
[181] NNLO w/ Ci ($\mu_{\text{EFT}} = 1$GeV)	—	—
Ding et al. [193]	42.8°	42.8°
ETMC [52]	38.8(3.3)°	38.8(3.3)°
Gu et al. [194] NNLO-A9p(F_π)	39.6(2.6)°	36.7(2.3)°
this work ($\mu = 1$GeV)	38.3(1.8)°	36.8(1.6)°
this work ($\mu = 2$ GeV)	36.2(1.9)°	37.9(1.6)°
this work ($\mu = \infty$)	33.3(2.0)°	39.3(1.7)°

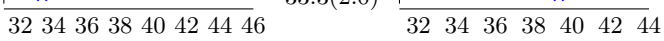
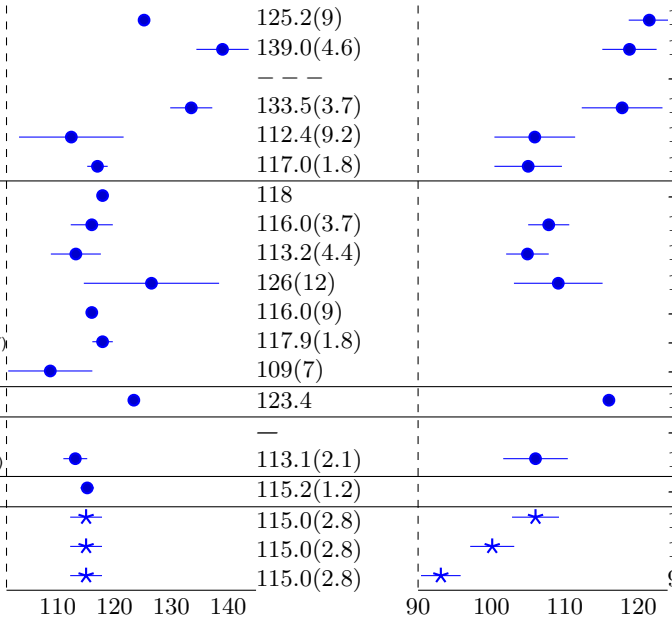


TABLE 5.10: Comparison of recent phenomenological and lattice results for the (scale dependent) decay constants in the angle representation for the light/strange flavour basis. The results from this work are presented at three different scales.

ref	F^8/MeV	F^0/MeV
Benayoun et al. [188]	125.2(9)	121.5(2.8)
Escribano and Frere [189]	139.0(4.6)	118.8(3.7)
Escribano et al. [190]	− − −	—
Chen et al. [191]	133.5(3.7)	117.8(5.5)
Escribano et al. [192]	112.4(9.2)	105.9(5.5)
Escribano et al. [185]	117.0(1.8)	105.0(4.6)
Leutwyler [32]	118	—
Feldmann [184]	116.0(3.7)	107.8(2.8)
Guo et al. [182] NLO-A	113.2(4.4)	104.9(2.9)
Guo et al. [182] NNLO-B	126(12)	109.1(6.0)
Bickert et al. [181] NLO-I	116.0(9)	—
[181] NNLO w/o Ci ($\mu_{\text{EFT}} = 1$GeV)	117.9(1.8)	—
[181] NNLO w/ Ci ($\mu_{\text{EFT}} = 1$GeV)	109(7)	—
Ding et al. [193]	123.4	116.0
ETMC [52]	—	—
Gu et al. [194] NNLO-A9p(F_π)	113.1(2.1)	106.0(4.4)
eq. (5.54)	115.2(1.2)	—
this work ($\mu = 1$GeV)	115.0(2.8)	106.0(3.2)
this work ($\mu = 2$ GeV)	115.0(2.8)	100.1(3.0)
this work ($\mu = \infty$)	115.0(2.8)	93.1(2.7)



ref	θ_8	θ_0
Benayoun et al. [188]	$-20.4(1.0)^\circ$	$-0.1(1.0)^\circ$
Escribano and Frere [189]	$-23.8(1.4)^\circ$	$-2.4(1.9)^\circ$
Escribano et al. [190]	—	—
Chen et al. [191]	$-26.7(1.8)^\circ$	$-11.0(1.0)^\circ$
Escribano et al. [192]	$-21.3(3.5)^\circ$	$-11.3(3.9)^\circ$
Escribano et al. [185]	$-21.2(1.9)^\circ$	$-6.9(2.4)^\circ$
Leutwyler [32]	-20°	-4°
Feldmann [184]	$-21.2(1.9)^\circ$	$-9.2(1.7)^\circ$
Guo et al. [182] NLO-A	$-21.5(4.5)^\circ$	$-7.2(2.5)^\circ$
Guo et al. [182] NNLO-B	$-27.9(1.7)^\circ$	$-6.8(3.8)^\circ$
Bickert et al. [181] NLO-I	$-21.7(7)^\circ$	$-0.5(7)^\circ$
[181] NNLO w/o Ci ($\mu_{\text{EFT}} = 1$GeV)	$-12.6(6.1)^\circ$	$-6.3(6.5)^\circ$
[181] NNLO w/ Ci ($\mu_{\text{EFT}} = 1$GeV)	$-34(22)^\circ$	$-33(24)^\circ$
Ding et al. [193]	-21°	-2.8°
ETMC [52]	—	—
Gu et al. [194] NNLO-A9p(F_π)	$-26.1(2.5)^\circ$	$-7.0(2.1)^\circ$
eq. (5.54)	—	—
this work ($\mu = 1$GeV)	$-25.8(2.3)^\circ$	$-8.1(1.8)^\circ$
this work ($\mu = 2$ GeV)	$-25.8(2.3)^\circ$	$-8.1(1.8)^\circ$
this work ($\mu = \infty$)	$-25.8(2.3)^\circ$	$-8.1(1.8)^\circ$

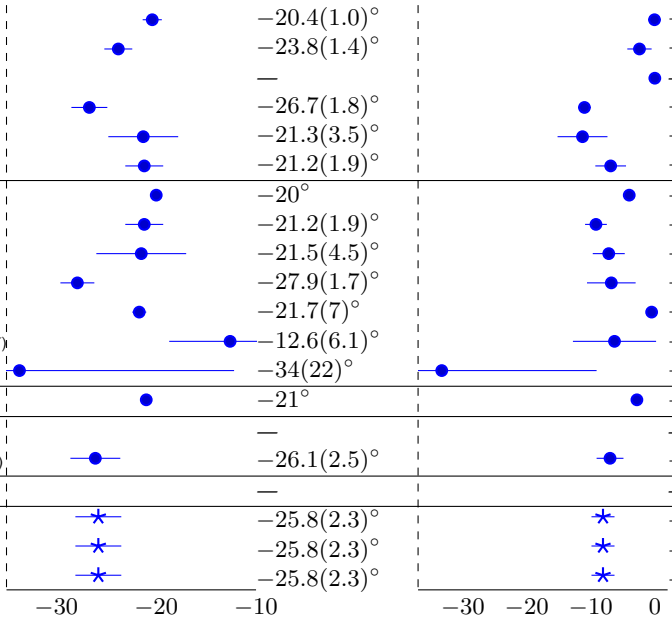


TABLE 5.11: Comparison of recent phenomenological and lattice results for the decay constants in the angle representation for the octet/singlet basis, where only F^0 depends on the scale. We use $F_{\pi^+} = 92.1$ MeV [19] to convert decay constants given as a multiple of F_π and eq. (5.54) refers to the NLO result from literature pion and kaon decay constants. The results from this work are presented at three different scales.

ref	Λ_1	$F^0/\sqrt{1+\Lambda_1}$/MeV
Benayoun et al. [188]	0.20(4)	*110.9(4.8)*
Escribano and Frere [189]	0.34(10)	*102.6(8.2)*
Escribano et al. [190]	0	—
Chen et al. [191]	—	—
Escribano et al. [192]	0	—
Escribano et al. [185]	0.01(13)	*105(11)*
Leutwyler [32]	—	101
Feldmann [184]	0.0(3)	107.8(2.8)
Guo et al. [182] NLO-A	0.02(8)	*102.8(7.0)*
Guo et al. [182] NNLO-B	−0.04(14)	*111(14)*
Bickert et al. [181] NLO-I	—	104.1(0)
[181] NNLO w/o Ci ($\mu_{\text{EFT}} = 1$GeV)	0	79.2(9)
[181] NNLO w/ Ci ($\mu_{\text{EFT}} = 1$GeV)	0	76.4(9)
Ding et al. [193]	—	—
ETMC [52]	0	—
Gu et al. [194] NNLO-A9p(F_π)	0.24(21)	*95(20)*
eq. (5.54)	—	104.3(1.1)
this work ($\mu = 1$GeV)	−0.03(5)	107.3(2.2)
this work ($\mu = 2$ GeV)	−0.13(5)	107.3(2.2)
this work ($\mu = \infty$)	−0.25(5)	107.3(2.2)

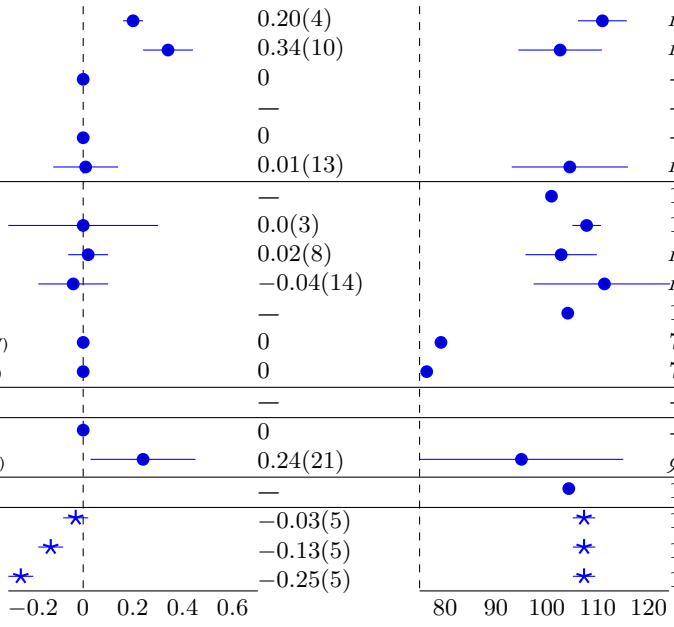


TABLE 5.12: Comparison of determinations of the (scale dependent) large-N_c ChPT LEC Λ_1 and the scale independent combination of Λ_1 and F^0, where eq. (5.54) refers to the NLO result from literature pion and kaon decay constants. The values indicated in Italics have been computed by us from F_0 and Λ_1 with naive error propagation. The results from this work are presented at three different scales.

ChPT is made and allows to predict some of the LECs. Chen et al. [191] couple large-N_c ChPT at NLO to vector resonances and extract the LECs, including these additional couplings, by simultaneously analysing in this framework radiative decay form factors of light vector mesons and charmonia into pseudoscalar final states. Finally, in their calculation Ding et al. [193] employ coupled gap and Bethe-Salpeter equations.

We also include the values $F^8 = 115.2(1.3)$ MeV and $F^0/\sqrt{1+\Lambda_1} = 104.3(1.1)$ MeV in tabs. 5.11 and 5.12 (labelled as "eq. (5.54)"). These are obtained from the identities

$$(F^8)^2 = \frac{4F_K^2 - F_\pi^2}{3} \quad \text{and} \quad (F^0)^2 = \frac{2F_K^2 + F_\pi^2}{3}\,(1+\Lambda_1)\,, \tag{5.54}$$

which hold at NLO in large-N_c ChPT [32], using the values $F_{K^+}/F_{\pi^+} = 1.193(2)$ [152] and $F_{\pi^+} = 92.1(8)$ MeV [19, 152] as input, neglecting electromagnetic and isospin breaking effects. The perfect agreement with our lattice QCD determination of F^8 and the agreement on the $1.3\,\sigma$ level for $F^0/\sqrt{1+\Lambda_1}$ indicates that NLO large-N_c ChPT is a good approximation for these quantities, at least near the physical quark mass point.

Overall, we find reasonable agreement between our results and NLO large-N_c ChPT determinations relying on experimental input. However, in the latter case the errors are often not easily quantifiable. For quantities that depend on the scale, the comparison should be made with our values determined at low scales (for which Λ_1 vanishes). Note that the mixing angles and fundamental decay constants F^q in the flavour basis all depend on the QCD scale, unless the FKS approximation is used. The other determinations, which rely on more complicated experimental analyses (incorporating processes at various scales), differ more significantly, in particular, for the scale independent octet decay constant and the octet mixing angle. The theoretical effort involved in working out the higher orders not withstanding, in general, it seems that the data available (lattice or experimental) are not sufficient to constrain the many parameters of NNLO large-N_c ChPT.

Hence, these studies tend to have larger errors while still relying on assumptions, such as setting individual parameters to zero [181, 182]. The lattice can help, for example, by determining the (QCD renormalization scale dependent) OZI violating parameters Λ_1 and Λ_2.

Relation to other lattice results. Matrix elements of the η and η' have previously been computed on the lattice in the context of an exploratory study of $D_s \to \eta, \eta'$ semileptonic decays [48]. However, only two ensembles were utilized, with pion masses far away from the physical point. ETMC calculated the decay constants utilizing an indirect approach in [51]. This study was updated in [52] to include a continuum extrapolation, employing seventeen $N_f = 2 + 1 + 1$ gauge ensembles at three different lattice spacings. Their results for the masses were discussed in sec. 5.5.1. Due to the level of noise in the axialvector channels, they utilize the FKS scheme to access the decay constants via the pseudoscalar matrix elements of the η and η' states [184]. Before summarizing their results, we will discuss the assumptions they make.

The FKS approximation neglects all OZI violating terms. This amounts to setting $\Lambda_1 = 0$. Rotating eq. (5.54) into the flavour basis results in

$$(F^\ell)^2 = F_\pi^2 + \frac{2}{3}\Lambda_1(2F_K^2 + F_\pi^2) \quad \text{and} \quad (F^s)^2 = 2F_K^2 - F_\pi^2 - \frac{1}{3}\Lambda_1(2F_K^2 + F_\pi^2). \tag{5.55}$$

Setting $\Lambda_1 = 0$ and plugging the experimental ratio $F_{K^+}/F_{\pi^+} = 1.193(2)$ [152] into eq. (5.55) gives

$$F^\ell/F_\pi = 1 \quad \text{and} \quad F^s/F_K = \sqrt{2 - \frac{F_{\pi^+}^2}{F_{K^+}^2}} = 1.139(2). \tag{5.56}$$

In the FKS model the flavour basis AWIs (6.5) factorize into anomalous and non-anomalous contributions and one obtains the relations [183, 184]

$$2m_\ell H^\ell = M_\pi^2 F^\ell \qquad \text{and} \qquad 2m_s H^s = (2M_K^2 - M_\pi^2)F^s, \tag{5.57}$$

where $H^q = \sqrt{(H_\eta^q)^2 + (H_{\eta'}^q)^2}$ for $q = \ell, s$ and the $H_{\eta^{(\prime)}}^q$ are defined in eq. (2.115). Therefore, in this approximation, in the flavour basis the mass mixing angle $\phi_{\text{SU(3)}}$, the mixing angles ϕ^ℓ and ϕ^s and the mixing angles for pseudoscalar matrix elements are all equal: $\phi_{\text{SU(3)}} = \phi = \phi_s = \phi_\ell = \phi_{\text{PS}}$, where [196]

$$\phi_{\text{SU(3)}} = \arcsin\sqrt{\frac{\left(M_{\eta'}^2 - (2M_K^2 - M_\pi^2)\right)(M_\eta^2 - M_\pi)}{(M_{\eta'}^2 - M_\eta^2)(2M_K^2 - 2M_\pi^2)}} = 42.4^\circ. \tag{5.58}$$

ETMC [52] compute the mixing angle from the averaged ratios of their pseudoscalar matrix elements:

$$\tan\phi_{\text{PS}} = \sqrt{\tan\phi_{\text{PS}}^\ell \tan\phi_{\text{PS}}^s} \quad \text{with} \quad \tan\phi_{\text{PS}}^\ell = \frac{H_{\eta'}^\ell}{H_\eta^\ell} \quad \text{and} \quad \tan\phi_{\text{PS}}^s = -\frac{H_\eta^s}{H_{\eta'}^s}. \tag{5.59}$$

The decay constants are obtained, using eq. (5.57). Subsequently, the ratios F^ℓ/F_π, F^s/F_K and the angle ϕ are extrapolated, using a fit that is linear in the quark masses and quadratic in the lattice spacing. At the physical point ETMC obtain

$$\text{ETMC [52]:} \qquad F^\ell/F_\pi = 0.960(59), \quad F^s/F_K = 1.143(24) \quad \text{and} \quad \phi = 38.8(3.3)^\circ, \tag{5.60}$$

which within errors agrees with the values computed within the FKS model in eqs. (5.56) and (5.58)

from the experimental kaon and pion decay constants and the experimental meson masses, respectively. The ETMC results also agree well with our results, that are obtained directly from the axialvector matrix elements, at the scale $\mu = 1$ GeV (see tab. 5.10), where we find $\Lambda_1 \approx 0$, whereas at higher scales the two sets of results differ somewhat.

Test of the FKS approximation, away from the physical point. The results in tabs. 5.10, 5.11 and 5.12 (see also eqs. (5.56) and (5.58)) show that our ab-initio values determined at $\mu = 1$ GeV agree well with those derived by employing the FKS scheme. We can go further and directly check the relations (5.57), (5.59) and (5.55), also away from the physical point. Figure 5.6 displays the two decay constants and two angles in the flavour basis determined from the pseudoscalar matrix elements against the direct results for a range of ensembles at $\mu = 1$ GeV. The values for F^ℓ and F^s obtained from the pion and kaon decay constants are also shown. Modulo the large errors for some ensembles, there is reasonable agreement between the direct results for the angles and the FKS expectation, with $\phi_{\mathrm{SU(3)}} \approx \phi_{\mathrm{PS}}$. Qualitative agreement is also found for the decay constants, however, some scatter in the results is visible, which may be due to discretization effects and/or the limitations of the FKS approximation. This is less significant for F^s and it is striking how well this quantity is reproduced by the combination $\sqrt{2F_K^2 - F_\pi^2}$. However, at higher scales, where Λ_1 can no longer be neglected, scale dependent quantities cannot be reliably extracted with any precision using the FKS method, as indicated in the comparison tables. In particular, for the singlet decay constant we observe the difference between the results at high and low scales $2(F^0(1\,\mathrm{GeV}) - F^0(\infty))/(F^0(1\,\mathrm{GeV}) + F^0(\infty)) = 12.9\,\%$. We remark that previously it was unclear at what scale the FKS approximation holds and this led to an additional unquantifiable uncertainty in phenomenological analyses.

Our direct QCD results can be used as input to theory calculations and we consider one important example in sec. 5.5.4, namely the light-cone sum rule computation of the $\gamma\gamma^* \to \eta^{(\prime)}$ transition form factors.

5.5.3 Large-N_c low energy constants

As part of our analysis we are able to extract the large-N_c ChPT LECs up to NLO. The singlet mass in the chiral limit M_0 and the two OZI-rule violating parameters Λ_1 and Λ_2 are all $\mathcal{O}(1/N_c)$ in the power counting. Besides these large-N_c specific LECs, L_5, L_8 and the decay constant in the chiral limit, F, also appear. These are present in ordinary SU(3) ChPT, although, their values can differ. In particular, L_5 and L_8 depend on the ChPT renormalization scale in the SU(3) theory, however, such scale dependence only arises at NNLO in large-N_c ChPT.

The $\mathcal{O}(1/N_c)$ LECs extracted from our NLO fits at $\mu = \infty$ for $N_f = 3$ read:

$$M_0 = 761\,(^{13}_{21})_{\mathrm{stat}}\,(^{18}_{11})_{\mathrm{syst}}\,(11)_{t_0}\,\mathrm{MeV}, \tag{5.61}$$

$$\Lambda_1 = -0.25\,(^{1}_{4})_{\mathrm{stat}}\,(^{6}_{2})_{\mathrm{syst}}\,, \tag{5.62}$$

$$\Lambda_2 = 0.11\,(^{5}_{5})_{\mathrm{stat}}\,(^{7}_{10})_{\mathrm{syst}}\,, \tag{5.63}$$

where the ChPT and lattice spacing errors have been combined into a single systematic uncertainty. The dependence of these quantities on the QCD renormalization scale is discussed in sec. 5.5.2. To aid comparison with literature values we consider the scale independent combinations [32]:

$$M_0/\sqrt{1+\Lambda_1} = 877\,(^{12}_{10})_{\mathrm{stat}}\,(^{21}_{8})_{\mathrm{syst}}\,(13)_{t_0}\,\mathrm{MeV} \quad \text{and} \quad \tilde{\Lambda} = \Lambda_1 - 2\Lambda_2 = -0.46(19). \tag{5.64}$$

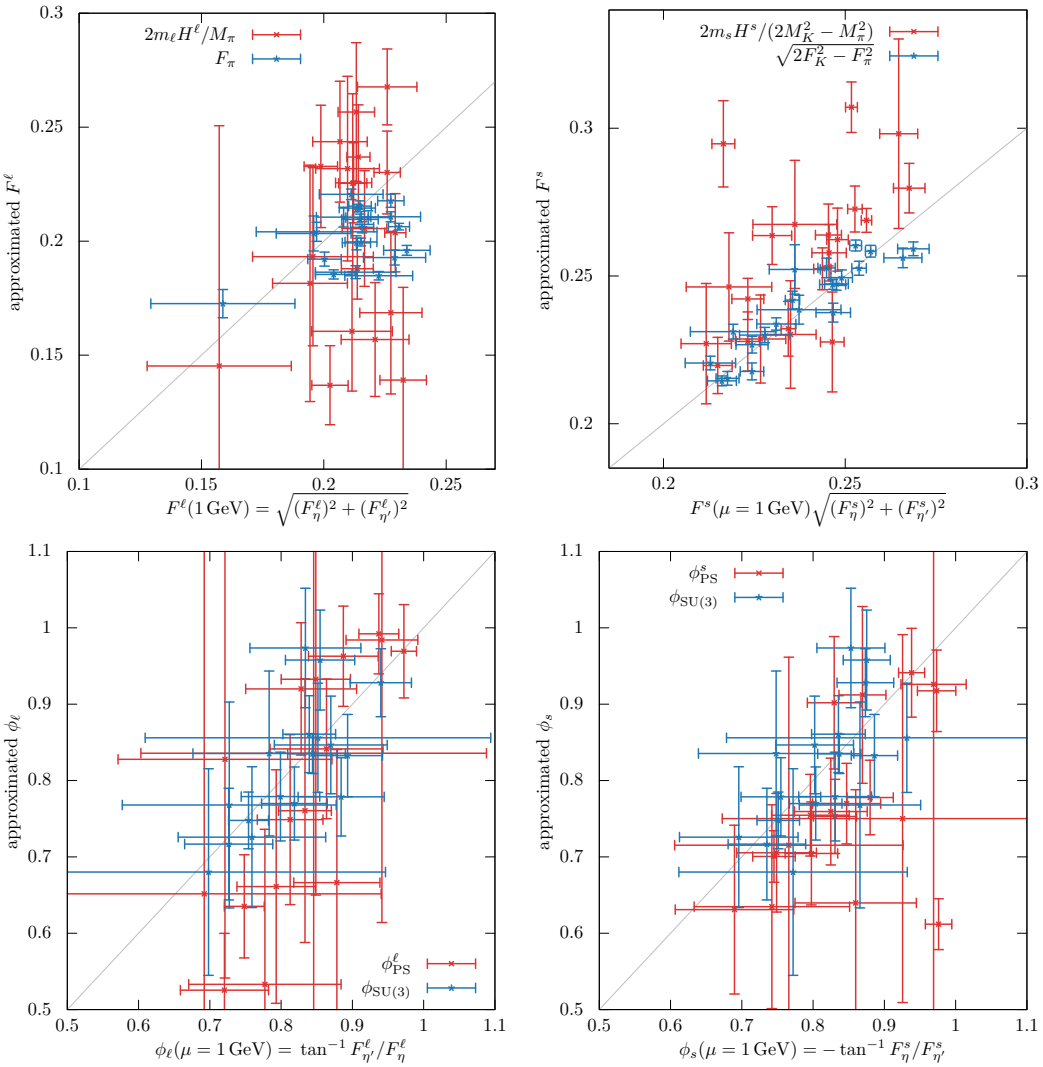


FIGURE 5.6: Light and strange decay constants (top) and angles (bottom) determined on each ensemble at the QCD scale $\mu = 1\,\mathrm{GeV}$. The central values and errors in the x-direction indicate the results obtained directly from the axialvector matrix elements, while the position and error in the y-direction indicates the values constructed from the pseudoscalar matrix elements (red, eqs. (5.57) and (5.59)). For the decay constants, the expectations derived from combinations of the pion and kaon masses and decay constants (blue, eq. (5.55)) are also displayed, while for the angles $\phi_{\mathrm{SU(3)}}$ (eq. (5.58)) is also shown. The blue points have been shifted slightly to the right for better visibility.

Previous determinations of these quantities include:

$$\begin{aligned}
&\text{Leutwyler [32]:} && M_0/\sqrt{1+\Lambda_1} \approx 899\,\text{MeV} && \text{and} && \tilde{\Lambda} = -0.31, && (5.65)\\
&\text{Benayoun et al. [188]:} && && && \tilde{\Lambda} = -0.42(6), && (5.66)\\
&\text{Guo et al. [182]:} && M_0/\sqrt{1+\Lambda_1} = 804(80)\,\text{MeV} && \text{and} && \tilde{\Lambda} = -0.37(17), && (5.67)\\
&\text{Bickert et al. [181]:} && M_0/\sqrt{1+\Lambda_1} = 950(7)\,\text{MeV} && \text{and} && \tilde{\Lambda} = -0.34(5), && (5.68)
\end{aligned}$$

where except for [181], we have constructed these scale independent quantities from the individual results quoted in the publications. Our central value for $M_0/\sqrt{1+\Lambda_1}$ is larger than the result of [182], which utilizes lattice data, however, considering the large uncertainty quoted in this reference, there is no significant disagreement. The determination from [181] lies roughly two standard deviations higher, where the LECs in this study are determined from experimental input which includes the singlet and non-singlet meson masses and non-singlet decay constants. We also find some disagreement with [181] when comparing predictions for the decay constants and angles, cf. tab. 5.11. Interestingly, our value for the combination of OZI-violating LECs $\tilde{\Lambda}$ is in good agreement with the above determinations.

For the decay constant in the chiral limit, we obtain

$$F = 87.71\,\binom{1.44}{1.57}_{\text{stat}}\,\binom{2.69}{81}_{\text{syst}}\,(1.31)_{t_0}\,\text{MeV}. \tag{5.69}$$

This result agrees within errors with the NLO values presented in [182] and [181]:

$$\text{Guo et al. [182]:}\quad F = 92.1(6)\,\text{MeV}, \qquad \text{Bickert et al. [181]:}\quad F = 90.73(11)\,\text{MeV}. \tag{5.70}$$

However, the corresponding NNLO analyses give somewhat lower values of F,

$$\text{Guo et al. [182]:}\quad F = 80.8(6.3)\,\text{MeV}, \qquad \text{Bickert et al. [181]:}\quad F = 79.46(6.59)\,\text{MeV}, \tag{5.71}$$

which within errors still agree with our result (5.69). In the simulations of [197] with $N_f = 4$ sea quarks, the number of colours $N_c \in \{2,3,4,5,6\}$ is varied. The pion decay constant and its mass are then fitted to the NNLO large-N_c U(4) ChPT prediction. From the expected dependence on N_f/N_c and $1/N_c^2$ (neglecting N_f^2/N_c^2 terms), the even lower value

$$\text{Hernandez et al. [197]:}\quad F(N_f = 3, N_c = 3) = 68(7)\,\text{MeV} \tag{5.72}$$

is inferred for $N_f = N_c = 3$ at the lattice spacing $a \approx 0.075\,\text{fm}$.

The additional terms appearing at NNLO comprise chiral logs and expressions which include the LECs L_4, L_6, L_7, L_{18} and L_{25}. In particular, in [198], it is argued that L_4 is anti-correlated with the decay constant in the chiral limit as seen, e.g., in fits to experimental data in [199]. Thus, neglecting NNLO contributions including L_4-terms may lead to larger values of F. However, also our fits in app. D, including only the NNLO loop contributions (see the discussion of sec. 5.3 as well as below), gives $F = 79.0(3.8)\,\text{MeV}$. In conclusion, both these effects may account for the reduction of the value of F within NNLO analyses, in comparison to results from NLO parametrizations.

For the LECs L_5 and L_8, we find

$$L_5 = 1.66\,(11)_{\text{stat}}\,(20)_{\text{syst}} \cdot 10^{-3} \quad \text{and} \quad L_8 = 1.08\,(9)_{\text{stat}}\,(9)_{\text{syst}} \cdot 10^{-3}. \tag{5.73}$$

These values agree reasonably well with those obtained from other NLO large-N_c ChPT studies, e.g.,

$$\text{Leutwyler [32]:} \quad L_5 = 2.2 \cdot 10^{-3} \quad \text{and} \quad L_8 = 1.0 \cdot 10^{-3}, \tag{5.74}$$

$$\text{Guo et al. [182]:} \quad L_5 = 1.47(29) \cdot 10^{-3} \quad \text{and} \quad L_8 = 1.08(6) \cdot 10^{-3}, \tag{5.75}$$

$$\text{Bickert et al. [181]:} \quad L_5 = 1.86(6) \cdot 10^{-3} \quad \text{and} \quad L_8 = 0.78(5) \cdot 10^{-3}. \tag{5.76}$$

A comparison can also be made with the LECs obtained using SU(3) ChPT. The LECs in the SU(3) and large-N_c theories are related via [33, 181, 200]

$$L_5(\mu_{\text{EFT}}) = L_5^{\text{SU(3)}}(\mu_{\text{SU(3)}}) + \frac{3}{8}\frac{1}{16\pi^2}\ln\left(\frac{\mu_{\text{SU(3)}}}{\mu_{\text{EFT}}}\right) \tag{5.77}$$

$$L_8(\mu_{\text{EFT}}) = L_8^{\text{SU(3)}}(\mu_{\text{SU(3)}}) + \frac{5}{48}\frac{1}{16\pi^2}\ln\left(\frac{\mu_{\text{SU(3)}}}{\mu_{\text{EFT}}}\right) + \frac{1}{12}\frac{1}{16\pi^2}\ln\left(\frac{\mu_{\text{match}}}{\mu_{\text{EFT}}}\right), \tag{5.78}$$

where $\mu_{\text{SU(3)}}$ is the SU(3) ChPT scale, μ_{EFT} is the scale of large-N_c ChPT (which is ill-defined at NLO) and μ_{match} is the scale at which the two theories are matched. We set $\mu_{\text{EFT}} = \mu_{\text{SU(3)}} = \mu_{\text{match}} = 0.770\,\text{GeV}$ such that $L_{5,8}(\mu_{\text{EFT}}) = L_{5,8}^{\text{SU(3)}}(\mu_{\text{SU(3)}})$. A direct comparison can then be made with the SU(3) values obtained in [201] from a $N_f = 2+1$ lattice study of the pion and kaon masses and decay constants. Here, we quote the values presented in the FLAG review [152] for $\mu_{\text{SU(3)}} = 0.770\,\text{GeV}$

$$\text{MILC [201]:} \quad L_5^{\text{SU(3)}} = 0.98(38) \cdot 10^{-3} \quad \text{and} \quad L_8^{\text{SU(3)}} = 0.42(27) \cdot 10^{-3}. \tag{5.79}$$

For $\mu_{\text{EFT}} < 0.770\,\text{GeV}$ and $\mu_{\text{match}} > 0.770\,\text{GeV}$ the agreement with our (scale independent) results improves.

Overall, our results for the large-N_c ChPT LECs are reasonably consistent with literature values. A direct comparison of NLO and NNLO results is difficult due to the scale dependence which arises at NNLO. Results from our fits including the NNLO loop contributions can be found in app. D. The inferior $\chi^2/N_{\text{df}} = 2.56$ indicates that this parametrization does not describe our data well and additional NNLO terms are required for a consistent description of the data. Note that this analysis gives values for the LECs L_5 and L_8 (see eq. (D.9)) that are slightly smaller and larger than our NLO values quoted in eq. (5.73), respectively. As also observed in our analysis, it appears difficult to reliably pin down the many NNLO LECs, and usually priors or assumptions are needed to carry out such fits, giving rise to additional uncertainties, see for example the scatter of NNLO results in [181, 182].

5.5.4 Transition form factors $F_{\gamma\gamma^*\to\eta}(Q^2)$ and $F_{\gamma\gamma^*\to\eta'}(Q^2)$

Finally we present an application of our decay constant results to compute the photoproduction transition form factor (TFF). Photoproduction is the simplest hard process involving $\eta^{(\prime)}$ mesons. The corresponding TFFs $F_{\gamma\gamma^*\to\eta}(Q^2)$ and $F_{\gamma\gamma^*\to\eta'}(Q^2)$ at large photon virtualities $Q^2 = |q^2|$ have been studied, e.g., in [193, 202, 203] and assessed phenomenologically for instance in [185, 192]. On the lattice the pion TFF has been studied [204] and work on the η and η' TFFs have started [205]. This is motivated by the TFF contributions to the theoretical estimate of the hadronic light-by-light contribution to the anomalous magnetic moment of the muon.

The special role of the transition form factors as "gold plated" observables for the study of meson light-cone distribution amplitudes (LCDAs) is widely recognized. The corresponding theory is similar to that for the more easily accessible pion transition form factor $F_{\gamma\gamma^*\to\pi^0}(Q^2)$ but the non-perturbative input encoded in the LCDAs is more complicated. The two outstanding issues are, first, whether the η and η' LCDAs follow the same mixing pattern as the decay constants at a

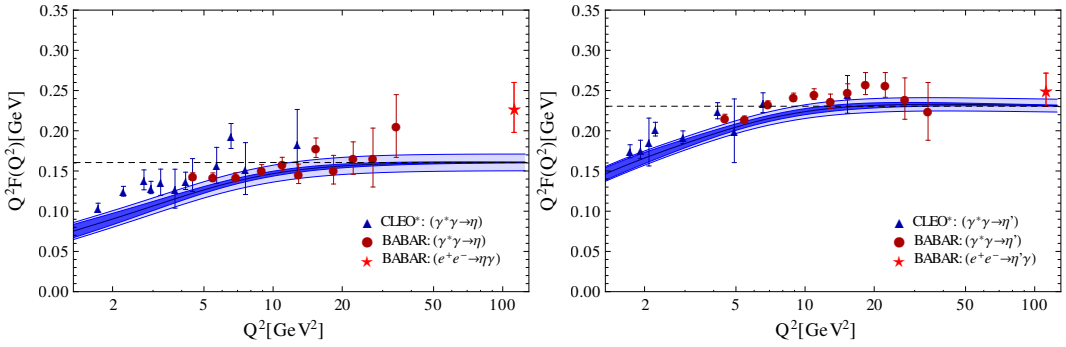


FIGURE 5.7: Predictions for the transition form factor $F_{\gamma\gamma^*\to\eta}(Q^2)$ (left) and $F_{\gamma\gamma^*\to\eta'}(Q^2)$ (right) following the approach in [203], using the decay constants determined in our work as input parameters.

low scale, and, second, whether η' contains a significant two-gluon component, see, e.g., [202, 203] for a discussion.

In fig. 5.7 we show the QCD prediction for the form factors, employing the light-cone sum rules calculation presented in [203]. The results are compared to the experimental data for the space-like form factors by the CLEO [206] and BaBar [22] collaborations, and we also include BaBar's time-like data points at $q^2 = 112\,\text{GeV}^2$ [21] as stars.

The difference with respect to the original calculation is that in fig. 5.7 the lattice values of the decay constants determined in this work are used as an input, whereas in [203] the decay constants were taken from [183] under the (ad hoc) assumption that they correspond to the scale 1 GeV. Using our lattice results removes this scale uncertainty. The overall difference is small, due to the good agreement with these results at 1 GeV with [183], cf. sec. 5.5.2 and tab. 5.11.

In view of the experimentally available range of Q^2, employing $N_f = 4$ appears reasonable. Therefore, we run our $N_f = 3$ values for $F^0_{\eta^{(\prime)}}$ (see tab. E.2) down to $\mu_0 = 1.51\,\text{GeV}$ [207] (see sec. 2.4.3), where we match to the $N_f = 4$ theory. This value is then taken as an input for the LCSR calculation.

The LCSR technique involves a certain model dependence in the calculation of the power-suppressed contributions to the form factors. This is indicated by the dark blue shaded regions in fig. 5.7 and can be regarded as an (at present) irreducible uncertainty of such calculations. The total uncertainty including that of the lattice values for the decay constants is shown in light blue (added in quadrature). Starting around $Q^2 \sim 10\,\text{GeV}^2$ this uncertainty dominates over the model dependence.

The calculation is carried out assuming that the shapes of the LCDAs of the η and the η' at low scales are the same as that of the pion. In our calculation, following [203], the corresponding parameters are chosen from the fit to the pion transition form factor in the same approach [208]. Moreover, the two-gluon LCDA at the low scale is set to zero. Under these assumptions the only additional non-perturbative input at the leading-twist level are the decay constants that we computed here. The comparison of fig. 5.7 between the predictions for $F_{\gamma\gamma^*\to\eta}(Q^2)$ and $F_{\gamma\gamma^*\to\eta'}(Q^2)$ and experimental data shows that the above approximation appears to work relatively well, although there is some tension with the available data. In the asymptotic limit $Q^2 \to \infty$ the dependence on the shape of the LCDAs is removed, and the decay constants provide the only necessary input:

$$\lim_{Q^2\to\infty} Q^2 F_{\gamma\gamma^*\to\eta^{(\prime)}} = \frac{2}{\sqrt{3}}\left(F^8_{\eta^{(\prime)}} + 2\sqrt{2}F^0_{\eta^{(\prime)}}(\mu=\infty, N_f=4)\right). \tag{5.80}$$

We obtain for $N_f = 4$:

$$\lim_{Q^2\to\infty} Q^2 F_{\gamma\gamma^*\to\eta}(Q^2) = 160.5(10.0)\,\text{MeV},$$
$$\lim_{Q^2\to\infty} Q^2 F_{\gamma\gamma^*\to\eta'}(Q^2) = 230.5(10.1)\,\text{MeV}. \tag{5.81}$$

These asymptotic values are shown as dashed lines in fig. 5.7. Regarding the latter form factor, it is particularly important to take the scale dependence of the singlet decay constant into account. This explains the relatively large values obtained for $Q^2 F_{\gamma\gamma^*\to\eta'}(Q^2)$ when neglecting the scale evolution, see, e.g., the predictions in [52]. Also the matching to the $N_f = 4$ theory somewhat reduces the value. As already emphasized in [203], the effect due to the scale dependence is enhanced for the η' form factor because in this case the two terms in eq. (5.80) have opposite signs.

The current experimental accuracy is not yet sufficient to draw definite conclusions. In the future, due to an increase of the statistics by a large factor and improved particle identification, the Belle II collaboration will be able to measure the pseudoscalar meson transition form factors with much higher precision [27]. A disagreement with QCD calculations using lattice input for the decay constants would either indicate qualitative differences between the shapes of the LCDAs for different pseudoscalar mesons or the presence of a large two-gluon contribution. Both would have important consequences for other hard processes involving η and η' mesons, e.g., in weak B-decays.

6 Gluonic matrix elements and axial Ward identities

In this final chapter we return to one of the fundamental features of the non-perturbative QCD vacuum, laid out in sec. 2.1.5, which is the anomalous breaking of the $\mathrm{U}(1)_A$ symmetry and the non-conservation of the axialvector current resulting from it, see eq. (2.34). The singlet axial Ward identity has received a lot of attention, also in different settings, e.g., regarding the spin structure of the nucleon [209]. The AWIs are operator identities and in this chapter we confront them with the physical η and η' states and test the validity on the level of matrix elements.

We start by discussing the axial Ward identity both in the octet/singlet as well as in the flavour basis on the lattice in sec. 6.1 and comment on the renormalization of the operators that appear. There, we also check the simpler octet case that relates just the pseudoscalar and axialvector matrix elements but does not contain the gluonic anomaly contribution $\langle\Omega|\omega|\eta^{(\prime)}\rangle$. Motivated by the success of this comparison in the octet case, we then first compute the anomaly contribution in the singlet channel indirectly from the fermionic matrix elements in sec. 6.2. In sec. 6.3 the NLO large-N_c ChPT formulae for the extrapolation to the physical point is derived and the continuum and chiral limit is taken. We then return to the complicated renormalization of the singlet AWI in sec. 6.4 to finally test the Ward identity. To this end we determine the gluonic matrix elements directly from correlators of the topological susceptibility with the η and η' states. For the renormalization of the susceptibility and to disentangle the mixing with the axialvector current we analyse the topological susceptibility. Equipped with the renormalization factors we can proceed and compare both sides of the singlet AWI in sec. 6.5. Finally, we summarize our results on pseudoscalar and anomalous matrix elements in sec. 6.6 and compare to literature values.

This chapter has already been published in similar or verbatim form in [60].

6.1 The axial Ward identities

The AWIs between renormalized operators (indicated by a hat) read

$$\partial_\mu \widehat{A}^a_\mu = \widehat{\left(\overline{\psi}\gamma_5\{M, t^a\}\psi\right)} + \sqrt{2N_f}\delta^{a0}\widehat{\omega}, \tag{6.1}$$

where $M = \mathrm{diag}(m_\ell, m_\ell, m_s)$ is the quark mass matrix, $a \in \{0, 1, \ldots, 8\}$, and the topological charge density is defined as

$$\omega(x) = -\frac{1}{16\pi^2}\,\mathrm{tr}\left[F_{\mu\nu}(x)\widetilde{F}_{\mu\nu}(x)\right] = -\frac{1}{32\pi^2}F^a_{\mu\nu}(x)\widetilde{F}^a_{\mu\nu}(x) = -\frac{1}{64\pi^2}\epsilon_{\mu\nu\rho\sigma}F^a_{\mu\nu}(x)F^a_{\rho\sigma}(x). \tag{6.2}$$

Since different conventions are used in the literature, for clarity we have written the right-hand side in three different ways. Regarding the octet and singlet AWIs, eq. (6.1) corresponds to

$$\partial_\mu \widehat{A}^8_\mu = \frac{2}{3}\left(\widehat{m}_\ell + 2\widehat{m}_s\right)\widehat{P}^8 - \frac{2\sqrt{2}}{3}\delta\widehat{m}\widehat{P}^0, \tag{6.3}$$

and

$$\partial_\mu \widehat{A}^0_\mu = \frac{2}{3}\left(2\widehat{m}_\ell + \widehat{m}_s\right)\widehat{P}^0 - \frac{2\sqrt{2}}{3}\delta\widehat{m}\widehat{P}^8 + \sqrt{6}\,\widehat{\omega}, \tag{6.4}$$

respectively. In the octet/singlet basis only the singlet AWI receives a contribution from the anomaly. Note, however, that the octet AWI receives pseudoscalar singlet contributions through the mass term and vice versa. Using eq. (2.103), the corresponding AWIs in the flavour basis can be written as

$$\partial_\mu \widehat{A}^s_\mu = 2\widehat{m}_s\widehat{P}^s + 2\,\widehat{\omega}, \quad \partial_\mu \widehat{A}^\ell_\mu = 2\widehat{m}_\ell\widehat{P}^\ell + 2\sqrt{2}\,\widehat{\omega}. \tag{6.5}$$

From a flavour perspective these are simpler because the quark flavours decouple, however, the anomaly contribution now enters both AWIs and the renormalization is more complicated.

We determine our quark masses, using the non-singlet AWIs for the light pseudoscalar mesons, eq. (3.6). We carry out the complete $\mathcal{O}(a)$ improvement of the currents, so that the above relations hold up to $\mathcal{O}(a^2)$ corrections. For this the (combinations of) improvement coefficients c_A, $b_A - b_P$ and $\tilde{b}_A - \tilde{b}_P$ are required, all of which are known non-perturbatively [119–121], see also sec. 2.4.2. The lattice AWI quark masses are related to the continuum masses via

$$\widehat{m}_q(\mu) = \frac{Z_A}{Z_P(\mu)}\widetilde{m}_q. \tag{6.6}$$

Again Z_A/Z_P is known non-perturbatively in the RGI scheme [210] and can be related to the $\overline{\text{MS}}$ scheme at a scale μ perturbatively at the five-loop level [129] whenever this is needed.

The octet AWI between lattice matrix elements reads

$$\partial_\mu \left\langle \Omega \left| A^8_\mu \right| \mathcal{M} \right\rangle = \frac{2}{3}\left(\widetilde{m}_\ell + 2\widetilde{m}_s\right)\left\langle \Omega \left| P^8 \right| \mathcal{M} \right\rangle - \frac{2\sqrt{2}}{3}\delta\widetilde{m}\, r_P \left\langle \Omega \left| P^0 \right| \mathcal{M} \right\rangle, \tag{6.7}$$

where $\delta\widetilde{m} = \widetilde{m}_s - \widetilde{m}_\ell$ denotes the difference between the lattice AWI quark masses. This expression is only non-trivial for η and η' states. The (scale independent) ratio $r_P = Z^s_P/Z_P$ appears since the renormalization of the singlet relative to that of the non-singlet pseudoscalar current can differ at $\mathcal{O}(g^6)$ for Wilson fermions. In addition to the known improvement coefficients, also c^s_P (which to $\mathcal{O}(a)$ is equivalent to g_P, see eq. (2.117)), d_P and $\tilde{d}_P$ contribute, while for $m_s \neq m_\ell$, f_A and f_P appear too. For $m_s = m_\ell$, the Ward identity is trivial if applied to the $\eta' = \eta_0$ state. Note that the left-hand side of eq. (6.7) can be replaced with

$$\partial_\mu \left\langle \Omega \left| A^8_\mu \right| \mathcal{M} \right\rangle = Z_A^{-1}\partial_\mu \left\langle \Omega \left| \widehat{A}^8_\mu \right| \mathcal{M} \right\rangle = Z_A^{-1} M^2_{\mathcal{M}} F^8_{\mathcal{M}}, \tag{6.8}$$

as discussed around eq. 2.118. Hence, the combination on the right-hand side of eq. (6.7) does not depend on the momentum of the meson $\mathcal{M}$.

In fig. 6.1 we check the octet AWI (6.7) at zero momentum directly against our data, utilizing the quark masses computed according to eq. 3.6. Note that the equality should hold without any renormalization, up to the ratio r_P. For the comparison, we set $r_P = 1$, $c^s_P = d_P = \tilde{d}_P = f_P = 0$ and $f_A = -0.689\,g^6$. The value of the latter coefficient is taken from the central fit of sec. 5.3 (fit 7). Throughout, we find reasonable agreement between the left and right-hand sides of the Ward identity, as shown in fig. 6.1. Only the pseudoscalar combination for the η' tends to result in slightly smaller values than those of the derivative of the axialvector current. Within our precision, we conclude that indeed $r_P = 1$ to a good approximation and that the effect of the three (for $m_s = m_\ell$) or four (for $m_s \neq m_\ell$) unknown improvement terms is moderate, even for our coarsest lattice spacing.

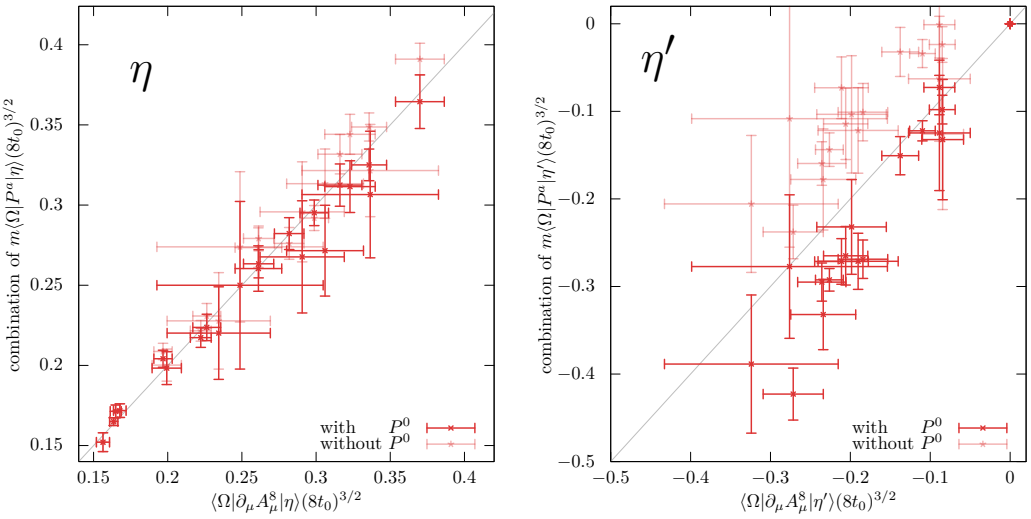


FIGURE 6.1: Check of the octet AWI (6.7) for the η and η' states. Light points correspond to the case when the singlet contribution to the octet AWI is neglected.

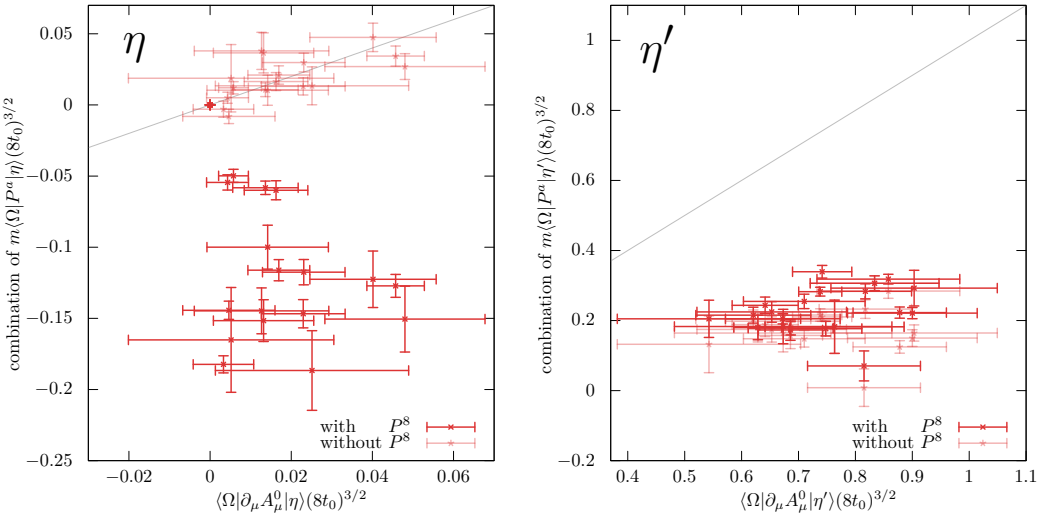


FIGURE 6.2: Fermionic contributions to the singlet AWI for the η and η' states. Light points correspond to only the singlet contribution whereas the darker points correspond to all pseudoscalar terms in eq. (6.4), but the anomalous contribution is neglected.

In the singlet case, due to the anomaly contribution to eq. (6.4), we would expect that

$$\partial_\mu \left\langle \Omega \left| A_\mu^0 \right| \mathcal{M} \right\rangle \neq 2\overline{\widetilde{m}}\, r_P \left\langle \Omega \left| P^0 \right| \mathcal{M} \right\rangle - \frac{2\sqrt{2}}{3}\delta\widetilde{m} \left\langle \Omega \left| P^8 \right| \mathcal{M} \right\rangle , \tag{6.9}$$

where $\overline{\widetilde{m}}$ denotes the average lattice AWI quark mass. Again we set $r_P = 1$ and ignore any unknown improvement terms. The comparison at zero momentum is shown in fig. 6.2. The difference is large for both states and does not significantly depend on the lattice spacing but mostly on the quark masses. This rules out the incomplete singlet $\mathcal{O}(a)$ improvement as a major cause for the disagreement. Interestingly, in the case of the η, the singlet pseudoscalar contribution coincides with the left-hand side of eq. (6.4): the (in this case) large octet pseudoscalar matrix element approximately cancels against the anomaly term. For the η' the octet contribution is much smaller and no such effect can be seen. In both cases, contributions from the anomalous matrix elements $\langle\Omega|\omega|\eta^{(\prime)}\rangle$ are large in comparison to the terms involving pseudoscalar matrix elements, and it is clear that the anomalous term must be included. The gluonic matrix element can be determined simply from the difference observed in these plots, a procedure that does not involve any additional renormalization. We follow this strategy in the next section.

6.2 Fermionic determination of $\langle\Omega|2\omega|\eta\rangle$ and $\langle\Omega|2\omega|\eta'\rangle$

Rather than determining the renormalized matrix elements

$$a_{\mathcal{M}} = 2\langle\Omega|\widehat{\omega}|\mathcal{M}\rangle \tag{6.10}$$

directly using gluonic correlators, we first compute them via the renormalized singlet AWI (6.4):

$$\begin{aligned} a_{\mathcal{M}}(\mu) &= \sqrt{\frac{2}{3}} Z_A^s(\mu)\partial_\mu \left\langle \Omega \left| A_\mu^0 \right| \mathcal{M} \right\rangle + \frac{2\sqrt{2}}{\sqrt{3}} Z_A \left[\frac{\sqrt{2}}{3}\delta\widetilde{m} \left\langle \Omega \left| P^8 \right| \mathcal{M} \right\rangle - r_P\overline{\widetilde{m}} \left\langle \Omega \left| P^0 \right| \mathcal{M} \right\rangle \right] \\ &= \sqrt{\frac{2}{3}} M_{\mathcal{M}}^2 F_{\mathcal{M}}^0(\mu) + \frac{4}{3\sqrt{3}}\frac{Z_A}{Z_P}\delta\widetilde{m} H_{\mathcal{M}}^8 - \frac{2\sqrt{2}}{\sqrt{3}} r_P \frac{Z_A}{Z_P}\overline{\widetilde{m}} H_{\mathcal{M}}^0, \end{aligned} \tag{6.11}$$

where $H_{\mathcal{M}}^a$ are the renormalized and $\mathcal{O}(a)$ improved pseudoscalar matrix elements, in analogy to eqs. (2.120) and (2.124). This fermionic definition has the advantage that no knowledge of the renormalization factors $Z_{\omega A}$ and Z_ω is needed (see sec. 6.4 below) and we only assume $r_P = 1$ as has been shown to hold to good accuracy in the previous section. Note that $a_{\mathcal{M}}$ depends on the renormalization scale μ and only in the modified scheme, discussed in sec. 2.4.3 (which corresponds to $\mu \to \infty$), do the gluonic matrix elements become scale independent.

6.3 Continuum limit and extrapolation to the physical point

To extrapolate these results to the physical point a continuum parametrization of their mass dependence is needed. To our knowledge no such large-N_c ChPT result exists, yet, and so we employed the expressions for the masses, eq. (5.11), and decay constants, eqs. (5.24) to (5.27), to solve the octet AWI, eq. (6.7), both in the $N_f = 3$ where the singlet and octet decouple and in the more generic $N_f = 2 + 1$ case for the pseudoscalar matrix elements times the quark masses, $m_q H_q^a$. We substitute these into the singlet AWI to obtain expressions for $a_{\eta'}$ order by order. We describe the calculation in more detail and give the expressions in app. F. Similar to sec. 5.2 we label the continuum limit expressions as $a_{\eta'}(\overline{M}, \delta M|\cdots)$, where the ellipsis represents the low energy constants and the masses are the arguments of the functions.

To include the $\mathcal{O}(a)$ improvement coefficients we proceed similarly as in sec. 2.4.2. We start by defining partially improved matrix elements at $\mu = \infty$:

$$\check{a}_{\eta^{(\prime)}} = \sqrt{\frac{2}{3}} M^2_{\eta^{(\prime)}} F^0_{\eta^{(\prime)}}(\mu) + \frac{4}{3\sqrt{3}} Z_A \delta\widetilde{m} \check{H}^8_{\eta^{(\prime)}} - \frac{2\sqrt{2}}{\sqrt{3}} r_P Z_A \overline{\widetilde{m}} \check{H}^0_{\eta^{(\prime)}}, \tag{6.12}$$

where $\check{H}^a_{\eta^{(\prime)}} = \langle \Omega | P^a | \eta^{(\prime)} \rangle$ are unimproved pseudoscalar lattice matrix elements and we assume $r_P = 1$.

The unknown axialvector coefficients that enter the computation of the singlet decay constants are taken from sec. 5.3, and we set $d_A = b_A + 1.84\, g^4$, $f_A = -0.689\, g^6$ and $\tilde{d}_A = \delta c_A = 0$ (as determined using fit 7). Six improvement coefficients are needed for the pseudoscalar currents: b_P, $\tilde{b}_P$, d_P, $\tilde{d}_P$, f_P and c^s_P. The latter was defined in eq. (2.117) and persists in the chiral limit. We remark again that by replacing $a g_P \mathrm{tr} F_{\mu\nu} \widetilde{F}_{\mu\nu} \mapsto a c^s_P \partial_\mu A^0_\mu$, the definition of the coefficients d_P and $\bar{d}_P$ (and therefore of $\tilde{d}_P$) with respect to [118] is altered.

We then carry out a fit according to

$$\begin{aligned}
\check{a}_{\eta^{(\prime)}}(a, \overline{M}^2, \delta M^2) &= a_{\eta^{(\prime)}}(\overline{M}^2, \delta M^2 | \cdots) \\
&- \frac{2\sqrt{2} Z_A}{\sqrt{3}} \overline{\widetilde{m}} \left[3a\tilde{d}_P \overline{m} \check{H}^0_{\eta^{(\prime)}} + a d_P \frac{1}{\sqrt{3}} \left(\sqrt{2} m_\ell \check{H}^\ell_{\eta^{(\prime)}} + m_s \check{H}^s_{\eta^{(\prime)}} \right) + a c^s_P M^2_{\eta^{(\prime)}} F^0_{\eta^{(\prime)}} \right] \\
&+ \frac{4 Z_A}{3\sqrt{3}} \delta\widetilde{m} \left[3a\tilde{b}_P \overline{m} \check{H}^8_{\eta^{(\prime)}} + a b_P \frac{1}{\sqrt{3}} \left(m_\ell \check{H}^\ell_{\eta^{(\prime)}} - \sqrt{2} m_s \check{H}^s_{\eta^{(\prime)}} \right) + \sqrt{2} a f_P \delta m \check{H}^0_{\eta^{(\prime)}} \right],
\end{aligned} \tag{6.13}$$

In keeping with the rest of our analysis, all dimensionful quantities appearing within this fit are multiplied by the appropriate powers of $\sqrt{8t_0}$. We parametrize the coefficients d_P, $\tilde{d}_P$, f_P and c^s_P (that are functions of g^2) similarly to eq. (5.31) with one parameter each, while b_P and $\tilde{b}_P$ are known non-perturbatively [121]:

$$\begin{aligned}
b_P(\beta = 3.4) &= 1.622(74), & b_P(\beta = 3.46) &= 1.592(213), \\
b_P(\beta = 3.55) &= 1.560(165), & b_P(\beta = 3.7) &= 1.696(78),
\end{aligned} \tag{6.14}$$

$$\begin{aligned}
\tilde{b}_P(\beta = 3.4) &= 0.39(27), & \tilde{b}_P(\beta = 3.46) &= 0.32(20), \\
\tilde{b}_P(\beta = 3.55) &= 0.40(23), & \tilde{b}_P(\beta = 3.7) &= 0.16(13).
\end{aligned} \tag{6.15}$$

The resulting 10-parameter fit is only weakly constrained, however, at NLO in large-N_c ChPT the LECs should be identical to those that we already determined in sec. 5.4.1. Therefore, in analogy to eq. (5.37), we add these results, given in eq. (5.38), as priors to the χ^2 function. The widths σ are set to the statistical and systematic errors, added in quadrature. In analogy to $\tilde{d}_A$ it turns out that we are still unable to resolve $\tilde{d}_P$ and fix $\tilde{d}_P = \tilde{b}_P$ instead.

The fit, shown in fig. 6.3, gives a valid description of the data, with a fully correlated $\chi^2/N_{\mathrm{df}} \approx 34/31 \approx 1.09$. We obtain

$$d_P(g^2) = b_P(g^2) + 6.6(6) g^4, \quad c^s_P(g^2) = -2.4(3) g^4 \quad \text{and} \quad f_P(g^2) = -26(6) g^6 \tag{6.16}$$

for the additional improvement coefficients, setting $\tilde{d}_P = \tilde{b}_P$. The corresponding LECs read

$$\begin{aligned}
L_5 &= 1.95\, \tbinom{7}{3}_{\mathrm{stat}} \cdot 10^{-3}, & L_8 &= 0.97\, \tbinom{4}{6}_{\mathrm{stat}} \cdot 10^{-3}, \\
M_0 &= 1.59\, \tbinom{1}{6}_{\mathrm{stat}} (8t_0^\chi)^{-1/2}, & F &= 0.1881\, \tbinom{9}{24}_{\mathrm{stat}} (8t_0^\chi)^{-1/2} \\
\Lambda_1 &= -0.10\, \tbinom{1}{1}_{\mathrm{stat}}, & \tilde{\Lambda} &= -0.21\, \tbinom{3}{4}_{\mathrm{stat}},
\end{aligned} \tag{6.17}$$

where the errors given are purely statistical and generally small, due to the priors. In particular, Λ_1 is by 2.8 standard deviations larger than its input value (5.38), obtained from the fit to the masses

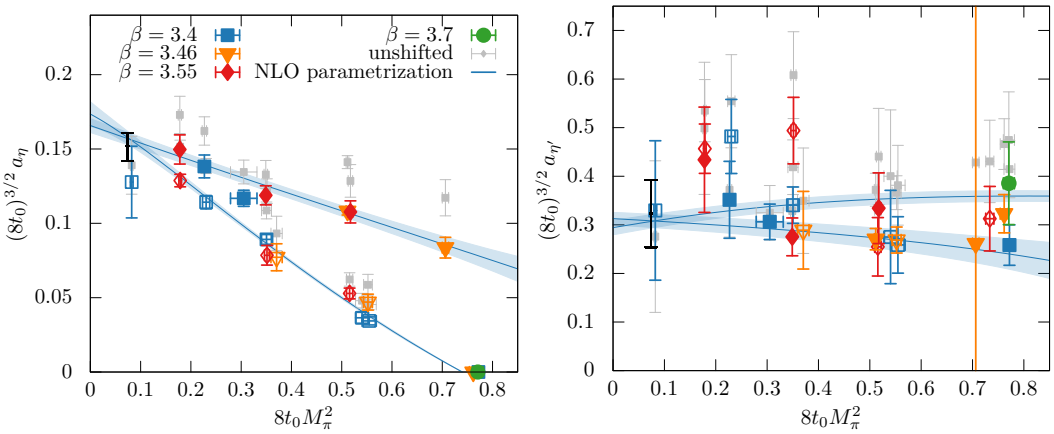


FIGURE 6.3: The anomalous gluonic matrix element a_η (left) and $a_{\eta'}$ (right) determined via the singlet AWI from fermionic matrix elements, eq. (6.12). The coloured points have been adjusted for lattice spacing effects, while the grey points indicate the unshifted data. The two curves correspond to the NLO large-N_c ChPT parametrization eqs. (F.9) and (F.10) for trajectories with a constant average quark mass and a constant strange quark mass. The black error bars indicate the final results at the physical point including statistical and systematic errors.

and decay constants, and $\tilde{\Lambda}$ moves up accordingly. Also L_5 is larger by about $1.3\,\sigma$. This indicates some tension between the data and the NLO expressions.

At the physical point and $\mu = \infty$, the fit gives

$$(8t_0^{\text{ph}})^{3/2} a_\eta = 0.1564\,(^{37}_{63}) \quad \text{and} \quad (8t_0^{\text{ph}})^{3/2} a_{\eta'} = 0.308\,(^{16}_{17})\,. \tag{6.18}$$

The NLO prediction eqs. (F.9) and (F.10), using the LECs of eq. (5.38), reads

$$(8t_0^{\text{ph}})^{3/2} a_\eta = 0.1609\,(^{17}_{27}) \quad \text{and} \quad (8t_0^{\text{ph}})^{3/2} a_{\eta'} = 0.383\,(^{11}_{17})\,. \tag{6.19}$$

Note that the latter values are based exclusively on the meson masses and their decay constants, with no input from the data on $a_{\eta^{(\prime)}}$. The predictions and fit results are close to each other. However, within the relatively small errors stated, the two results on $a_{\eta'}$ differ by several standard deviations, which indicates the limitations of the NLO continuum parametrization within our range of quark masses. Therefore, we assign the difference between eqs. (6.18) and (6.19) as the systematic error associated with taking the physical limit. We discuss the results and quote values in physical units in sec. 6.6.

6.4 Renormalization of the anomaly term and the topological susceptibility

After the indirect determination of the gluonic matrix elements via pseudoscalar and axialvector matrix elements, we now determine them directly, destroying $\eta^{(\prime)}$ states by the topological charge density operator. This allows us to validate the singlet AWI in our lattice study, also in view of confirming a consistent continuum limit extrapolation of the lattice data. However, this requires an analysis of the renormalization of the anomaly term and its mixing with the derivative of the axialvector current.

We compute the topological charge density according to eq. (6.2), using the field theoretical definition of the field strength tensor, eq. (2.63), after flowing the gauge fields to a Wilson flow time [144] of approximately $t/a^2 = t_0^*/a^2$ at each lattice spacing, cf. also sec. 3.4. This removes possible mixing with operators like $a^{-1}P^0$ and regularizes contributions from contact terms like

$\langle\omega(0)\omega(0)\rangle$. From this flowed charge density we can compute correlators just as with ordinary loops, cf. sec. 4.2.1 and the matrix elements can be computed according to eqs. (4.31) and (4.32), setting $J = \omega$. Again for ensembles with open boundary conditions we keep the same distance $b \gtrsim 1.9\,\mathrm{fm}$ to the boundaries as we did in the computation of the fermion loops, cf. sec. 4.1 and tab. 4.1.

To investigate the renormalization of ω, we start from the singlet AWI in the massless case

$$\partial_\mu \widehat{A}^0_\mu = \sqrt{2N_f}\,\widehat{\omega}. \tag{6.20}$$

Since $\omega(x)$ can be written as the divergence of a topological current, the associated Pontryagin index

$$Q = \int \mathrm{d}^4x\, \omega(x) \tag{6.21}$$

is integer-valued on $\mathbb{R}^4$ in the continuum limit and scale independent such that $\omega(x)$ itself will not acquire an anomalous dimension. $\partial_\mu A^0_\mu$ can and will mix into ω:

$$\widehat{\omega} = Z_\omega \omega + Z_{\omega A}\partial_\mu A^0_\mu, \tag{6.22}$$

up to gradient flow time dependent $\mathcal{O}(a)$ corrections. We remark that the anomalous dimensions of A^0_μ and of ω differ from each other in lattice regularization as well as in naive dimensional regularization. The singlet AWI holds exactly when defining the topological charge density using overlap fermions [211], without any factor Z_ω and the term containing $Z_{\omega A}$ cancels when computing the topological susceptibility τ, defined in eq. (6.23), with periodic boundary conditions. Since the topological susceptibility obtained from employing the overlap definition and the field theoretical definition after cooling (which is equivalent to the gradient flow) appear to agree in the continuum limit [212], it is likely that actually $Z_\omega = 1$.

Note that the running of $Z_{\omega A}$ with the scale is the same as that of Z^s_A, which is consistent with eq. (6.20).[1] An alternative scheme of renormalizing the singlet axialvector current is discussed in sec. 2.4.3. In that case, both Z^s_A and $Z_{\omega A}$ have no anomalous dimension. We remark that the $Z_{\omega A}\partial_\mu A^0_\mu$ term will not affect the topological susceptibility

$$\hat{\tau} = \sum_x \langle\widehat{\omega}(0)\widehat{\omega}(x)\rangle = \frac{1}{V}\sum_{x,y}\langle\widehat{\omega}(x)\widehat{\omega}(y)\rangle = \frac{\langle\widehat{Q}^2\rangle}{V} \tag{6.23}$$

since this term does not contribute to the volume sum, due to translational invariance. In the bulk of the lattice this also holds for open boundary conditions in time, provided $L_t - 2b$ is much larger than the relevant correlation lengths:

$$\sum_{x_0=-L_t/2+b}^{L_t/2-b}\sum_{\vec{x}}\partial_\mu A^0_\mu(x) = \frac{1}{2a}\sum_{\vec{x}}\left[A^0_0(L_t/2-b) - A^0_0(-L_t/2+b)\right] \xrightarrow{L_t\to\infty} 0. \tag{6.24}$$

As a first sanity check, we plot the topological susceptibility in fig. 6.4, where we include most of the CLS ensembles analysed in [90], which adds additional points at finer and coarser lattice

[1] In [124] somewhat different conventions are used that correspond to $4\pi^2 a_s G\widetilde{G} = F^a_{\mu\nu}F^a_{\mu\nu}$, where $a_s = \alpha_s/\pi = g_s^2/(4\pi^2)$. Therefore, in that case the γ-function for $G\widetilde{G}$ reads $-\beta(a_s)/a_s$, while our Z_ω does not carry any anomalous dimension. Likewise, in that article the anomalous dimension of the off-diagonal element is proportional to γ^s_A/a_s, while here $\gamma_{\omega A} = \gamma^s_A$.

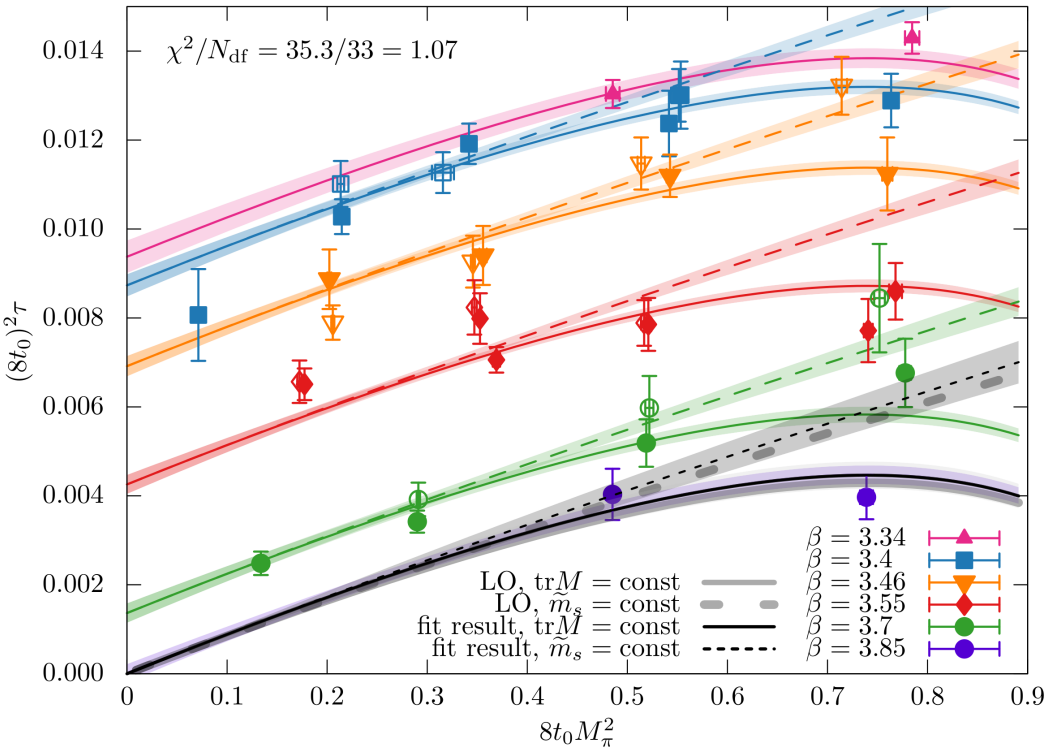


FIGURE 6.4: Topological susceptibility for many of the CLS ensembles described in [90]. Filled symbols mark ensembles that are simulated with a constant sum of quark masses (solid lines), open symbols correspond to ensembles with the strange quark mass fixed to approximately the physical value (dashed lines). Lines and shaded regions are the result of a fit to eq. (6.26). The continuum limit result (black lines) is very close to both the fit result at $\beta = 3.85$ as well as the leading order expectation (grey lines), when using $\sqrt{8t_0}F = 0.1866$, see eq. (5.38), and setting $Z_\omega = 1$.

spacings. We find large cut-off effects with our definition of the susceptibility, shifting points considerably away from the $N_f = 3$ continuum expectation [29, 213],

$$\hat{\tau} = \frac{F^2}{2}\left(\frac{1}{2M_K^2 - M_\pi^2} + \frac{2}{M_\pi^2}\right)^{-1}. \tag{6.25}$$

Indeed, large cut-off effects have been reported in unquenched simulations previously [214–220]. To confirm $Z_\omega = 1$ numerically, we attempt a simple fit to

$$(8t_0)^2\tau = \frac{(8t_0)^2F^2}{2Z_\omega^2}\left(\frac{1}{2M_K^2 - M_\pi^2} + \frac{2}{M_\pi^2}\right)^{-1} + l_\tau^{(2)}\frac{a^2}{t_0^*} + l_\tau^{(3)}\frac{a^3}{(t_0^*)^{3/2}} + l_\tau^{(4)}\frac{a^4}{(t_0^*)^2}. \tag{6.26}$$

From this four parameter fit with $\chi^2/N_{\rm df} \approx 35.3/33 \approx 1.07$, we obtain in the continuum limit

$$\frac{\sqrt{8t_0^\chi}F}{Z_\omega} = 0.190(13). \tag{6.27}$$

When assuming $Z_\omega = 1$, this value agrees with our previous result $\sqrt{8t_0^\chi}F = 0.1866(48)$ (see eq. (5.38) of sec. 5.4.1). The coefficients of the terms parameterizing the lattice spacing dependence are

$$l_\tau^{(2)} = -0.072(10), \quad l_\tau^{(3)} = 0.355(34) \quad \text{and} \quad l_\tau^{(4)} = -0.324(30), \tag{6.28}$$

resulting in the non-monotonous behaviour observed in fig. 6.4. The alternating sign also explains how the susceptibilities at our finest lattice spacing $a \approx 0.039\,\text{fm}$ ($\beta = 3.85$) can agree with the continuum limit expectation. We also tried to add mass-dependent terms to our parametrization of lattice artefacts, however, the resulting coefficients turned out to be small and the quality of the fit did not improve. Equation (6.26) with four parameters turned out to be the minimal ansatz that resulted in a valid description of all our 37 data points. Interestingly, the leading order continuum limit expectation for the dependence of $\hat{\tau}$ on the pion and kaon masses already gives a very adequate description of the data.

6.5 Direct determination of the gluonic matrix elements

We wish to check if the fermionic results that were obtained in sec. 6.2 from employing the singlet AWI are consistent with a direct determination of the gluonic matrix elements. The renormalized matrix elements are given as

$$a_{\mathcal{M}}(\mu) = 2\,Z_\omega\langle\Omega|\omega|\mathcal{M}\rangle + 2\frac{Z_{\omega A}}{Z_A^s}M_{\mathcal{M}}^2F_{\mathcal{M}}^0(\mu), \tag{6.29}$$

see eq. (6.22). In the previous section, we have found $Z_\omega = 1$ from a fit to the topological susceptibility. As an additional cross check, we also simultaneously solve the above equation for $\mathcal{M} = \eta$ and $\mathcal{M} = \eta'$ to obtain $Z_{\omega A}$ and Z_ω. We plot the resulting values for Z_ω in the left panel of fig. 6.5. Qualitatively these are in agreement with $Z_\omega = 1$ and we suspect that the two outliers are due to lattice artefacts.

Based on the evidence presented above, we assume $Z_\omega = 1$, however, $Z_{\omega A}$ is not known and therefore comparing the direct determination (6.29) of the anomaly terms with the corresponding predictions from the singlet AWI eq. (6.11) cannot be entirely independent. Fortunately, the ratio $Z_{\omega A}/Z_A^s$ only depends on the inverse lattice coupling, β, but not on the pion and kaon masses. Moreover, the renormalization is independent of the meson. Rearranging eq. (6.29), we can isolate

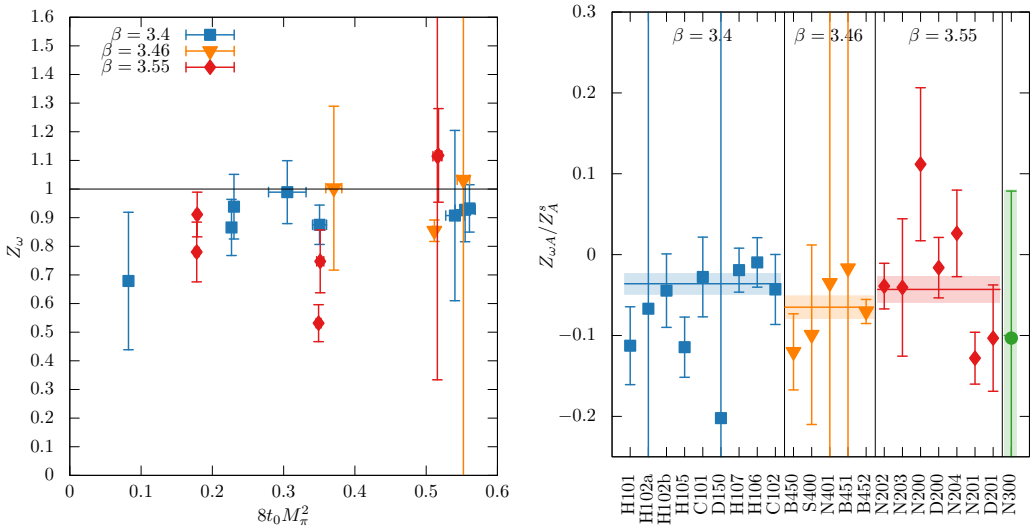


FIGURE 6.5: **Left panel**: Z_ω from solving eq. (6.29). The $m_s = m_\ell$ points are not shown since in these cases $F_\eta^0 = 0$ and the equation system is singular. **Right panel**: Values of $Z_{\omega A}/Z_A^s$ from eq. (6.30), assuming $Z_\omega = 1$.

the renormalization scale independent ratio

$$\frac{Z_{\omega A}}{Z_A^s} = \frac{a_{\mathcal{M}} - 2\, Z_\omega \langle \Omega | \omega | \mathcal{M} \rangle}{2 M_{\mathcal{M}}^2 F_{\mathcal{M}}^0}. \tag{6.30}$$

Since F_η^0 in the denominator is close to zero and has large relative errors, we only use the η' matrix elements for which $F_{\eta'}^0 \gg 0$ on all our ensembles. We plot this ratio for $Z_\omega = 1$ in the right panel of fig. 6.5. Indeed, the data for each β-value are compatible with a constant. Taking a weighted average over all points at each of our four lattice spacing, we obtain

$$\begin{aligned} (Z_{\omega_A}/Z_A^s)\,(\beta = 3.4) &= -0.036(13), \quad & (Z_{\omega_A}/Z_A^s)\,(\beta = 3.46) &= -0.065(14), \\ (Z_{\omega_A}/Z_A^s)\,(\beta = 3.55) &= -0.043(16), \quad & (Z_{\omega_A}/Z_A^s)\,(\beta = 3.7) &= -0.10(18). \end{aligned} \tag{6.31}$$

Using these values (and $Z_\omega = 1$), we evaluate eq. (6.29) with the anomalous matrix elements computed at the gradient flow time $t \approx t_0^*$ on the individual ensembles. We compare our results on every ensemble with the fermionic determination of sec. 6.2 in the scatter plot fig. 6.6. Our gluonic results agree qualitatively with the fermionic determination. The mixing with the axialvector current is non-negligible, i.e. $Z_{\omega A} \neq 0$. Had we ignored this mixing, the gluonic determinations would have undershot the fermionic ones by roughly 30 % both for the η and the η'. We stress that agreement can only be expected in the continuum limit since both definitions are subject to different discretization effects. We have observed considerable lattice spacing effects both for the topological susceptibility in sec. 6.4 and the singlet pseudoscalar matrix elements in sec. 6.2. The qualitative agreement suggests that some of the discretization effects may be similar for both definitions.

6.6 Pseudoscalar gluonic and fermionic matrix elements

We determined the anomaly matrix elements a_η and $a_{\eta'}$ in sec. 6.2 from a fit to combinations of axialvector and pseudoscalar matrix elements, eq. (6.11). The fit is performed for data at the QCD renormalization scale $\mu = \infty$, and we carry out the conversion to lower scales, using the fact that the combinations $m_f H_n^f$ are scale independent. We first determine these combinations by plugging

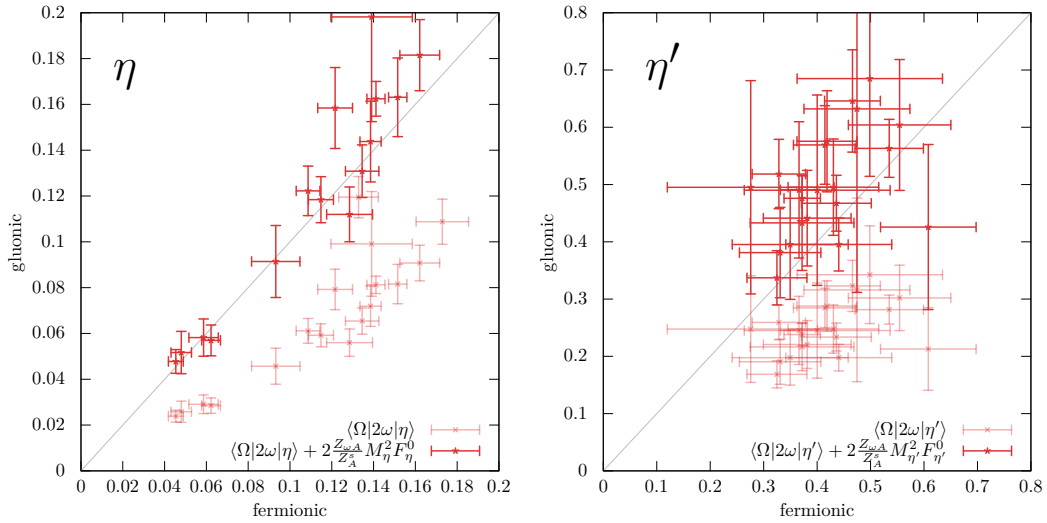


FIGURE 6.6: Scatter plot of the fermionic (eq. (6.12), horizontally) and gluonic (eq. (6.29), vertically) determinations of the gluonic matrix elements $a_{\eta^{(\prime)}}$. The mixing with the derivative of the axialvector current is non-negligible: the unrenormalized lattice matrix elements (pale red points) do not agree with the fermionic definition.

	a_η	$a_{\eta'}$
$\mu = 1\,\text{GeV}$	$0.01720\,(^{40}_{69})_{\text{stat}}\,(48)_{\text{syst}}\,(67)_{t_0}\ \text{GeV}^3$	$0.0424\,(^{19}_{17})_{\text{stat}}\,(80)_{\text{syst}}\,(19)_{t_0}\ \text{GeV}^3$
$\mu = 2\,\text{GeV}$	$0.01700\,(^{40}_{69})_{\text{stat}}\,(48)_{\text{syst}}\,(66)_{t_0}\ \text{GeV}^3$	$0.0381\,(^{18}_{17})_{\text{stat}}\,(80)_{\text{syst}}\,(17)_{t_0}\ \text{GeV}^3$
$\mu = 10\,\text{GeV}$	$0.01688\,(^{40}_{69})_{\text{stat}}\,(48)_{\text{syst}}\,(66)_{t_0}\ \text{GeV}^3$	$0.0356\,(^{18}_{17})_{\text{stat}}\,(80)_{\text{syst}}\,(17)_{t_0}\ \text{GeV}^3$
$\mu = \infty$	$0.01676\,(^{40}_{67})_{\text{stat}}\,(48)_{\text{syst}}\,(65)_{t_0}\ \text{GeV}^3$	$0.0330\,(^{18}_{17})_{\text{stat}}\,(80)_{\text{syst}}\,(16)_{t_0}\ \text{GeV}^3$
	θ_y	$a_{\eta'}^2/a_\eta^2$
$\mu = 1\,\text{GeV}$	$-22.1\,(^{3}_{5})_{\text{stat}}\,(2.8)^\circ_{\text{syst}}$	$6.09\,(^{27}_{53})_{\text{stat}}\,(2.05)_{\text{syst}}$
$\mu = 2\,\text{GeV}$	$-24.0\,(^{4}_{1.0})_{\text{stat}}\,(3.2)^\circ_{\text{syst}}$	$5.03\,(^{19}_{45})_{\text{stat}}\,(1.94)_{\text{syst}}$
$\mu = 10\,\text{GeV}$	$-25.3\,(^{4}_{1.1})_{\text{stat}}\,(3.6)^\circ_{\text{syst}}$	$4.46\,(^{16}_{41})_{\text{stat}}\,(1.86)_{\text{syst}}$
$\mu = \infty$	$-26.9\,(^{4}_{1.2})_{\text{stat}}\,(4.1)^\circ_{\text{syst}}$	$3.88\,(^{14}_{38})_{\text{stat}}\,(1.78)_{\text{syst}}$

TABLE 6.1: Gluonic matrix elements of the η and η' and combinations thereof at various scales.

our physical point results on the masses, decay constants and the gluonic matrix elements into the AWIs in the flavour basis, eq. (6.5). Following this, we reconstruct $a_{\eta^{(\prime)}}$ at different scales using the known running of the singlet axialvector current. With $N_f = 3$ active quark flavours, at 2 GeV we obtain:

$$a_\eta(\mu = 2\,\text{GeV}) = 0.01700\,(^{40}_{69})_{\text{stat}}\,(48)_{\text{syst}}\,(66)_{t_0}\ \text{GeV}^3, \tag{6.32}$$

$$a_{\eta'}(\mu = 2\,\text{GeV}) = 0.0381\,(^{18}_{17})_{\text{stat}}\,(80)_{\text{syst}}\,(17)_{t_0}\ \text{GeV}^3. \tag{6.33}$$

The systematic error is computed as the difference between our results from a direct NLO fit to the $a_{\eta^{(\prime)}}$ data (see eq. (6.18)), that included lattice correction terms, and the continuum NLO large-N_c ChPT prediction (see eq. (6.19)), based on the set of LECs that we obtained from our simultaneous fits to the masses and decay constants. We list our results at various scales in tab. 6.1 and compare to literature values in tab. 6.2, where the scale is not specified. These analyses are based on, e.g., QCD sum rule calculations [221, 222], large-N_c ChPT [184, 223] and related state mixing models that include a pseudoscalar glueball [224, 225]. We find agreement with the references that give error estimates, with the exception of [223].

Combining our physical point results on $a_{\eta^{(\prime)}}$, the $\eta^{(\prime)}$ masses and their decay constants with eq. (6.5) gives the following predictions

$$\begin{aligned}
m_\ell H^\ell_\eta &= \quad 0.0021\,(^{3}_{2})_{\text{stat}}\,(13)_{\text{syst}}(0)_{t_0}\,\text{GeV}^3,\\
m_s H^s_\eta &= -0.0173\,(^{3}_{2})_{\text{stat}}\,(17)_{\text{syst}}(7)_{t_0}\,\text{GeV}^3,\\
m_\ell H^\ell_{\eta'} &= \quad 0.0045\,(^{10}_{8})_{\text{stat}}\,(40)_{\text{syst}}(0)_{t_0}\,\text{GeV}^3,\\
m_s H^s_{\eta'} &= \quad 0.0309\,(^{15}_{5})_{\text{stat}}\,(50)_{\text{syst}}(10)_{t_0}\,\text{GeV}^3
\end{aligned} \tag{6.34}$$

for the pseudoscalar fermionic matrix elements, where again the systematic error is the difference with respect to the NLO ChPT predictions eqs. (F.5)–(F.8), obtained using our set of LECs. Since the values of the above combinations are smaller in the light quark sector than for strange quarks, and the absolute error on $a_{\eta^{(\prime)}}$ is the major contribution to their uncertainty, the relative precision that we can achieve is limited for the light quark combinations. Note that this is a statement about the physical mass continuum limit; on individual ensembles also the light quark matrix elements can be quite precise. While there is some tension for the combination $m_s H^s_\eta$, most of our results agree with the estimate in the FKS approximation, where the pseudoscalar matrix element is taken in the SU(2) isospin limit,

$$\begin{aligned}
\text{Feldmann [183]:}\quad m_\ell H^\ell_\eta &= 0.0010\,\text{GeV}^3, & m_s H^s_\eta &= -0.026\,\text{GeV}^3,\\
m_\ell H^\ell_{\eta'} &= 0.0008\,\text{GeV}^3, & m_s H^s_{\eta'} &= \quad 0.032\,\text{GeV}^3,
\end{aligned} \tag{6.35}$$

and the very similar numbers of a QCD sum rule calculation,

$$\begin{aligned}
\text{Singh [222]:}\quad m_\ell H^\ell_\eta &= 0.00105(14)\,\text{GeV}^3, & m_s H^s_\eta &= -0.0284(55)\,\text{GeV}^3,\\
m_\ell H^\ell_{\eta'} &= 0.000782(250)\,\text{GeV}^3, & m_s H^s_{\eta'} &= \quad 0.0379(71)\,\text{GeV}^3.
\end{aligned} \tag{6.36}$$

We again emphasize that the above combinations are renormalization group invariants.

It is particularly interesting to inspect the ratio of the gluonic matrix elements that can be used to define a mixing angle in the gluonic sector [184],

$$\theta_y(\mu = 2\,\text{GeV}) = -\arctan\left(\frac{a_\eta(2\,\text{GeV})}{a_{\eta'}(2\,\text{GeV})}\right) = -24.0\,(^{4}_{1.0})_{\text{stat}}\,(3.2)^\circ_{\text{syst}}\,. \tag{6.37}$$

ref	a_η/GeV^3	$a_{\eta'}/\mathrm{GeV}^3$
Novikov et al. [221]	0.021	0.035
Feldmann [184]	0.023	0.058
Beneke and Neubert [223]	0.022(2)	0.057(2)
Cheng et al. [224]	0.026(28)	0.054(57)
Singh [222]	0.0220(50)	0.037(10)
Qin et al. [225]	0.016	0.051
Ding et al. [193]	0.024	0.051
this work at $\mu = 1\mathrm{GeV}$	0.0172(10)	0.0424(84)
this work at $\mu = 2\mathrm{GeV}$	0.0170(10)	0.0381(84)
this work at $\mu = \infty$	0.0168(10)	0.0330(83)

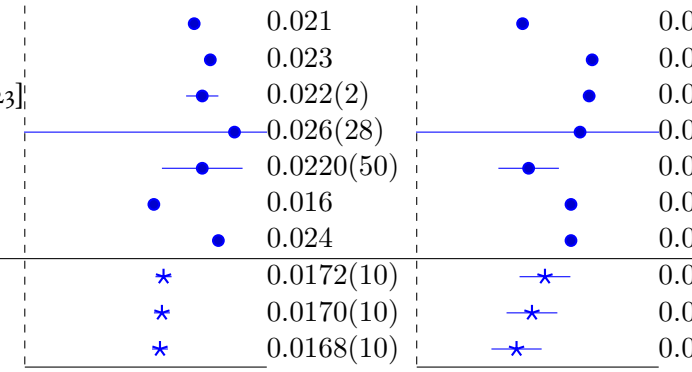


TABLE 6.2: Literature values for the anomaly matrix elements in comparison with our results at various scales. Note that the error bars of [224] are cut off at both ends.

The squared ratio $(a_{\eta'}/a_\eta)^2 = (\cot\theta_y)^2$ is closely related to the ratio of decay widths of $J/\psi \to \eta^{(\prime)}\gamma$ when assuming that the anomaly dominates [221, 226],

$$R(J/\psi) = \frac{\Gamma[J/\psi \to \eta'\gamma]}{\Gamma[J/\psi \to \eta\gamma]} = \frac{a_{\eta'}^2}{a_\eta^2}\left(\frac{k_{\eta'}}{k_\eta}\right)^3, \tag{6.38}$$

where $k_{\mathcal{M}} = \frac{1}{2}\frac{M_{J/\psi}^2 - M_{\mathcal{M}}^2}{2\,M_{J/\psi}}$ is the momentum of the meson $\mathcal{M}$ in the rest frame of the J/ψ. Using the experimental masses of [19], we obtain $(k_{\eta'}/k_\eta)^3 \approx 0.8137$. Our result for the ratio $a_{\eta'}^2/a_\eta^2$, listed in tab. 6.1, gives at $\mu = 2\,\mathrm{GeV}$

$$R(J/\psi, \mu = 2\,\mathrm{GeV}) = 5.03\,\left(\tfrac{19}{45}\right)_{\mathrm{stat}}\,(1.94)_{\mathrm{syst}}\,. \tag{6.39}$$

Note that $a_{\eta'}$ depends strongly on the scale, see. tab. 6.1, and the most relevant scale for this decay is probably below 2 GeV, which would somewhat increase the prediction. The most recent PDG averages for the partial widths $\Gamma(J/\psi \to \eta\gamma)/\Gamma_{\mathrm{total}} = 1.108(27)\cdot 10^{-3}$ and $\Gamma(J/\psi \to \eta'\gamma)/\Gamma_{\mathrm{total}} = 5.25(7)\cdot 10^{-3}$ result in

$$\text{PDG [19]:} \quad R(J/\psi) = 4.74(13), \tag{6.40}$$

which is very close to our value with $N_f = 3$ at $\mu = 2\,\mathrm{GeV}$. Clearly, a more precise comparison should take $\mathcal{O}(\alpha_s)$ corrections into account.

7 Conclusions

In this thesis we determined important properties of the η and η' mesons, in particular their masses, decay constants and anomalous matrix elements. The latter two determinations are the first from lattice QCD and also the first without making model assumptions. This demonstrates the unique position of lattice QCD as the principal tool to study the deeply non-perturbative regime of QCD, dynamically reproducing aspects such as the axial anomaly that gives the η' a large fraction of its mass. In our calculation the total precision that was achieved for all these observables is in many cases higher than in previous — lattice and phenomenological — studies.

The reason for this is two-fold: First, we applied efficient noise reduction techniques and developed improved analysis methods that rely on fitting the data simultaneously, fully taking into account the correlations and the different noise levels of the entries of the matrix of correlation functions. A new matrix generalization of the effective mass method improves the stability of the fits.

Second, we take advantage of a large subset of the plethora of CLS ensembles that were generated during the course of this thesis and for which we set the relative scale in sec. 3.4. These ensembles employ $N_f = 2 + 1$ flavours of non-perturbatively Sheikholeslami-Wohlert improved Wilson fermions and are generated with a tree-level improved gauge action. The four lattice spacings analysed here cover the range $0.050\,\text{fm} \lesssim a \lesssim 0.086\text{fm}$. Quark masses are tuned to follow two distinct mass trajectories that start at $M_\pi \approx 422\,\text{MeV}$ and both lead to the physical point. Five of the analysed ensembles are at particularly small pion masses $M_\pi < 250\,\text{MeV}$ of which one is very close to the physical mass point.

These features put us in a position to reduce and reliably quantify systematic errors: The effect of the finite lattice spacing can be extrapolated away including known linear improvement coefficients and parametrizing unknown and higher order effects in a simultaneous fit to the decay constants and masses employing four different lattice spacings. By varying the parametrizations of these, the uncertainty related to the discretization could be estimated with good confidence.

The chiral extrapolation, connecting the ensembles at unphysical quark masses to the physical point, employs next-to-leading order large-N_c ChPT parametrizations that allow us to combine the two η and η' masses and four decay constants along both mass trajectories in a single simultaneous fit that in the continuum is parametrized by just six low energy constants. This fit is tightly constrained, and the precision of the inherently noisy mass of the η', for example, benefits from this global fit approach. The deviations from this fit — the value of χ^2/N_{df} — indicate only slight truncation effects of the continuum parametrization and we quantified the systematic error associated to this by successively removing points at large average quark masses. This reduces the χ^2/N_{df}; however, central values have changed only slightly, and the partial inclusion of NNLO contributions did not improve the fit.

The immediate result of the fits is a set of the six low energy constants of large-N_c ChPT, F, L_5, L_8, M_0, Λ_1 and Λ_2. We list their values in eq. (5.38) and discuss them in more detail in sec. 5.5.3. For some of these only very rough estimates existed previously, and for the first time the QCD renormalization scale dependence of Λ_1, Λ_2 and M_0 could be taken into account. This allowed us to test model assumptions as that of the FKS scheme which could be verified to hold at low energies.

These fit results for the low energy constants in the continuum yield our physical point values for the masses,

$$M_\eta = 554.7(9.2)\,\text{MeV} \quad \text{and} \quad M_{\eta'} = 930(21)\,\text{MeV}, \tag{7.1}$$

where the errors were added in quadrature. These values determined with $N_f = 2 + 1$ flavours are in agreement with experiment within about one standard error and in line with previous lattice results that have larger uncertainties, see sec. 5.5.1 and fig. 5.3.

Our results for the octet/singlet basis decay constants in the $\overline{\text{MS}}$ scheme at $\mu = 2\,\text{GeV}$ and in the angle representation read

$$\begin{aligned} F^8 &= 115.0(2.8)\,\text{MeV}, & \theta_8 &= -25.8(2.3)^\circ, \\ F^0(\mu = 2\,\text{GeV}) &= 100.1(3.0)\,\text{MeV}, & \theta_0 &= -8.1(1.8)^\circ. \end{aligned} \tag{7.2}$$

The same results in various other representations are given in tab. E.2. These are the first determinations of these couplings directly from the axialvector matrix elements and without resorting to model assumptions. Using a modified renormalization scheme we take the scale dependence of the singlet axialvector current in the $\overline{\text{MS}}$ scheme fully into account and can compare to previous phenomenological determinations that had to assume a scale to connect to physical processes, see sec. 5.5.2. These results are important input for further theory predictions and future experiments like Belle II [26, 27] and we demonstrated their use in the calculation of the transition form factor $\gamma\gamma^* \to \eta^{(\prime)}$ in sec. 5.5.4.

Finally, in ch. 6 we computed the anomalous matrix elements that contribute to the singlet axial Ward identity from first principles,

$$a_\eta(\mu = 2\,\text{GeV}) = 0.0170(10)\,\text{GeV}^3 \quad \text{and} \quad a_{\eta'}(\mu = 2\,\text{GeV}) = 0.0381(84)\,\text{GeV}^3, \tag{7.3}$$

where the gluonic mixing angle is given by

$$\theta_y(\mu = 2\,\text{GeV}) = -\arctan\left(\frac{a_\eta}{a_{\eta'}}\right) = -24.0(3.3)^\circ. \tag{7.4}$$

At the scale $\mu = 2\,\text{GeV}$ the latter is in good agreement with the octet angle, see eq. (7.2), again confirming the validity of the FKS state mixing model [183, 196]. The matrix elements $a_{\eta^{(\prime)}}$ are the main couplings parametrizing the anomalous strong decay $J/\psi \to \eta^{(\prime)}$ and we see agreement with experimental decay rates, see sec. 6.6.

Our determinations of the masses and decay constants benefit from the comparably simple parametrization by NLO ChPT, however, some tension with the data is indicated by the χ^2/N_{df} of the fits that reduces when removing heavy points for which truncation errors may be significant. Also, the prediction of the anomalous matrix elements using the LECs from the fit to masses and decay constants gives results that are numerically relatively close to the fit to the gluonic matrix elements themselves, but disagree given their errors. Both these observations are indications that higher orders in ChPT may have to be included. Adding NNLO mesonic loop terms did, however, not improve the situation and a full NNLO fit would introduce many more LECs that with the currently available data we are unable to fix. The inclusion of decay constants and masses of the full meson-nonet may help in that respect and provide enough information to determine all LECs of NNLO large-N_c ChPT. Also, more points along the $m_s = m_\ell$ line may help to pin down some parameters — and in particular those that exclusively appear in the singlet channels — more reliably since there the singlet and the octet channels decouple. Naturally, the inclusion of even more ensembles at the other trajectories and at smaller lattice spacings and light quark masses would also be beneficial in that respect. Given the good precision already at NLO that we observe, this may provide important input to existing NNLO large-N_c calculations [181, 182, 194, 227] to enable

predictions with high precision.

For some of our results, the overall scale uncertainty amounts for a large fraction of the total error. This uncertainty will likely reduce in the future when new determinations of t_0^{ph} in physical units become available. In case of the determined values for the masses, eq. (7.1), it will be interesting to see if the central values move closer to the η (when the value of t_0^{ph} is reduced) or closer to the η'. In both cases the disagreement with one of the results will likely increase. We simulate $N_f = 2 + 1$ QCD and while electromagnetic contributions presumably do not play a role for the neutral η and η', it is possible that the inclusion of a charm quark and isospin breaking effects may become significant. The contribution of the charm quark has been shown to be small [50] but in the future this may need to be reassessed in view of the reduced errors.

A Conventions

Here we briefly lay out the conventions that we use throughout the text. In general repeated indices are summed over. Greek letters usually label space-time indices, roman letters $a, b, c, \ldots$ are used for colour (both in fundamental and adjoint representation) and $i, j, k, \ldots$ are used for more general, ordinal (multi-) indices.

A.1 Natural units

As usual in this field of physics, natural units are used, i.e., we set the speed of light and the reduced Planck constant to one,

$$c = \hbar = 1. \tag{A.1}$$

We can relate energies and lengths by

$$c\hbar = 197.3\,\text{MeV fm} = 1. \tag{A.2}$$

A.2 Dirac gamma matrices

We use the following four dimensional chiral basis in Euclidean time that is used in the software packages `QDP++` and `chroma` [228]:

$$\gamma_t = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \qquad \gamma_x = \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix},$$
$$\gamma_y = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \qquad \gamma_z = \begin{pmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & -i \\ -i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix}. \tag{A.3}$$

These obey the usual anti-commutation relations in Euclidean time,

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}\mathbb{1}. \tag{A.4}$$

The Minkowski metric can be recovered by multiplying the spatial matrices with $-i = 1/i$,

$$\gamma_x \mapsto \gamma_x^M = -i\gamma_x, \quad \gamma_y \mapsto \gamma_y^M = -i\gamma_y, \quad \gamma_z \mapsto \gamma_z^M = -i\gamma_z, \tag{A.5}$$

which leads to

$$\{\gamma_\mu^M, \gamma_\nu^M\} = 2g_{\mu\nu}\mathbb{1}, \tag{A.6}$$

where $g_{\mu\nu}$ is the metric tensor with signature $(1,-1,-1,-1)$. The fifth matrix is defined as the product of the other four,

$$\gamma_5 = \gamma_x\gamma_y\gamma_z\gamma_t = \begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & -\mathbb{1}_2 \end{pmatrix}. \tag{A.7}$$

In Minkowski time the definition is slightly different, but leads to the same matrix

$$\gamma_5^M = i\gamma_t\gamma_x^M\gamma_y^M\gamma_z^M = \gamma_5. \tag{A.8}$$

With this, the chiral projectors of eq. (2.23) are given as

$$\mathcal{P}_L = \frac{1}{2}\left(\mathbb{1} - \gamma_5\right) = \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{1}_2 \end{pmatrix}, \tag{A.9}$$

$$\mathcal{P}_R = \frac{1}{2}\left(\mathbb{1} + \gamma_5\right) = \begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & 0 \end{pmatrix}. \tag{A.10}$$

A.3 Gell-Mann matrices

We use the following set of hermitian, traceless 3×3 matrices that span the $\mathfrak{su}(3)$ algebra,

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix},$$

$$\lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda^8 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{A.11}$$

These Gell-Mann matrices obey the commutation relation

$$\left[\lambda^a, \lambda^b\right] = 2if^{abc}\lambda^c, \tag{A.12}$$

where the antisymmetric structure constants are given by

$$f^{abc} = -\frac{1}{4}i \ \mathrm{tr}\left(\lambda^a[\lambda^b, \lambda^c]\right). \tag{A.13}$$

We typically use $t^a = \lambda^a/2$ as generators of SU(3) so that the normalization is $\mathrm{tr}\, t^a t^b = \delta^{ab}/2$. We also define $t^0 = \mathbb{1}/\sqrt{6}$ to complete the set of generators for U(3).

B Statistical analysis and fitting

On one ensemble e we obtain a sequence of n measurements for a set of observables $\{y\}_i$ (correlators, for example) for every configuration i separated in Monte-Carlo time,

$$e = (e_0, e_{n-1}), \qquad e_i = (\{y\}_i, w_{\ell,i}, w_{s,i}) = (\{y\}_i, w_i), \tag{B.1}$$

with associated light and strange reweighting factors $w_{\ell,i}$ and $w_{s,i}$, respectively. The latter we combine immediately to a joint reweighting factor for that configuration, $w_i = w_{\ell,i} w_{s,i}$. See sec. 2.2.4 for the origin of and the reason for the reweighting factors.

Averages of a particular observable y over a set of indices $\mathcal{I}$ can be computed by taking the weighted mean,

$$\langle y \rangle_{\mathcal{I}} = \frac{\sum_{i\in\mathcal{I}} w_i y_i}{\sum_{i\in\mathcal{I}} w_i}. \tag{B.2}$$

We denote the mean over the full set $[0, n-1]$ of configurations as $\overline{y} = \langle y \rangle_{[0,n-1]}$.

An estimate for the sample variance is given by

$$\mathrm{var}(y) = \frac{1}{n-1} \sum_{i=0}^{n-1} \left(\frac{w_i y_i}{\sum_{j=0}^{n-1} w_j} - \overline{y} \right)^2. \tag{B.3}$$

B.1 Binning

Due to the Monte-Carlo process e_i is not independent of e_{i-j} for $j > 0$ and the naive error underestimates the true error.

Therefore, we bin the data into n_{bin} bins that each contains averages over $\Delta_{\mathrm{bin}} = n/n_{\mathrm{bin}}$ of the original measurements (the measurements may be truncated, if n is not divisible by n_{bin}),

$$e_i \mapsto \left(\{ \langle y \rangle_{\mathcal{B}_i} \}, \sum_{j\in\mathcal{B}_i} w_j \right), \tag{B.4}$$

where $\mathcal{B}_i$ is the set of indices in the i-th bin, $\mathcal{B}_i = (i \cdot \Delta_{\mathrm{bin}}, \ldots, (i+1) \cdot \Delta_{\mathrm{bin}} - 1)$.

With increasing values of Δ_{bin}, the error increases and the variance of the binned data approaches $\mathrm{var}_{\Delta_{\mathrm{bin}}} y \to 2\tau_{\mathrm{int}} \mathrm{var}_0 y$, where τ_{int} is the integrated autocorrelation length, var_0 is variance using the unbinned data, eq. (B.3) and $\mathrm{var}_{\Delta_{\mathrm{bin}}}$ the same but after binning with binsize Δ_{bin}. In this way Δ_{bin} is determined as the smallest value that gives the asymptotic value of the error. See tab. 4.1 for the values of Δ_{bin} for the ensembles analysed in this thesis.

B.2 Resampling

To estimate errors of derived quantities correctly, we also resample the data into n_{boot} bootstrap samples. As a rule of thumb, typically $n_{\mathrm{boot}} \gtrsim 2n$ and it is useful (see below) if all ensembles

employ the same number of bootstrap samples. Within our analysis, we use $n_{\text{boot}} = 500$ after binning.

We label the set of usually n random indices in sample i as $\mathcal{S}_i$. The elements are drawn uniformly from $[0, n-1]$ with replacement.

To compute central values, we also keep track of the means of the observables on the ensemble and compute the individual samples in the usual way, see eq. (B.2),

$$e \mapsto (\{\overline{y}\}, \{\tilde{e}_i\}), \qquad \tilde{e}_i = \left(\{\langle y \rangle_{\mathcal{S}_i}\}, \sum_{j \in \mathcal{S}_i} w_j \right) =: (\{\tilde{y}\}_i, \tilde{w}_i), \tag{B.5}$$

where typically but not necessarily resampling is performed on the binned ensembles. We stress that each of these samples is assigned a new reweighting factor to account for the unequal sum of reweighting factors within each sample. Note, however, that $\operatorname{var} \tilde{w}_i \leq \operatorname{var} w_i$ and for typical ensemble sizes $n \gg 1$, $\operatorname{var} \tilde{w}_i \ll \operatorname{var} w_i$. This is not the case for eq. (B.4) where $\Delta_{\text{bin}} = \mathcal{O}(10)$. For resampling, however, the reweighting has mostly been absorbed into the bootstrap samples and the effect of further reweighting of the samples is in practice negligible. In the following we therefore set $\tilde{w}_i = \sum_{j \in \mathcal{S}_i} = 1$ to combine results from several ensembles. In any case the effect of the reweighting of these bootstrap samples will only affect the errors but not the central values, see below.

B.2.1 Combining results

We can compute secondary observables by combining existing results on the mean and each sample,

$$\{\overline{y}'\} = \phi(\{\overline{y}\}), \qquad \{y'\} = \phi(\{\tilde{y}\}_i), \tag{B.6}$$

where ϕ is a suitable function, e.g., it may be a fitting function that determines masses and amplitudes from correlators.

The bootstrap method allows us also to combine multiple ensembles using some other function ψ that takes the observables $\{y\}^e$ of ensemble e and $\{y\}^f$ determined on another ensemble f, etc. as inputs and let's us compute a new bootstrap data set g one-by-one,

$$g = \left(\psi(\{\overline{y}\}^e, \{\overline{y}\}^f, \ldots), \{g_i\} \right), \qquad g_i = (\psi(\{y\}_i^e, \{y\}_i^f, \ldots)), \tag{B.7}$$

where we have omitted the reweighting now with the above argument. The function ψ may, for instance, be a fitting function that estimates joint parameters based on the ensembles e, f, . . ., and it is applied to both the means of the ensembles to return the combined central value as well as to the individual samples to facilitate the computation of the errors. Note that in general the mean over the derived samples is not equal to the function applied to the means, $\psi(\{\overline{y}\}^e, \{\overline{y}\}^f, \ldots) \neq \overline{\psi(\{y\}_i^e, \{y\}_i^f, \ldots)}$, but both are estimators for the expectation value, and working on the central values has the advantage of increased stability and being a better estimator, whereas for the errors the bootstrap method is used.

B.2.2 Determination of errors

Error estimates can be computed using eq. (B.3), $\sigma(y) = \sqrt{\operatorname{var} y}$. Assuming a Gaussian distribution of the samples, these estimate the central 1-σ confidence band, i.e., 68.3 % of all values are expected to lie within a distance of one σ to the mean.

In this thesis we usually give asymmetric confidence intervals, and we compute them in the following way: We sort the bootstrap samples by the magnitude of the observable y we are interested in. We then determine the index corresponding to the lower confidence bound as $i_{\text{low}} =$

$\mathrm{floor}(0.1587 \cdot n_{\mathrm{boot}} + 0.5)$; the upper is defined as $i_{\mathrm{up}} = \mathrm{floor}(0.8413 \cdot n_{\mathrm{boot}} + 0.5)$, where the numerical constants correspond to the cumulative distribution function of a Gaussian normal distribution at $1\,\sigma$ below and above the mean, respectively, i.e., 15.87 % of all points lie outside of below and above the confidence band, each. We quote central values and their asymmetric errors as

$$\overline{y}\begin{pmatrix} y_{i_{\mathrm{up}}} - \overline{y} \\ \overline{y} - y_{i_{\mathrm{low}}} \end{pmatrix}, \tag{B.8}$$

i.e, we do not give a negative sign for the lower bound. Note that $y_{i_{\mathrm{up}}} - y_{i_{\mathrm{low}}} \approx 2\sigma(y)$.

B.3 Fitting multi-variate functions with correlations

We fit by minimizing the quadratic differences of a fit function $\phi(\{x\}|\{\theta\})$ that depends on n arguments $\{x\} = [x_0, \ldots, x_{n-1}]$ that are given by our data and m parameters $\{\theta\} = [\theta_0, \ldots, \theta_{m-1}]$ that we wish to determine from our available data $y(\{x\})$. As an illustrative example, we may consider a simple correlator $C(t)$ that we fit in terms of an amplitude and an energy, $\{\theta\} = \{A, E\}$: $f(t|A, E) = A\exp(-Et)$ to a simple correlator $y(t)$.

The solution is given by the minimum of

$$\{\theta\} = \mathrm{argmin}_\theta \chi^2 = \mathrm{argmin}_\theta \sum_{\{x\}} \frac{\left(y(\{x\}) - \phi(\{x\}|\{\theta\})\right)^2}{\mathrm{var}(\, y(\{x\}))}, \tag{B.9}$$

and we find this minimum numerically using Minuit [229]. We call this a simple uncorrelated fit, and its validity can be evaluated by computing χ^2/N_{df} at the optimal values of θ and the number of degrees of freedom is the difference of the number of data points and the number of parameters that are fitted, $N_{\mathrm{df}} = |\{y\}| - |\{\theta\}|$. For uncorrelated data we expect $\chi^2/N_{\mathrm{df}} \approx 1$. For simultaneous fits to a set of functions that share some of the parameters we simply sum the corresponding χ^2 functions within the minimization.

To impose priors on the fit parameters, we add penalty terms similar to eq. (5.37) to the χ^2 function that increase the value of χ^2 when the parameters move away from their priors.

In practice, there are several problems with this simple approach:

- The data itself may be correlated, both within a data set and across simultaneous fits, i.e., $y(\{x\})$ is not independent of $y(\{x'\})$ and $\mathrm{cov}\,\{y\} \neq \mathrm{diag}(\mathrm{var}\,\{y\})$.
- Often the arguments themselves may be inflicted with errors, $\mathrm{var}\,\{x\} \neq 0$.
- Furthermore, if there is more than one argument, then the arguments are usually even correlated, $\mathrm{cov}(\{x\}) \neq 0$. This is of practical relevance in our case, since we extrapolate in a plane of quark masses that is parametrized by two masses, e.g., the average and the difference of non-singlet pseudoscalar meson masses, cf. eq. (3.2), and both are determined on the same configurations.

We address the first problem by replacing the uncorrelated χ^2 function with

$$\chi^2 = \delta^\intercal \mathrm{cov}^{-1}\{y\}\delta, \tag{B.10}$$

where $\mathrm{cov}(\{y\})$ is the estimated sample covariance matrix among all appearing data in the same ordering as the difference vector $\delta = y(\{x\}) - \phi(\{x\}|\{\theta\})$. If simultaneous fits are performed and the individual fits are uncorrelated (e.g., because they are determined on independent ensembles), then the corresponding blocks in the covariance matrix can be set to zero, and we refer to this situation as blockdiagonal fitting.

The remaining two issues can be addressed by a variant of Orear's method [230], taking the additional uncertainties into account,

$$\chi^2 = \delta^\intercal \left[\mathrm{cov}(\{y\}) + \left(\frac{\partial \phi}{\partial x}\right) \mathrm{cov}\,\{x\} \left(\frac{\partial \phi}{\partial x}\right)^\intercal\right]^{-1} \delta, \tag{B.11}$$

where the derivative terms are understood as a matrix where the derivative of the functions with respect to the arguments is in the columns for each argument,

$$\left[\left(\frac{\partial \phi}{\partial x}\right) \mathrm{cov}\,\{x\} \left(\frac{\partial \phi}{\partial x}\right)^\intercal\right]_{ij} = \sum_{ab} \frac{\partial \phi_i}{\partial x_a} \frac{\partial \phi_j}{\partial x_b} \mathrm{cov}(x_a, x_b). \tag{B.12}$$

This is computationally more expensive since now also the denominator of the function depends on the parameters $\{\theta\}$ and needs to be reevaluated in every iteration of the optimization. The case $\mathrm{cov}\,\{x\} = \mathrm{diag}\,(\mathrm{var}\,\{x\})$ is included as a special case.

C Continuum limit fit parameters

The parametrizations of lattice artefacts within our simultaneous fits to the masses and decay constants are defined in sec. 5.2. In sec. 5.3 we explain how 17 different parametrizations were selected. These are enumerated and defined in tab. 5.1. The six continuum limit fit parameters (LECs) for each of these fits are given in tab. 5.2. Here, in tab. C.1 we list the results for the unknown $\mathcal{O}(a)$ improvement coefficients within eqs. (5.29) and (5.30). Their parametrizations are given in eq. (5.31). In tabs. C.2 and C.3 we list the $\mathcal{O}(a^2)$ coefficients, defined in eq. (5.32) for both octet and both singlet decay constants, respectively.

id	χ^2/N_{df}	d_A^l	$\tilde{d}_A^l$	δc_A^l	f_A^l
1	1.75	—	—	—	—
2	1.63	$1.2\,(^{7}_{1.0})$	$-0.33\,(^{77}_{54})$	$0.047\,(^{46}_{31})$	$-0.511\,(^{241}_{99})$
3	1.64	$1.68\,(^{22}_{94})$	$-0.65\,(^{45}_{1.25})$	—	$-0.498\,(^{483}_{68})$
4	1.64	$1.30\,(^{29}_{57})$	—	—	$-0.456\,(^{217}_{78})$
5	1.52	$1.58\,(^{16}_{60})$	—	—	$-0.833\,(^{366}_{60})$
6	1.58	$1.49\,(^{15}_{56})$	—	—	$-0.714\,(^{397}_{48})$
7	1.47	$1.84\,(^{27}_{51})$	—	—	$-0.689\,(^{229}_{63})$
8	1.56	$1.66\,(^{40}_{52})$	—	—	$-0.429\,(^{202}_{87})$
9	1.49	$1.66\,(^{17}_{57})$	—	—	$-0.782\,(^{336}_{35})$
10	1.63	$0.82\,(^{30}_{52})$	—	—	$-0.500\,(^{239}_{95})$
11	1.65	$1.11\,(^{30}_{57})$	—	—	$-0.478\,(^{259}_{89})$
12	1.64	$1.13\,(^{37}_{49})$	—	—	$-0.452\,(^{227}_{81})$
13	1.47	$1.72\,(^{32}_{57})$	—	—	$-0.696\,(^{208}_{60})$
14	1.48	$1.76\,(^{27}_{57})$	—	—	$-0.702\,(^{220}_{51})$
15	1.47	$1.67\,(^{16}_{57})$	—	—	$-0.734\,(^{242}_{43})$
16	1.48	$1.88\,(^{25}_{55})$	—	—	$-0.698\,(^{223}_{45})$
17	1.46	$1.62\,(^{19}_{52})$	—	—	$-0.728\,(^{224}_{47})$

TABLE C.1: Fit results for the unknown $\mathcal{O}(a)$ improvement coefficients, see eqs. (5.29), (5.30) and (5.31). The fit ids are defined in tab. 5.1.

id	χ^2/N_{df}	$l_{F^8_\eta}$	$m_{F^8_\eta}$	$n_{F^8_\eta}$	$l_{F^8_{\eta'}}$	$m_{F^8_{\eta'}}$	$n_{F^8_{\eta'}}$
1	1.75	—	—	—	—	—	—
2	1.63	—	—	—	—	—	—
3	1.64	—	—	—	—	—	—
4	1.64	—	—	—	—	—	—
5	1.52	$-0.051\,(^{49}_{37})$	—	—	$-0.58\,(^{29}_{15})$	—	—
6	1.58	—	—	—	—	$-0.54\,(^{54}_{14})$	—
7	1.47	—	—	$-0.029\,(^{17}_{17})$	—	—	$-0.367\,(^{99}_{93})$
8	1.56	$-0.057\,(^{36}_{42})$	—	$-0.038\,(^{20}_{22})$	—	—	—
9	1.49	—	—	—	$-0.039\,(^{557}_{192})$	—	$-0.40\,(^{12}_{25})$
10	1.63	—	—	—	—	—	—
11	1.65	—	—	—	—	—	—
12	1.64	—	—	—	—	—	—
13	1.47	$-0.047\,(^{32}_{43})$	—	$-0.015\,(^{20}_{19})$	—	—	$-0.360\,(^{100}_{95})$
14	1.48	$-0.046\,(^{33}_{44})$	—	$-0.015\,(^{21}_{21})$	—	—	$-0.363\,(^{108}_{98})$
15	1.47	$-0.059\,(^{28}_{37})$	—	—	—	—	$-0.382\,(^{110}_{90})$
16	1.48	—	—	$-0.028\,(^{16}_{17})$	—	—	$-0.371\,(^{104}_{82})$
17	1.46	$-0.061\,(^{24}_{33})$	—	—	—	—	$-0.379\,(^{98}_{96})$

TABLE C.2: Fit parameters, accompanying quadratic lattice effects for the octet decay constants, see eq. (5.32). The fit ids are defined in tab. 5.1.

id	χ^2/N_{df}	$l_{F^0_\eta}$	$m_{F^0_\eta}$	$n_{F^0_\eta}$	$l_{F^0_{\eta'}}$	$m_{F^0_{\eta'}}$	$n_{F^0_{\eta'}}$
1	1.75	—	—	—	—	—	—
2	1.63	—	—	—	—	—	—
3	1.64	—	—	—	—	—	—
4	1.64	—	—	—	—	—	—
5	1.52	—	—	—	—	—	—
6	1.58	—	—	—	—	—	—
7	1.47	—	—	—	—	—	—
8	1.56	—	—	—	—	—	—
9	1.49	—	—	—	—	—	—
10	1.63	$0.062\,(^{1.304}_{983})$	$0.90\,(^{1.11}_{1.12})$	—	—	—	—
11	1.65	—	—	$0.25\,(^{26}_{25})$	—	$-0.14\,(^{18}_{12})$	$0.065\,(^{44}_{52})$
12	1.64	—	—	—	$-0.106\,(^{106}_{41})$	—	—
13	1.47	—	—	—	—	—	—
14	1.48	—	—	—	—	—	$0.0085\,(^{219}_{355})$
15	1.47	—	—	—	—	—	$0.011\,(^{21}_{37})$
16	1.48	—	—	—	—	—	$0.013\,(^{21}_{36})$
17	1.46	—	—	—	—	—	—

TABLE C.3: Fit parameters, accompanying quadratic lattice effects for the singlet decay constants, see eq. (5.32). The fit ids are defined in tab. 5.1.

D NNLO loop corrections: parametrization and fit results

Unlike in SU(3) ChPT, in large-N_c U(3) ChPT meson loops only enter at NNLO in the power counting because formally these contributions are of $\mathcal{O}(\delta^2)$. Therefore, the expressions in sec. 5.1.2 do not contain chiral logarithms or a dependence on the EFT renormalization scale μ_{EFT}. We define the loop functions

$$A_0(M^2) = -M^2 \log\left(\frac{M^2}{\mu_{\text{EFT}}^2}\right). \tag{D.1}$$

The octet and singlet decay constants of eqs. (5.24)–(5.27) receive the additional contributions [181, 182]

$$F_\eta^{8\,\text{NLO+loops}} = F_\eta^{8\,\text{NLO}} + \frac{3}{32\pi^2 F}\cos(\theta) A_0(M_K^2), \tag{D.2}$$

$$F_{\eta'}^{8\,\text{NLO+loops}} = F_{\eta'}^{8\,\text{NLO}} + \frac{3}{32\pi^2 F}\sin(\theta) A_0(M_K^2), \tag{D.3}$$

$$F_\eta^{0\,\text{NLO+loops}} = F_\eta^{0\,\text{NLO}} - \frac{1}{32\pi^2 F}\sin(\theta) A_0(M_\pi^2), \tag{D.4}$$

$$F_{\eta'}^{0\,\text{NLO+loops}} = F_{\eta'}^{0\,\text{NLO}} + \frac{1}{32\pi^2 F}\cos(\theta) A_0(M_\pi^2). \tag{D.5}$$

Moreover, the quark mass dependence of the mass mixing angle θ, defined in eq. (5.8), changes as the entries of the square mass matrix eq. (5.3) also receive additional contributions. Specifically, we

have to add to eqs. (5.21)–(5.23) [181]:[1]

$$
\begin{aligned}
(\mu_8^{\text{NLO+loops}})^2 = (\mu_8^{\text{NLO}})^2 + \frac{1}{48\pi^2 F^2}\Bigg[& \left(\frac{3}{2}\overline{M}^2 - \frac{1}{2}\delta M^2\right) A_0(M_\pi^2) \\
& - \left(4\overline{M}^2 + \frac{2}{3}\delta M^2\right) A_0(M_K^2) \\
& + \left(\frac{5}{4}\overline{M}^2 + \frac{7}{12}\delta M^2\right)\left(A_0(\tilde{M}_\eta^2) + A_0(\tilde{M}_{\eta'}^2)\right) \\
& + \frac{2\sqrt{2}\sin(2\theta^{\text{LO}}) + \cos(2\theta^{\text{LO}})}{4}\left(\overline{M}^2 + \delta M^2\right)\left(A_0(\tilde{M}_\eta^2) - A_0(\tilde{M}_{\eta'}^2)\right)\Bigg],
\end{aligned}
\tag{D.6}
$$

$$
\begin{aligned}
(\mu_0^{\text{NLO+loops}})^2 = (\mu_0^{\text{NLO}})^2 + \frac{1}{48\pi^2 F^2}\Bigg[& \left(3\overline{M}^2 - \delta M^2\right) A_0(M_\pi^2) \\
& + \left(4\overline{M}^2 + \frac{2}{3}\delta M^2\right) A_0(M_K^2) \\
& + \left(\overline{M}^2 + \frac{1}{6}\delta M^2\right)\left(A_0(\tilde{M}_\eta^2) + A_0(\tilde{M}_{\eta'}^2)\right) \\
& + \frac{2\sqrt{2}\sin(2\theta^{\text{LO}}) + \cos(2\theta^{\text{LO}})}{6}\delta M^2\left(A_0(\tilde{M}_\eta^2) - A_0(\tilde{M}_{\eta'}^2)\right)\Bigg],
\end{aligned}
\tag{D.7}
$$

$$
\begin{aligned}
(\mu_{80}^{\text{NLO+loops}})^2 = (\mu_{80}^{\text{NLO}})^2 + \frac{\sqrt{2}}{48\pi^2 F^2}\Bigg[& \left(\frac{3}{2}\overline{M}^2 - \frac{1}{2}\delta M^2\right) A_0(M_\pi^2) \\
& - \left(\overline{M}^2 + \frac{1}{6}\delta M^2\right) A_0(M_K^2) \\
& - \left(\frac{1}{4}\overline{M}^2 + \frac{5}{12}\delta M^2\right)\left(A_0(\tilde{M}_\eta^2) + A_0(\tilde{M}_{\eta'}^2)\right) \\
& - \frac{2\sqrt{2}\sin(2\theta^{\text{LO}}) + \cos(2\theta^{\text{LO}})}{4}\left(\overline{M}^2 + \frac{1}{3}\delta M^2\right)\left(A_0(\tilde{M}_\eta^2) - A_0(\tilde{M}_{\eta'}^2)\right)\Bigg].
\end{aligned}
\tag{D.8}
$$

θ^{LO} corresponds to the mass mixing angle eq. (5.8), evaluated at LO, eqs. (5.17)–(5.19). $\tilde{M}_\eta$ and $\tilde{M}_{\eta'}$ denote the η and η' masses, computed at LO via eqs. (5.9)–(5.11) and (5.17)–(5.19).

[1] Note that there are misprints within the normalizations of eqs. (C11)–(C13) in [181].

Carrying out the analysis of our masses and decay constants, including the NNLO loops, we obtain for the LECs at $\mu_{\rm EFT} = 0.770\,\text{GeV}$ in the $N_f = 3\,\overline{\text{MS}}$ scheme at $\mu = \infty$:

$$\begin{aligned}
L_5 &= 1.97\,\tbinom{16}{11}_{\text{stat}}\,\tbinom{0}{19}_{a}\,\tbinom{0}{23}_{\chi} \times 10^{-3},\\
L_8 &= 0.848\,\tbinom{126}{109}_{\text{stat}}\,\tbinom{0}{124}_{a}\,\tbinom{0}{113}_{\chi} \times 10^{-3},\\
M_0 &= 1.78\,\tbinom{3}{5}_{\text{stat}}\,\tbinom{2}{2}_{a}\,\tbinom{0}{0}_{\chi}\,\tbinom{1}{1}_{\text{renorm}}\,(8t_0^{\chi})^{-1/2}\\
&= 837\,\tbinom{13}{23}_{\text{stat}}\,\tbinom{11}{11}_{\text{syst}}\,(12)_{t_0}\,\text{MeV},\\
F &= 0.1680\,\tbinom{38}{66}_{\text{stat}}\,\tbinom{69}{0}_{a}\,\tbinom{39}{0}_{\chi}\,\tbinom{4}{1}_{\text{renorm}}\,(8t_0^{\chi})^{-1/2}\\
&= 78.97\,\tbinom{1.78}{3.10}_{\text{stat}}\,\tbinom{3.71}{17}_{\text{syst}}\,(1.18)_{t_0}\,\text{MeV},\\
\Lambda_1 &= -0.10\,\tbinom{2}{4}_{\text{stat}}\,\tbinom{5}{4}_{a}\,\tbinom{0}{2}_{\chi}\,\tbinom{4}{2}_{\text{renorm}}\,,\\
\tilde{\Lambda} &= -1.0\,\tbinom{2}{3}_{\text{stat}}\,\tbinom{4}{0}_{a}\,\tbinom{2}{0}_{\chi}\,\tbinom{0}{1}_{\text{renorm}}\,,\\
\Lambda_2 &= 0.45\,\tbinom{16}{12}_{\text{stat}}\,\tbinom{1}{19}_{a}\,\tbinom{0}{12}_{\chi}\,\tbinom{6}{3}_{\text{renorm}}
\end{aligned} \tag{D.9}$$

with $\chi^2/N_{\text{df}} \approx 312/122 \approx 2.56$. Note that only M_0, Λ_1 and Λ_2 depend on the QCD scale μ.

E Continuum limit results for decay constants in various representations

We list the four decay constants, in units of $(8t_0^{\rm ph})^{-1/2}$ and in MeV. In tab. E.1 we collect the results of our simultaneous fits to the masses and decay constants, including the statistical and systematic errors, while in tab. E.2 the corresponding results are shown, using the experimental masses of the η and the η' mesons as an additional input (priors).

In each table we list the decay constants in both the octet/singlet and the light/strange flavour bases. The conversion is given in eq. (2.103). In addition, we give the parameters of the corresponding two-angle representations eqs. (2.100) and (2.102). All these results are given at four distinct renormalization scales: $\mu = 1\,\text{GeV}$, $\mu = 2\,\text{GeV}$, $\mu = 10\,\text{GeV}$ and $\mu = \infty$, where all the values refer to the $\overline{\text{MS}}$ scheme for $N_f = 3$ active flavours. Only the octet decay constants F_η^8, $F_{\eta'}^8$ and F^8 as well as the angles θ_8 and θ_0 are scale independent. We remark that in the latter case the scale dependence cancels since $\tan(\theta_0) = -F_\eta^0/F_{\eta'}^0$.

Table E.1: Decay constants in various representations and at several renormalization scales.

octet/singlet basis, state representation				
F_η^8			$0.2219\,\binom{18}{37}_{\text{stat}}\binom{17}{24}_{a}\binom{10}{2}_{\chi}(8t_0^{\rm ph})^{-1/2}$	$105.4\,\binom{9}{1.8}_{\text{stat}}\binom{9}{1.1}_{\text{syst}}(1.3)_{t_0}$ MeV
$F_{\eta'}^8$			$-0.0939\,\binom{28}{100}_{\text{stat}}\binom{84}{0}_{a}\binom{58}{82}_{\chi}(8t_0^{\rm ph})^{-1/2}$	$-44.6\,\binom{1.3}{4.8}_{\text{stat}}\binom{4.9}{3.9}_{\text{syst}}(6)_{t_0}$ MeV
F_η^0	$\mu =$	∞	$0.0224\,\binom{53}{30}_{\text{stat}}\binom{28}{0}_{a}\binom{5}{21}_{\chi}\binom{20}{8}_{\text{renorm}}(8t_0^{\rm ph})^{-1/2}$	$10.6\,\binom{2.5}{1.4}_{\text{stat}}\binom{1.4}{1.4}_{\text{syst}}(1)_{t_0}$ MeV
		1 GeV	$0.0255\,\binom{60}{35}_{\text{stat}}\binom{32}{0}_{a}\binom{5}{24}_{\chi}\binom{23}{9}_{\text{renorm}}(8t_0^{\rm ph})^{-1/2}$	$12.1\,\binom{2.9}{1.6}_{\text{stat}}\binom{1.6}{1.6}_{\text{syst}}(2)_{t_0}$ MeV
		2 GeV	$0.0241\,\binom{57}{33}_{\text{stat}}\binom{30}{0}_{a}\binom{5}{23}_{\chi}\binom{22}{9}_{\text{renorm}}(8t_0^{\rm ph})^{-1/2}$	$11.4\,\binom{2.7}{1.5}_{\text{stat}}\binom{1.5}{1.5}_{\text{syst}}(1)_{t_0}$ MeV
		10 GeV	$0.0233\,\binom{55}{32}_{\text{stat}}\binom{29}{0}_{a}\binom{5}{22}_{\chi}\binom{21}{8}_{\text{renorm}}(8t_0^{\rm ph})^{-1/2}$	$11.1\,\binom{2.6}{1.5}_{\text{stat}}\binom{1.5}{1.5}_{\text{syst}}(1)_{t_0}$ MeV
$F_{\eta'}^0$	$\mu =$	∞	$0.1974\,\binom{14}{48}_{\text{stat}}\binom{0}{31}_{a}\binom{4}{27}_{\chi}\binom{52}{26}_{\text{renorm}}(8t_0^{\rm ph})^{-1/2}$	$93.77\,\binom{67}{2.29}_{\text{stat}}\binom{1.24}{3.18}_{\text{syst}}(1.18)_{t_0}$ MeV
		1 GeV	$0.2247\,\binom{16}{55}_{\text{stat}}\binom{0}{36}_{a}\binom{5}{31}_{\chi}\binom{60}{29}_{\text{renorm}}(8t_0^{\rm ph})^{-1/2}$	$106.7\,\binom{8}{2.6}_{\text{stat}}\binom{1.4}{3.6}_{\text{syst}}(1.3)_{t_0}$ MeV
		2 GeV	$0.2122\,\binom{15}{52}_{\text{stat}}\binom{0}{34}_{a}\binom{5}{29}_{\chi}\binom{56}{28}_{\text{renorm}}(8t_0^{\rm ph})^{-1/2}$	$100.8\,\binom{7}{2.5}_{\text{stat}}\binom{1.3}{3.4}_{\text{syst}}(1.3)_{t_0}$ MeV
		10 GeV	$0.2051\,\binom{15}{50}_{\text{stat}}\binom{0}{33}_{a}\binom{5}{28}_{\chi}\binom{54}{27}_{\text{renorm}}(8t_0^{\rm ph})^{-1/2}$	$97.41\,\binom{69}{2.38}_{\text{stat}}\binom{1.29}{3.30}_{\text{syst}}(1.23)_{t_0}$ MeV
octet/singlet basis, angle representation				
F^8			$0.2410\,\binom{23}{16}_{\text{stat}}\binom{11}{50}_{a}\binom{38}{12}_{\chi}(8t_0^{\rm ph})^{-1/2}$	$114.5\,\binom{1.1}{8}_{\text{stat}}\binom{1.9}{2.5}_{\text{syst}}(1.4)_{t_0}$ MeV
θ_8			$-0.400\,\binom{9}{45}_{\text{stat}}\binom{30}{0}_{a}\binom{24}{30}_{\chi}$	$-22.9\,\binom{5}{2.6}_{\text{stat}}\binom{2.2}{1.7}^{\circ}_{\text{syst}}$
F^0	$\mu =$	∞	$0.1987\,\binom{12}{42}_{\text{stat}}\binom{0}{31}_{a}\binom{2}{29}_{\chi}\binom{54}{27}_{\text{renorm}}(8t_0^{\rm ph})^{-1/2}$	$94.37\,\binom{57}{2.01}_{\text{stat}}\binom{1.26}{3.26}_{\text{syst}}(1.19)_{t_0}$ MeV
		1 GeV	$0.2262\,\binom{14}{48}_{\text{stat}}\binom{0}{35}_{a}\binom{2}{33}_{\chi}\binom{62}{30}_{\text{renorm}}(8t_0^{\rm ph})^{-1/2}$	$107.4\,\binom{6}{2.3}_{\text{stat}}\binom{1.4}{3.7}_{\text{syst}}(1.4)_{t_0}$ MeV
		2 GeV	$0.2136\,\binom{13}{46}_{\text{stat}}\binom{0}{33}_{a}\binom{2}{31}_{\chi}\binom{58}{29}_{\text{renorm}}(8t_0^{\rm ph})^{-1/2}$	$101.5\,\binom{6}{2.2}_{\text{stat}}\binom{1.4}{3.5}_{\text{syst}}(1.3)_{t_0}$ MeV
		10 GeV	$0.2064\,\binom{12}{44}_{\text{stat}}\binom{0}{32}_{a}\binom{2}{30}_{\chi}\binom{56}{28}_{\text{renorm}}(8t_0^{\rm ph})^{-1/2}$	$98.04\,\binom{59}{2.09}_{\text{stat}}\binom{1.31}{3.39}_{\text{syst}}(1.24)_{t_0}$ MeV
θ_0			$-0.113\,\binom{15}{29}_{\text{stat}}\binom{0}{15}_{a}\binom{11}{3}_{\chi}$	$-6.5\,\binom{9}{1.7}_{\text{stat}}\binom{6}{9}^{\circ}_{\text{syst}}$

Continued on next page

Table E.1 (continued): Decay constants at various scales.

light/strange basis, state representation				
F^ℓ_η	$\mu =$	∞	$0.1464\,\binom{27}{21}_{\text{stat}}\binom{23}{2}_{a}\binom{3}{14}_{\chi}\binom{17}{7}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$69.56\,\binom{1.28}{1.01}_{\text{stat}}\binom{1.13}{1.05}_{\text{syst}}(88)_{t_0}$ MeV
		1 GeV	$0.1490\,\binom{34}{23}_{\text{stat}}\binom{25}{2}_{a}\binom{3}{17}_{\chi}\binom{19}{8}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$70.76\,\binom{1.60}{1.10}_{\text{stat}}\binom{1.24}{1.21}_{\text{syst}}(89)_{t_0}$ MeV
		2 GeV	$0.1478\,\binom{31}{23}_{\text{stat}}\binom{24}{2}_{a}\binom{3}{16}_{\chi}\binom{18}{7}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$70.21\,\binom{1.46}{1.10}_{\text{stat}}\binom{1.19}{1.13}_{\text{syst}}(89)_{t_0}$ MeV
		10 GeV	$0.1471\,\binom{29}{22}_{\text{stat}}\binom{23}{2}_{a}\binom{3}{15}_{\chi}\binom{17}{7}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$69.89\,\binom{1.37}{1.05}_{\text{stat}}\binom{1.16}{1.09}_{\text{syst}}(88)_{t_0}$ MeV
$F^\ell_{\eta'}$	$\mu =$	∞	$0.1070\,\binom{24}{88}_{\text{stat}}\binom{26}{3}_{a}\binom{11}{44}_{\chi}\binom{37}{20}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$50.82\,\binom{1.16}{4.18}_{\text{stat}}\binom{1.65}{2.75}_{\text{syst}}(64)_{t_0}$ MeV
		1 GeV	$0.1293\,\binom{24}{92}_{\text{stat}}\binom{23}{5}_{a}\binom{8}{43}_{\chi}\binom{43}{22}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$61.41\,\binom{1.15}{4.36}_{\text{stat}}\binom{1.59}{2.93}_{\text{syst}}(78)_{t_0}$ MeV
		2 GeV	$0.1191\,\binom{24}{89}_{\text{stat}}\binom{25}{4}_{a}\binom{10}{44}_{\chi}\binom{41}{21}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$56.57\,\binom{1.16}{4.23}_{\text{stat}}\binom{1.62}{2.84}_{\text{syst}}(71)_{t_0}$ MeV
		10 GeV	$0.1132\,\binom{24}{88}_{\text{stat}}\binom{26}{4}_{a}\binom{10}{44}_{\chi}\binom{39}{20}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$53.79\,\binom{1.16}{4.16}_{\text{stat}}\binom{1.64}{2.80}_{\text{syst}}(68)_{t_0}$ MeV
F^s_η	$\mu =$	∞	$(-0.1683\,\binom{59}{30}_{\text{stat}}\binom{36}{5}_{a}\binom{4}{19}_{\chi}\binom{12}{4}_{\text{renorm}}\,8t_0^{\text{ph}})^{-1/2}$	$-79.93\,\binom{2.78}{1.43}_{\text{stat}}\binom{1.72}{1.09}_{\text{syst}}(1.01)_{t_0}$ MeV
		1 GeV	$-0.1665\,\binom{63}{32}_{\text{stat}}\binom{38}{4}_{a}\binom{5}{20}_{\chi}\binom{13}{5}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$-79.08\,\binom{3.01}{1.50}_{\text{stat}}\binom{1.83}{1.17}_{\text{syst}}(100)_{t_0}$ MeV
		2 GeV	$-0.1673\,\binom{61}{31}_{\text{stat}}\binom{37}{5}_{a}\binom{5}{20}_{\chi}\binom{13}{5}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$-79.46\,\binom{2.91}{1.47}_{\text{stat}}\binom{1.78}{1.13}_{\text{syst}}(1.00)_{t_0}$ MeV
		10 GeV	$-0.1678\,\binom{60}{30}_{\text{stat}}\binom{36}{5}_{a}\binom{4}{19}_{\chi}\binom{12}{4}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$-79.69\,\binom{2.85}{1.45}_{\text{stat}}\binom{1.75}{1.11}_{\text{syst}}(1.01)_{t_0}$ MeV
$F^s_{\eta'}$	$\mu =$	∞	$0.1906\,\binom{71}{31}_{\text{stat}}\binom{0}{75}_{a}\binom{70}{63}_{\chi}\binom{38}{17}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$90.55\,\binom{3.37}{1.48}_{\text{stat}}\binom{3.41}{5.00}_{\text{syst}}(1.14)_{t_0}$ MeV
		1 GeV	$0.2064\,\binom{68}{32}_{\text{stat}}\binom{0}{77}_{a}\binom{70}{65}_{\chi}\binom{42}{19}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$98.04\,\binom{3.21}{1.53}_{\text{stat}}\binom{3.45}{5.20}_{\text{syst}}(1.24)_{t_0}$ MeV
		2 GeV	$0.1992\,\binom{70}{33}_{\text{stat}}\binom{0}{76}_{a}\binom{70}{64}_{\chi}\binom{40}{18}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$94.62\,\binom{3.30}{1.58}_{\text{stat}}\binom{3.43}{5.11}_{\text{syst}}(1.20)_{t_0}$ MeV
		10 GeV	$0.1951\,\binom{70}{33}_{\text{stat}}\binom{0}{76}_{a}\binom{70}{64}_{\chi}\binom{39}{18}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$92.65\,\binom{3.34}{1.55}_{\text{stat}}\binom{3.42}{5.05}_{\text{syst}}(1.17)_{t_0}$ MeV
light/strange basis, angle representation				
F^ℓ	$\mu =$	∞	$0.1814\,\binom{12}{49}_{\text{stat}}\binom{25}{1}_{a}\binom{0}{37}_{\chi}\binom{35}{17}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$86.14\,\binom{59}{2.32}_{\text{stat}}\binom{1.45}{2.45}_{\text{syst}}(1.09)_{t_0}$ MeV
		1 GeV	$0.1972\,\binom{12}{56}_{\text{stat}}\binom{27}{0}_{a}\binom{0}{41}_{\chi}\binom{43}{21}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$93.69\,\binom{59}{2.67}_{\text{stat}}\binom{1.62}{2.81}_{\text{syst}}(1.18)_{t_0}$ MeV
		2 GeV	$0.1898\,\binom{13}{53}_{\text{stat}}\binom{27}{1}_{a}\binom{0}{39}_{\chi}\binom{39}{19}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$90.16\,\binom{62}{2.52}_{\text{stat}}\binom{1.55}{2.65}_{\text{syst}}(1.14)_{t_0}$ MeV
		10 GeV	$0.1857\,\binom{13}{51}_{\text{stat}}\binom{26}{1}_{a}\binom{0}{38}_{\chi}\binom{37}{18}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$88.20\,\binom{61}{2.42}_{\text{stat}}\binom{1.50}{2.55}_{\text{syst}}(1.11)_{t_0}$ MeV
ϕ_ℓ		$\mu = \infty$	$0.631\,\binom{12}{44}_{\text{stat}}\binom{8}{6}_{a}\binom{8}{15}_{\chi}\binom{11}{6}_{\text{renorm}}$	$36.2\,\binom{7}{2.5}_{\text{stat}}\binom{9}{1.0}^{\circ}_{\text{syst}}$
		1 GeV	$0.715\,\binom{12}{42}_{\text{stat}}\binom{5}{6}_{a}\binom{7}{11}_{\chi}\binom{11}{6}_{\text{renorm}}$	$41.0\,\binom{7}{2.4}_{\text{stat}}\binom{8}{8}^{\circ}_{\text{syst}}$
		2 GeV	$0.678\,\binom{12}{43}_{\text{stat}}\binom{6}{6}_{a}\binom{7}{13}_{\chi}\binom{11}{6}_{\text{renorm}}$	$38.8\,\binom{7}{2.5}_{\text{stat}}\binom{8}{9}^{\circ}_{\text{syst}}$
		10 GeV	$0.656\,\binom{12}{44}_{\text{stat}}\binom{7}{6}_{a}\binom{7}{14}_{\chi}\binom{11}{6}_{\text{renorm}}$	$37.6\,\binom{7}{2.5}_{\text{stat}}\binom{8}{9}^{\circ}_{\text{syst}}$
F^s	$\mu =$	∞	$0.2543\,\binom{40}{28}_{\text{stat}}\binom{0}{81}_{a}\binom{64}{34}_{\chi}\binom{20}{10}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$120.8\,\binom{1.9}{1.3}_{\text{stat}}\binom{3.1}{4.3}_{\text{syst}}(1.5)_{t_0}$ MeV
		1 GeV	$0.2652\,\binom{40}{30}_{\text{stat}}\binom{0}{85}_{a}\binom{66}{38}_{\chi}\binom{24}{12}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$126.0\,\binom{1.9}{1.4}_{\text{stat}}\binom{3.2}{4.6}_{\text{syst}}(1.6)_{t_0}$ MeV
		2 GeV	$0.2601\,\binom{40}{29}_{\text{stat}}\binom{0}{83}_{a}\binom{65}{36}_{\chi}\binom{22}{11}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$123.6\,\binom{1.9}{1.4}_{\text{stat}}\binom{3.1}{4.4}_{\text{syst}}(1.6)_{t_0}$ MeV
		10 GeV	$0.2573\,\binom{40}{29}_{\text{stat}}\binom{0}{82}_{a}\binom{64}{35}_{\chi}\binom{21}{11}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$122.2\,\binom{1.9}{1.4}_{\text{stat}}\binom{3.1}{4.4}_{\text{syst}}(1.5)_{t_0}$ MeV
ϕ_s	$\mu =$	∞	$0.723\,\binom{13}{30}_{\text{stat}}\binom{20}{3}_{a}\binom{22}{13}_{\chi}\binom{6}{13}_{\text{renorm}}$	$41.4\,\binom{7}{1.7}_{\text{stat}}\binom{1.7}{1.1}^{\circ}_{\text{syst}}$
		1 GeV	$0.679\,\binom{12}{29}_{\text{stat}}\binom{18}{4}_{a}\binom{22}{11}_{\chi}\binom{6}{14}_{\text{renorm}}$	$38.9\,\binom{7}{1.7}_{\text{stat}}\binom{1.7}{1.0}^{\circ}_{\text{syst}}$
		2 GeV	$0.699\,\binom{12}{29}_{\text{stat}}\binom{19}{4}_{a}\binom{22}{12}_{\chi}\binom{6}{14}_{\text{renorm}}$	$40.0\,\binom{7}{1.7}_{\text{stat}}\binom{1.7}{1.1}^{\circ}_{\text{syst}}$
		10 GeV	$0.710\,\binom{12}{30}_{\text{stat}}\binom{19}{3}_{a}\binom{22}{12}_{\chi}\binom{6}{14}_{\text{renorm}}$	$40.7\,\binom{7}{1.7}_{\text{stat}}\binom{1.7}{1.1}^{\circ}_{\text{syst}}$

TABLE E.2: Decay constants, using the experimental η and η' masses as additional input (priors, see sec. 5.4.1) in various representations and at several renormalization scales.

octet/singlet basis, state representation				
F_η^8			$0.2180\,(^{25}_{32})_{\text{stat}}\,(^{15}_{22})_a\,(^{33}_{0})_\chi\,(8t_0^{\text{ph}})^{-1/2}$	$103.5\,(^{1.2}_{1.5})_{\text{stat}}\,(^{1.7}_{1.1})_{\text{syst}}\,(1.3)_{t_0}$ MeV
$F_{\eta'}^8$			$-0.105\,(^{5}_{9})_{\text{stat}}\,(^{7}_{0})_a\,(^{7}_{0})_\chi\,(8t_0^{\text{ph}})^{-1/2}$	$-50.0\,(^{2.2}_{4.2})_{\text{stat}}\,(^{4.8}_{3})_{\text{syst}}\,(6)_{t_0}$ MeV
F_η^0	$\mu =$	∞	$0.0276\,(^{34}_{36})_{\text{stat}}\,(^{52}_{0})_a\,(^{0}_{50})_\chi\,(^{26}_{11})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$13.1\,(^{1.6}_{1.7})_{\text{stat}}\,(^{2.5}_{2.7})_{\text{syst}}\,(2)_{t_0}$ MeV
		1 GeV	$0.0314\,(^{39}_{41})_{\text{stat}}\,(^{60}_{0})_a\,(^{0}_{56})_\chi\,(^{30}_{12})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$14.9\,(^{1.9}_{2.0})_{\text{stat}}\,(^{2.9}_{3.0})_{\text{syst}}\,(2)_{t_0}$ MeV
		2 GeV	$0.0297\,(^{37}_{39})_{\text{stat}}\,(^{56}_{0})_a\,(^{0}_{53})_\chi\,(^{28}_{11})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$14.1\,(^{1.7}_{1.8})_{\text{stat}}\,(^{2.7}_{2.9})_{\text{syst}}\,(2)_{t_0}$ MeV
		10 GeV	$0.0287\,(^{36}_{38})_{\text{stat}}\,(^{54}_{0})_a\,(^{0}_{51})_\chi\,(^{27}_{11})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$13.6\,(^{1.7}_{1.8})_{\text{stat}}\,(^{2.6}_{2.8})_{\text{syst}}\,(2)_{t_0}$ MeV
$F_{\eta'}^0$	$\mu =$	∞	$0.1941\,(^{15}_{39})_{\text{stat}}\,(^{4}_{10})_a\,(^{34}_{4})_\chi\,(^{44}_{23})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$92.21\,(^{69}_{1.85})_{\text{stat}}\,(^{1.96}_{2.14})_{\text{syst}}\,(1.16)_{t_0}$ MeV
		1 GeV	$0.2210\,(^{17}_{44})_{\text{stat}}\,(^{5}_{34})_a\,(^{39}_{4})_\chi\,(^{50}_{26})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$105.0\,(^{8}_{2.1})_{\text{stat}}\,(^{2.2}_{2.9})_{\text{syst}}\,(1.3)_{t_0}$ MeV
		2 GeV	$0.2087\,(^{16}_{42})_{\text{stat}}\,(^{4}_{32})_a\,(^{37}_{4})_\chi\,(^{47}_{24})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$99.14\,(^{74}_{1.99})_{\text{stat}}\,(^{2.11}_{2.70})_{\text{syst}}\,(1.25)_{t_0}$ MeV
		10 GeV	$0.2017\,(^{15}_{40})_{\text{stat}}\,(^{4}_{31})_a\,(^{36}_{4})_\chi\,(^{45}_{24})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$95.79\,(^{72}_{1.92})_{\text{stat}}\,(^{2.04}_{2.61})_{\text{syst}}\,(1.21)_{t_0}$ MeV
octet/singlet basis, angle representation				
F^8			$0.2421\,(^{22}_{26})_{\text{stat}}\,(^{8}_{50})_a\,(^{32}_{12})_\chi\,(8t_0^{\text{ph}})^{-1/2}$	$115.0\,(^{1.1}_{1.2})_{\text{stat}}\,(^{1.6}_{2.4})_{\text{syst}}\,(1.5)_{t_0}$ MeV
θ_8			$-0.450\,(^{21}_{36})_{\text{stat}}\,(^{24}_{0})_a\,(^{29}_{0})_\chi\,(^{1}_{5})_{\text{renorm}}$	$-25.8\,(^{1.2}_{2.1})_{\text{stat}}\,(^{2.2}_{0.3})^{\circ}_{\text{syst}}$
F^0	$\mu =$	∞	$0.1961\,(^{13}_{37})_{\text{stat}}\,(^{12}_{6})_a\,(^{28}_{7})_\chi\,(^{47}_{24})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$93.14\,(^{62}_{1.75})_{\text{stat}}\,(^{1.83}_{2.27})_{\text{syst}}\,(1.18)_{t_0}$ MeV
		1 GeV	$0.2232\,(^{15}_{42})_{\text{stat}}\,(^{13}_{26})_a\,(^{32}_{8})_\chi\,(^{53}_{27})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$106.0\,(^{7}_{2.0})_{\text{stat}}\,(^{2.1}_{2.9})_{\text{syst}}\,(1.3)_{t_0}$ MeV
		2 GeV	$0.2108\,(^{14}_{40})_{\text{stat}}\,(^{12}_{25})_a\,(^{30}_{8})_\chi\,(^{50}_{26})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$100.1\,(^{7}_{1.9})_{\text{stat}}\,(^{2.0}_{2.7})_{\text{syst}}\,(1.3)_{t_0}$ MeV
		10 GeV	$0.2037\,(^{14}_{38})_{\text{stat}}\,(^{12}_{24})_a\,(^{29}_{7})_\chi\,(^{49}_{25})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$96.76\,(^{65}_{1.82})_{\text{stat}}\,(^{1.90}_{2.60})_{\text{syst}}\,(1.22)_{t_0}$ MeV
θ_0			$-0.141\,(^{18}_{20})_{\text{stat}}\,(^{0}_{27})_a\,(^{27}_{0})_\chi$	$-8.1\,(^{1.0}_{1.1})_{\text{stat}}\,(^{1.5}_{1.5})^{\circ}_{\text{syst}}$
light/strange basis, state representation				
F_η^ℓ	$\mu =$	∞	$0.1484\,(^{20}_{21})_{\text{stat}}\,(^{31}_{0})_a\,(^{0}_{21})_\chi\,(^{21}_{9})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$70.48\,(^{93}_{99})_{\text{stat}}\,(^{1.53}_{1.42})_{\text{syst}}\,(89)_{t_0}$ MeV
		1 GeV	$0.1515\,(^{23}_{26})_{\text{stat}}\,(^{37}_{0})_a\,(^{0}_{27})_\chi\,(^{24}_{10})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$71.96\,(^{1.11}_{1.22})_{\text{stat}}\,(^{1.81}_{1.71})_{\text{syst}}\,(91)_{t_0}$ MeV
		2 GeV	$0.1501\,(^{22}_{23})_{\text{stat}}\,(^{34}_{0})_a\,(^{0}_{24})_\chi\,(^{22}_{9})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$71.29\,(^{1.03}_{1.10})_{\text{stat}}\,(^{1.68}_{1.58})_{\text{syst}}\,(90)_{t_0}$ MeV
		10 GeV	$0.1493\,(^{21}_{22})_{\text{stat}}\,(^{33}_{0})_a\,(^{0}_{23})_\chi\,(^{22}_{9})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$70.90\,(^{98}_{1.04})_{\text{stat}}\,(^{1.61}_{1.50})_{\text{syst}}\,(90)_{t_0}$ MeV
$F_{\eta'}^\ell$	$\mu =$	∞	$0.09773\,(^{362}_{686})_{\text{stat}}\,(^{441}_{12})_a\,(^{364}_{0})_\chi\,(^{285}_{164})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$46.42\,(^{1.72}_{3.26})_{\text{stat}}\,(^{2.83}_{1.35})_{\text{syst}}\,(59)_{t_0}$ MeV
		1 GeV	$0.1197\,(^{40}_{71})_{\text{stat}}\,(^{44}_{2})_a\,(^{36}_{0})_\chi\,(^{33}_{19})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$56.84\,(^{1.88}_{3.38})_{\text{stat}}\,(^{2.86}_{1.59})_{\text{syst}}\,(72)_{t_0}$ MeV
		2 GeV	$0.1096\,(^{40}_{70})_{\text{stat}}\,(^{44}_{1})_a\,(^{36}_{0})_\chi\,(^{31}_{18})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$52.08\,(^{1.90}_{3.32})_{\text{stat}}\,(^{2.84}_{1.48})_{\text{syst}}\,(66)_{t_0}$ MeV
		10 GeV	$0.1039\,(^{38}_{69})_{\text{stat}}\,(^{44}_{1})_a\,(^{36}_{0})_\chi\,(^{30}_{17})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$49.35\,(^{1.82}_{3.29})_{\text{stat}}\,(^{2.83}_{1.42})_{\text{syst}}\,(62)_{t_0}$ MeV
F_η^s	$\mu =$	∞	$-0.1620\,(^{44}_{40})_{\text{stat}}\,(^{48}_{7})_a\,(^{0}_{56})_\chi\,(^{16}_{6})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$-76.97\,(^{2.07}_{1.91})_{\text{stat}}\,(^{2.29}_{2.76})_{\text{syst}}\,(97)_{t_0}$ MeV
		1 GeV	$-0.1598\,(^{46}_{42})_{\text{stat}}\,(^{52}_{6})_a\,(^{0}_{60})_\chi\,(^{18}_{7})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$-75.93\,(^{2.20}_{2.00})_{\text{stat}}\,(^{2.49}_{2.97})_{\text{syst}}\,(96)_{t_0}$ MeV
		2 GeV	$-0.1609\,(^{45}_{41})_{\text{stat}}\,(^{50}_{6})_a\,(^{0}_{58})_\chi\,(^{17}_{7})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$-76.41\,(^{2.14}_{1.96})_{\text{stat}}\,(^{2.40}_{2.87})_{\text{syst}}\,(97)_{t_0}$ MeV
		10 GeV	$-0.1614\,(^{44}_{41})_{\text{stat}}\,(^{49}_{7})_a\,(^{0}_{57})_\chi\,(^{16}_{7})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$-76.68\,(^{2.11}_{1.93})_{\text{stat}}\,(^{2.35}_{2.82})_{\text{syst}}\,(97)_{t_0}$ MeV
$F_{\eta'}^s$	$\mu =$	∞	$0.1980\,(^{59}_{44})_{\text{stat}}\,(^{0}_{74})_a\,(^{24}_{58})_\chi\,(^{35}_{16})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$94.06\,(^{2.80}_{2.08})_{\text{stat}}\,(^{1.37}_{4.77})_{\text{syst}}\,(1.19)_{t_0}$ MeV
		1 GeV	$0.2135\,(^{58}_{47})_{\text{stat}}\,(^{0}_{76})_a\,(^{27}_{58})_\chi\,(^{39}_{18})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$101.4\,(^{2.8}_{2.2})_{\text{stat}}\,(^{1.5}_{4.9})_{\text{syst}}\,(1.3)_{t_0}$ MeV
		2 GeV	$0.2064\,(^{58}_{45})_{\text{stat}}\,(^{0}_{75})_a\,(^{25}_{58})_\chi\,(^{37}_{17})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$98.06\,(^{2.78}_{2.15})_{\text{stat}}\,(^{1.45}_{4.85})_{\text{syst}}\,(1.24)_{t_0}$ MeV
		10 GeV	$0.2024\,(^{59}_{45})_{\text{stat}}\,(^{0}_{74})_a\,(^{25}_{58})_\chi\,(^{36}_{17})_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$96.13\,(^{2.79}_{2.12})_{\text{stat}}\,(^{1.41}_{4.81})_{\text{syst}}\,(1.21)_{t_0}$ MeV

Continued on next page

Table E.2 (continued): Decay constants at various scales.

light/strange basis, angle representation				
F^ℓ	$\mu =$	∞	$0.1777\,\binom{25}{40}_{\text{stat}}\,\binom{55}{0}_{a}\,\binom{12}{4}_{\chi}\,\binom{33}{16}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$84.40\,\binom{1.17}{1.91}_{\text{stat}}\,\binom{2.78}{1.58}_{\text{syst}}\,(1.07)_{t_0}$ MeV
		1 GeV	$0.1931\,\binom{26}{44}_{\text{stat}}\,\binom{63}{0}_{a}\,\binom{13}{3}_{\chi}\,\binom{39}{19}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$91.70\,\binom{1.22}{2.08}_{\text{stat}}\,\binom{3.17}{1.87}_{\text{syst}}\,(1.16)_{t_0}$ MeV
		2 GeV	$0.1859\,\binom{25}{43}_{\text{stat}}\,\binom{59}{0}_{a}\,\binom{13}{3}_{\chi}\,\binom{36}{18}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$88.28\,\binom{1.20}{2.02}_{\text{stat}}\,\binom{3.00}{1.74}_{\text{syst}}\,(1.12)_{t_0}$ MeV
		10 GeV	$0.1819\,\binom{25}{42}_{\text{stat}}\,\binom{57}{0}_{a}\,\binom{13}{3}_{\chi}\,\binom{35}{17}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$86.38\,\binom{1.19}{1.99}_{\text{stat}}\,\binom{2.90}{1.66}_{\text{syst}}\,(1.09)_{t_0}$ MeV
ϕ_ℓ	$\mu =$	∞	$0.582\,\binom{19}{36}_{\text{stat}}\,\binom{11}{4}_{a}\,\binom{20}{0}_{\chi}\,\binom{7}{5}_{\text{renorm}}$	$33.3\,\binom{1.1}{2.1}_{\text{stat}}\,\binom{1.3}{0.4}^{\diamond}_{\text{syst}}$
		1 GeV	$0.669\,\binom{19}{34}_{\text{stat}}\,\binom{6}{5}_{a}\,\binom{21}{0}_{\chi}\,\binom{6}{5}_{\text{renorm}}$	$38.3\,\binom{1.1}{1.9}_{\text{stat}}\,\binom{1.3}{0.4}^{\diamond}_{\text{syst}}$
		2 GeV	$0.631\,\binom{19}{35}_{\text{stat}}\,\binom{8}{4}_{a}\,\binom{20}{0}_{\chi}\,\binom{7}{5}_{\text{renorm}}$	$36.2\,\binom{1.1}{2.0}_{\text{stat}}\,\binom{1.3}{0.4}^{\diamond}_{\text{syst}}$
		10 GeV	$0.608\,\binom{19}{35}_{\text{stat}}\,\binom{9}{4}_{a}\,\binom{19}{0}_{\chi}\,\binom{7}{5}_{\text{renorm}}$	$34.8\,\binom{1.1}{2.0}_{\text{stat}}\,\binom{1.3}{0.4}^{\diamond}_{\text{syst}}$
F^s	$\mu =$	∞	$0.2559\,\binom{35}{34}_{\text{stat}}\,\binom{0}{83}_{a}\,\binom{54}{25}_{\chi}\,\binom{17}{9}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$121.5\,\binom{1.7}{1.6}_{\text{stat}}\,\binom{2.6}{4.2}_{\text{syst}}\,(1.5)_{t_0}$ MeV
		1 GeV	$0.2667\,\binom{37}{35}_{\text{stat}}\,\binom{0}{86}_{a}\,\binom{57}{27}_{\chi}\,\binom{20}{10}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$126.7\,\binom{1.8}{1.7}_{\text{stat}}\,\binom{2.8}{4.4}_{\text{syst}}\,(1.6)_{t_0}$ MeV
		2 GeV	$0.2617\,\binom{37}{34}_{\text{stat}}\,\binom{0}{85}_{a}\,\binom{56}{26}_{\chi}\,\binom{19}{9}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$124.3\,\binom{1.7}{1.6}_{\text{stat}}\,\binom{2.7}{4.3}_{\text{syst}}\,(1.6)_{t_0}$ MeV
		10 GeV	$0.2589\,\binom{36}{34}_{\text{stat}}\,\binom{0}{84}_{a}\,\binom{55}{26}_{\chi}\,\binom{18}{9}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$123.0\,\binom{1.7}{1.6}_{\text{stat}}\,\binom{2.6}{4.3}_{\text{syst}}\,(1.6)_{t_0}$ MeV
ϕ_s	$\mu =$	∞	$0.686\,\binom{17}{23}_{\text{stat}}\,\binom{6}{4}_{a}\,\binom{24}{0}_{\chi}\,\binom{6}{14}_{\text{renorm}}$	$39.3\,\binom{1.0}{1.3}_{\text{stat}}\,\binom{1.5}{0.8}^{\diamond}_{\text{syst}}$
		1 GeV	$0.643\,\binom{16}{22}_{\text{stat}}\,\binom{6}{4}_{a}\,\binom{23}{0}_{\chi}\,\binom{6}{14}_{\text{renorm}}$	$36.8\,\binom{0.9}{1.3}_{\text{stat}}\,\binom{1.4}{0.8}^{\diamond}_{\text{syst}}$
		2 GeV	$0.662\,\binom{17}{23}_{\text{stat}}\,\binom{6}{4}_{a}\,\binom{23}{0}_{\chi}\,\binom{6}{14}_{\text{renorm}}$	$37.9\,\binom{1.0}{1.3}_{\text{stat}}\,\binom{1.4}{0.8}^{\diamond}_{\text{syst}}$
		10 GeV	$0.673\,\binom{17}{23}_{\text{stat}}\,\binom{6}{4}_{a}\,\binom{23}{0}_{\chi}\,\binom{6}{14}_{\text{renorm}}$	$38.6\,\binom{1.0}{1.3}_{\text{stat}}\,\binom{1.4}{0.8}^{\diamond}_{\text{syst}}$

F NLO expressions for pseudoscalar and gluonic matrix elements

In this appendix we derive the large-N_c ChPT expressions of the products of the pseudoscalar matrix elements with the quark masses $m_f H^f_{\eta^{(\prime)}}$ and the anomalous matrix elements $a_{\eta^{(\prime)}}$ to NLO.

This appendix has already been published in similar or verbatim form in [60].

We start from the octet and singlet AWIs eqs. (6.3) and (6.4). We apply these to states $|n\rangle \in \{|\eta\rangle, |\eta'\rangle\}$ (see eq. (6.7) for the octet case) and replace $\langle\Omega|\partial_\mu A^a_\mu|n\rangle = M_n^2 F_n^a$, where $a \in \{8, 0\}$. This gives

$$F_n^8 M_n^2 = \frac{2}{\sqrt{3}} m_\ell H_n^\ell - 2\sqrt{\frac{2}{3}} m_s H_n^s, \tag{F.1}$$

$$F_n^0 M_n^2 = 2\sqrt{\frac{2}{3}} m_\ell H_n^\ell + \frac{2}{\sqrt{3}} m_s H_n^s + \sqrt{\frac{2}{3}} a_n, \tag{F.2}$$

where $H_n^q = \langle\Omega|P^q|n\rangle$ are the pseudoscalar matrix elements and $H_n^\ell = (H_n^u + H_n^d)/\sqrt{2}$. The anomaly terms $a_n = 2\langle\Omega|\omega|n\rangle$ are the matrix elements of the topological charge density ω. The left-hand sides of the above equations are functions of $\overline{M}^2$ and δM^2 and can be parameterized in terms of the six LECs F, M_0^2, L_5, L_8, Λ_1 and Λ_2.

In terms of the large-N_c ChPT power counting, one finds $\{L_5, L_8\} \sim \delta^{-1}$, $F \sim \delta^{-1/2}$, $\{\sin\theta, \cos\theta, B_0\} \sim \delta^0$ and $\{m_q, M^2, M_0^2, \Lambda_1, \Lambda_2\} \sim \delta^1$. This counting is consistent with the LO GMOR relation $M_\pi^2 = 2B_0 m_\ell$, where $\langle S\rangle = \langle\bar{q}q\rangle \sim N_c = \mathcal{O}(\delta^{-1})$ and $B_0 = -\langle\bar{q}q\rangle/F^2 = \mathcal{O}(\delta^0)$. Using the AWIs, $P \sim \delta^{-1}$ implies that $A_\mu \sim \delta^{-1/2}$ and $\omega \sim \delta^0$. The latter is consistent with the topological susceptibility $\tau \sim \delta^0$ as one would expect from the Witten-Veneziano relation $M_0^2 = 6\tau_0/F^2$, where $M_0^2 \sim \delta^1$ and $F^2 \sim \delta^{-1}$. Finally, the parametrization of the axial matrix elements $\langle\Omega|A^a_\mu|n\rangle = ip_\mu F_n^a \sim \delta^0$ means that $|n\rangle \sim \delta^{1/2}$. Note that for the meson masses and the matrix elements the above counting applies to the LO expressions and there will be higher order corrections in δ.

Using these counting rules, eqs. (F.1) and (F.2) start at $\mathcal{O}(\delta^{1/2})$ and should be expanded up to $\mathcal{O}(\delta^{3/2})$ to obtain predictions at NLO. We carried out the matching, first at LO and then at NLO. The NLO results are presented below. The decay constants can be expressed in terms of the LECs and meson masses at $\mathcal{O}(\delta^{-1/2})$ and $\mathcal{O}(\delta^{1/2})$ (LO and NLO) via eqs. (5.24)–(5.27) while M_η^2 and $M_{\eta'}^2$ can be parameterized in terms of the squared mass matrix (with elements eqs. (5.21)–(5.23)) via the rotation (5.4) with the angle defined in eq. (5.8) at $\mathcal{O}(\delta^1)$ and $\mathcal{O}(\delta^2)$. We truncate the product at $\mathcal{O}(\delta^{3/2})$. Accordingly, we replace the quark masses on the right-hand sides by combinations of

$\overline{M}^2$ and δM^2 via the NLO large-N_c GMOR relations

$$2m_\ell B_0 = M_\pi^2 \left(1 - 8\frac{2L_8 - L_5}{F^2} M_\pi^2\right), \tag{F.3}$$

$$(m_\ell + m_s)\, B_0 = M_K^2 \left(1 - 8\frac{2L_8 - L_5}{F^2} M_K^2\right). \tag{F.4}$$

Then both sides are polynomials in δM^2 and $\overline{M}^2$ of degree one and two at LO and at NLO, respectively. The pion masses also enter through $\sin\theta$ and $\cos\theta$. Since we carry out the matching in terms of powers of δ, we keep the sine and the cosine (that are of $\mathcal{O}(\delta^0)$) in the coefficient functions. Note that at LO θ only depends on δM^2 (as well as on the LEC M_0^2).

Equations (F.1)–(F.2) amount to four identities ($n \in \{\eta, \eta'\}$ and $a \in \{8, 0\}$) but we have six unknown functions on the right-hand sides (four H_n^q and two a_n). Nevertheless, we are able to determine these unambiguously since the relations should hold for any combination of δM^2 and $\overline{M}^2 > 3\delta M^2$. It is instructive first to inspect the special case $\delta M = 0$, where $\sin\theta = 0$. Then the substitution of the meson masses simplifies: $M_\eta^2 = \mu_8^2 = \overline{M}^2 + \delta M^2/3 + \ldots$, $M_{\eta'}^2 = \mu_0^2 = \overline{M}^2 + M_0^2 + \ldots$. Moreover, in this limit there exist only two non-trivial relations (for F_η^8 and $F_{\eta'}^0$) since $H_\eta^s + \sqrt{2} H_\eta^\ell = 0$, $H_{\eta'}^\ell - \sqrt{2} H_\eta^s = 0$ and $a_\eta = 0$. These three equalities are also obvious from the respective Wick contractions. In the vicinity of this limit, to leading order, these combinations must be proportional to $\sin\theta$ or to δM^2, where $\sin\theta \propto \delta M^2$ for small θ. One can easily see that in the limit $\delta M = 0$, to leading order also $H_\eta^s = -(\sqrt{2}/3) B_0 F = -\sqrt{2} H_{\eta'}^s$ holds and therefore, $a_{\eta'} = \sqrt{2/3}\, F M_{\eta'}^2 = \sqrt{2/3}\, F(M_0^2 + \overline{M}^2)$. Starting from these identifications and sorting all terms accordingly, where in the end we substitute back the GMOR relations and eliminate B_0, gives the following expressions for the pseudoscalar combinations:

$$
\begin{aligned}
m_\ell H_\eta^\ell = \cos\theta F \Bigg\{ & \left(\overline{M}^2 - \frac{\delta M^2}{3}\right)\frac{\sqrt{3}}{12}\left[(1-2\Lambda_1+4\Lambda_2)\left(1-\frac{1}{\cos(2\theta)}\right)+2\right] \\
& + M_0^2\left(1-\frac{3\overline{M}^2}{\delta M^2}\right)\frac{\sqrt{3}}{12}(1-\Lambda_1)\left(1-\frac{1}{\cos(2\theta)}\right) \\
& + \frac{\overline{M}^2\delta M^2}{3}\left[\frac{5\sqrt{3}}{3}\left(2\frac{2L_8-L_5}{F^2}+\frac{L_5}{F^2}\right)\left(1+\frac{1}{5\cos(2\theta)}\right)\right. \\
& \qquad \left. + \sqrt{6}\frac{L_5}{F^2}\tan\theta\left(1-\frac{1}{3\cos(2\theta)}\right)\right] \\
& + \overline{M}^2 M_0^2\left[\frac{2L_8-L_5}{F^2}\left(\frac{5\sqrt{3}}{3}+2\sqrt{3}\frac{\overline{M}^2}{\delta M^2}\right)\right. \\
& \qquad \left. - \left(\frac{\sqrt{6}}{3}\tan\theta+\sqrt{3}\right)\frac{L_5}{F^2}\right]\left(1-\frac{1}{\cos(2\theta)}\right) \\
& - \frac{\delta M^2 M_0^2}{3}\left[\frac{7\sqrt{3}}{3}\frac{2L_8-L_5}{F^2} - \left(\sqrt{3}+\frac{\sqrt{6}}{3}\tan\theta\right)\frac{L_5}{F^2}\right]\left(1-\frac{1}{\cos(2\theta)}\right) \\
& - 2\sqrt{3}\,\overline{M}^4\frac{2L_8-L_5}{F^2}\left(1-\frac{1}{3\cos(2\theta)}\right) \\
& - \delta M^4\left[\frac{4\sqrt{3}}{27}\frac{2L_8-L_5}{F^2}\left(1+\frac{1}{\cos(2\theta)}\right)+\frac{\sqrt{3}}{27}\frac{L_5}{F^2}\left(5+\frac{1}{\cos(2\theta)}\right)\right. \\
& \qquad \left. + \frac{\sqrt{6}}{9}\tan\theta\frac{L_5}{F^2}\left(1-\frac{1}{3\cos(2\theta)}\right)\right]\Bigg\},
\end{aligned}
\tag{F.5}
$$

$$
\begin{aligned}
m_s H_\eta^s = \cos\theta F \Bigg\{ & -\frac{\sqrt{6}}{8}\left(\overline{M}^2+\frac{2}{3}\delta M^2\right)\left[\left(1-\frac{\Lambda_1}{3}+\frac{2\Lambda_2}{3}\right)\left(1-\frac{1}{\cos(2\theta)}\right)+\frac{4}{3\cos(2\theta)}\right] \\
& -\frac{\sqrt{6}}{12}M_0^2(1-\Lambda_1)\left(1+\frac{3\overline{M}^2}{2\delta M^2}\right)\left(1-\frac{1}{\cos(2\theta)}\right) \\
& +\frac{\overline{M}^2\delta M^2}{3}\left[\frac{5\sqrt{6}}{3}\frac{2L_8-L_5}{F^2}\left(1+\frac{1}{5\cos(\theta)}\right)\right. \\
& \qquad -\frac{L_5}{F^2}\left(\frac{2\sqrt{6}}{3}\left(1+\frac{1}{2\cos(2\theta)}\right)\right. \\
& \qquad\quad \left.\left.+\sqrt{3}\tan\theta\left(1+\frac{1}{3\cos(2\theta)}\right)\right)\right]+ \\
& +\overline{M}^2M_0^2\left[\left(\frac{5\sqrt{6}}{6}+\sqrt{6}\frac{\overline{M}^2}{\delta M^2}\right)\frac{2L_8-L_5}{F^2}-\frac{\sqrt{3}}{3}\frac{L_5}{F^2}\tan\theta\right]\left(1-\frac{1}{\cos(2\theta)}\right) \\
& -\delta M^2M_0^2\left(\frac{7\sqrt{6}}{18}\frac{2L_8-L_5}{F^2}+\frac{2\sqrt{3}}{9}\frac{L_5}{F^2}\tan\theta\right)\left(1-\frac{1}{\cos(2\theta)}\right) \\
& +\sqrt{6}\overline{M}^4\frac{2L_8-L_5}{F^2}\left(1+\frac{1}{3\cos(2\theta)}\right) \\
& -\frac{\delta M^4}{9}\left[\frac{\sqrt{6}}{6}\frac{2L_8-L_5}{F^2}\left(31+\frac{13}{\cos(2\theta)}\right)+\frac{2\sqrt{6}}{3}\frac{L_5}{F^2}\left(2+\frac{1}{\cos(2\theta)}\right)\right. \\
& \qquad \left.+2\sqrt{3}\tan\theta\frac{L_5}{F^2}\left(1+\frac{1}{3\cos(2\theta)}\right)\right]\Bigg\},
\end{aligned}
\tag{F.6}
$$

$$
\begin{aligned}
m_\ell H^\ell_{\eta'} = \sin\theta F \Bigg\{ & \left(\overline{M}^2 - \frac{\delta M^2}{3}\right)\frac{\sqrt{3}}{12}\left[(1-2\Lambda_1+4\Lambda_2)\left(1+\frac{1}{\cos(2\theta)}\right)+2\right] \\
& + M_0^2\left(1-\frac{3\overline{M}^2}{\delta M^2}\right)\frac{\sqrt{3}}{12}(1-\Lambda_1)\left(1+\frac{1}{\cos(2\theta)}\right) \\
& + \frac{\overline{M}^2\delta M^2}{3}\left[\frac{5\sqrt{3}}{3}\left(2\frac{2L_8-L_5}{F^2}+\frac{L_5}{F^2}\right)\left(1-\frac{1}{5\cos(2\theta)}\right)\right. \\
& \qquad \left. -\sqrt{6}\frac{L_5}{F^2}\cot\theta\left(1+\frac{1}{3\cos(2\theta)}\right)\right] \\
& + \overline{M}^2 M_0^2\left[\frac{2L_8-L_5}{F^2}\left(\frac{5\sqrt{3}}{3}+2\sqrt{3}\frac{\overline{M}^2}{\delta M^2}\right)\right. \\
& \qquad \left. +\left(\frac{\sqrt{6}}{3}\cot\theta-\sqrt{3}\right)\frac{L_5}{F^2}\right]\left(1+\frac{1}{\cos(2\theta)}\right) \\
& - \frac{\delta M^2 M_0^2}{3}\left[\frac{7\sqrt{3}}{3}\frac{2L_8-L_5}{F^2}+\left(-\sqrt{3}+\frac{\sqrt{6}}{3}\cot\theta\right)\frac{L_5}{F^2}\right]\left(1+\frac{1}{\cos(2\theta)}\right) \\
& - 2\sqrt{3}\overline{M}^4\frac{2L_8-L_5}{F^2}\left(1+\frac{1}{3\cos(2\theta)}\right) \\
& - \frac{\sqrt{3}}{27}\delta M^4\left[4\frac{2L_8-L_5}{F^2}\left(1-\frac{1}{\cos(2\theta)}\right)+\left(5-\frac{1}{\cos(2\theta)}\right)\frac{L_5}{F^2}\right. \\
& \qquad \left. -3\sqrt{2}\cot\theta\frac{L_5}{F^2}\left(1+\frac{1}{3\cos(2\theta)}\right)\right]\Bigg\},
\end{aligned}
\tag{F.7}
$$

$$
\begin{aligned}
m_s H^s_{\eta'} = \sin\theta F \Bigg\{ & -\frac{\sqrt{6}}{8}\left(\overline{M}^2+\frac{2}{3}\delta M^2\right)\left[\left(1-\frac{\Lambda_1}{3}+\frac{2\Lambda_2}{3}\right)\left(1+\frac{1}{\cos(2\theta)}\right)-\frac{4}{3\cos(2\theta)}\right] \\
& -\frac{\sqrt{6}}{12}M_0^2(1-\Lambda_1)\left(1+\frac{3\overline{M}^2}{2\delta M^2}\right)\left(1+\frac{1}{\cos(2\theta)}\right) \\
& +\frac{\overline{M}^2\delta M^2}{3}\left[\frac{5\sqrt{6}}{3}\frac{2L_8-L_5}{F^2}\left(1-\frac{1}{5\cos(\theta)}\right)\right. \\
& \qquad \left. -\frac{L_5}{F^2}\left(\frac{2\sqrt{6}}{3}\left(1-\frac{1}{2\cos(2\theta)}\right)-\sqrt{3}\cot\theta\left(1-\frac{1}{3\cos(2\theta)}\right)\right)\right] \\
& +\overline{M}^2M_0^2\left[\left(\frac{5\sqrt{6}}{6}+\sqrt{6}\frac{\overline{M}^2}{\delta M^2}\right)\frac{2L_8-L_5}{F^2}+\frac{\sqrt{3}}{3}\frac{L_5}{F^2}\cot\theta\right]\left(1+\frac{1}{\cos(2\theta)}\right) \\
& -\delta M^2M_0^2\left(\frac{7\sqrt{6}}{18}\frac{2L_8-L_5}{F^2}-\frac{2\sqrt{3}}{9}\frac{L_5}{F^2}\cot\theta\right)\left(1+\frac{1}{\cos(2\theta)}\right) \\
& +\sqrt{6}\,\overline{M}^4\frac{2L_8-L_5}{F^2}\left(1-\frac{1}{3\cos(2\theta)}\right) \\
& -\frac{\delta M^4}{9}\left[\frac{\sqrt{6}}{6}\frac{2L_8-L_5}{F^2}\left(31-\frac{13}{\cos(2\theta)}\right)+\frac{4\sqrt{6}}{6}\frac{L_5}{F^2}\left(2-\frac{1}{\cos(2\theta)}\right)\right. \\
& \qquad \left. -2\sqrt{3}\frac{L_5}{F^2}\cot\theta\left(1-\frac{1}{3\cos(2\theta)}\right)\right]\Bigg\}.
\end{aligned}
\tag{F.8}
$$

The NLO expressions contain M^4 terms because the left-hand sides are already proportional to quark masses. One remark is in order: the η' matrix elements all start with $\sin\theta$. This does not mean that they vanish in the limit $\delta M = 0$ (where $\sin\theta = 0$). For small δM^2 one can expand $\sin\theta = -\sqrt{2}\delta M^2/(3M_0^2) + \ldots$, which cancels against a term $\propto \overline{M}^2M_0^2/\delta M^2$, resulting in the limiting case discussed above.

Finally, the gluonic matrix elements can be obtained via the singlet AWI:

$$
\begin{aligned}
a_\eta = \cos\theta F \Bigg\{ & \frac{\sqrt{3}}{6}\overline{M}^2\left(-1+\Lambda_1-2\Lambda_2+3\frac{M_0^2}{\delta M^2}\left(1-\Lambda_1\right)\right)\left(1-\frac{1}{\cos\left(2\theta\right)}\right) \\
& + \frac{\sqrt{3}}{9}\delta M^2\left[\left(-\Lambda_1+2\Lambda_2\right)\left(1-\frac{1}{\cos\left(2\theta\right)}\right)+2\right] \\
& - \frac{4\sqrt{3}}{9}\overline{M}^2\delta M^2\left[\frac{2L_8-L_5}{F^2}\left(5+\frac{1}{\cos\left(2\theta\right)}\right)+3\frac{L_5}{F^2}\right] \\
& - \frac{2\sqrt{3}}{3}\overline{M}^2 M_0^2\left(5\frac{2L_8-L_5}{F^2}-2\frac{L_5}{F^2}\right)\left(1-\frac{1}{\cos\left(2\theta\right)}\right) \\
& + \frac{2\sqrt{3}}{9}\delta M^2 M_0^2\left(7\frac{2L_8-L_5}{F^2}-4\frac{L_5}{F^2}\right)\left(1-\frac{1}{\cos\left(2\theta\right)}\right) \\
& + \frac{4\sqrt{3}}{3}\overline{M}^4\left(1-3\frac{M_0^2}{\delta M^2}\right)\frac{2L_8-L_5}{F^2}\left(1-\frac{1}{\cos\left(2\theta\right)}\right) \\
& + \frac{\sqrt{3}}{27}\delta M^4\left[\frac{2L_8-L_5}{F^2}\left(26+\frac{14}{\cos\left(2\theta\right)}\right)+8\frac{L_5}{F^2}\right]\Bigg\} \\
+ \sin\theta F \Bigg\{ & -\frac{\sqrt{6}}{3}\overline{M}^2\left[1+\frac{\Lambda_1}{2\cos\left(2\theta\right)}+\Lambda_2\left(1-\frac{1}{\cos\left(2\theta\right)}\right)\right] \\
& - \frac{\sqrt{6}}{18}\delta M^2\left(1+\frac{\Lambda_1}{2}\right)\left(1+\frac{1}{\cos\left(2\theta\right)}\right) \\
& - \frac{\sqrt{6}}{6}M_0^2\left(1-\frac{\Lambda_1}{2}\right)\left(1-\frac{1}{\cos\left(2\theta\right)}\right) \\
& - \frac{4\sqrt{6}}{9}\frac{L_5}{F^2}\overline{M}^2\delta M^2+\frac{4\sqrt{6}}{3}\overline{M}^2 M_0^2\frac{L_5}{F^2}\left(1-\frac{1}{\cos\left(2\theta\right)}\right)-\frac{4\sqrt{6}}{3}\frac{L_5}{F^2}\overline{M}^4 \\
& - \frac{2\sqrt{6}}{3}\delta M^4\left[\left(1+\frac{1}{3\cos\left(2\theta\right)}\right)\frac{2L_8-L_5}{F^2}-\frac{4}{9}\frac{L_5}{F^2}\right]\Bigg\},
\end{aligned}
\tag{F.9}
$$

$$
\begin{aligned}
a_{\eta'} = \cos\theta F \Bigg\{ & \frac{\sqrt{6}}{3}\overline{M}^2 \left[1 - \frac{\Lambda_1}{2\cos(2\theta)} + \Lambda_2\left(1 + \frac{1}{\cos(2\theta)}\right)\right] \\
& + \frac{\sqrt{6}}{18}\delta M^2 \left(1 + \frac{\Lambda_1}{2}\right)\left(1 - \frac{1}{\cos(2\theta)}\right) \\
& + \frac{\sqrt{6}}{6} M_0^2 \left(1 - \frac{\Lambda_1}{2}\right)\left(1 + \frac{1}{\cos(2\theta)}\right) \\
& + \frac{4\sqrt{6}}{9}\overline{M}^2\delta M^2 \frac{L_5}{F^2} - \frac{4\sqrt{6}}{3}\overline{M}^2 M_0^2 \frac{L_5}{F^2}\left(1 + \frac{1}{\cos(2\theta)}\right) \\
& + \frac{4\sqrt{6}}{3}\overline{M}^4 \frac{L_5}{F^2} \\
& + \frac{2\sqrt{6}}{3}\delta M^4 \left[\frac{2L_8 - L_5}{F^2}\left(1 - \frac{1}{3\cos(2\theta)}\right) - \frac{4}{9}\frac{L_5}{F^2}\right] \Bigg\} \\
+ \sin\theta F \Bigg\{ & - \frac{\sqrt{3}}{6}\overline{M}^2 \left(1 - \Lambda_1 + 2\Lambda_2 - 3\frac{M_{0^2}}{\delta M^2}(1 - \Lambda_1)\right)\left(1 + \frac{1}{\cos(2\theta)}\right) \\
& + \frac{2\sqrt{3}}{9}\delta M^2 \left[1 + \left(-\frac{\Lambda_1}{2} + \Lambda_2\right)\left(1 + \frac{1}{\cos(2\theta)}\right)\right] \\
& - \frac{4\sqrt{3}}{9}\overline{M}^2\delta M^2 \left[\frac{2L_8 - L_5}{F^2}\left(5 - \frac{1}{\cos(2\theta)}\right) + 3\frac{L_5}{F^2}\right] \\
& - \frac{2\sqrt{3}}{3}\overline{M}^2 M_0^2 \left(5\frac{2L_8 - L_5}{F^2} - 2\frac{L_5}{F^2}\right)\left(1 + \frac{1}{\cos(2\theta)}\right) \\
& + \frac{2\sqrt{3}}{9}\delta M^2 M_0^2 \left(7\frac{2L_8 - L_5}{F^2} - 4\frac{L_5}{F^2}\right)\left(1 + \frac{1}{\cos(2\theta)}\right) \\
& + \frac{4\sqrt{3}}{3}\overline{M}^4 \frac{2L_8 - L_5}{F^2}\left(1 - 3\frac{M_0^2}{\delta M^2}\right)\left(1 + \frac{1}{\cos(2\theta)}\right) \\
& + \frac{2\sqrt{3}}{27}\delta M^4 \left[\frac{2L_8 - L_5}{F^2}\left(13 - \frac{7}{\cos(2\theta)}\right) + 4\frac{L_5}{F^2}\right] \Bigg\}.
\end{aligned}
\tag{F.10}
$$

Note that a_η vanishes for $\theta = 0$, as it should.

The LO results can easily be obtained, setting $L_5 = L_8 = \Lambda_1 = \Lambda_2 = 0$. These only depend on the LECs F and M_0. The mass mixing angle θ at LO is given in eqs. (5.8) and (5.17)–(5.19) as a function of the δM^2 and M_0^2. Above, we use the NLO expression for θ, eqs. (5.8) and (5.21)–(5.23).

The parametrizations of the anomalous matrix elements eqs. (F.9) and (F.10) contain leading order terms $\propto F M_0^2$, $\propto F\overline{M}^2$ and $\propto F\delta M^2$. When taking the continuum limit using dimensionless combinations with $8t_0$ as discussed at the end of sec. 5.1.2, we need to add corresponding terms $\propto -(3k/2)FM_0^2\overline{M}^2 8t_0$, $\propto -(k/2)F\overline{M}^4 8t_0$ and $\propto -(k/2)F\delta M^2\overline{M}^2 8t_0$ which we do in our analysis of sec 6.2.

Bibliography

[1] Murray Gell-Mann. "The Eightfold Way: A Theory of strong interaction symmetry". In: (1961). DOI: 10.2172/4008239.

[2] Yuval Ne'eman. "Derivation of strong interactions from a gauge invariance". In: *Nucl. Phys.* 26 (1961). Ed. by R. Ruffini and Y. Verbin, pp. 222–229. DOI: 10.1016/0029-5582(61)90134-1.

[3] A. Pevsner et al. "Evidence for a Three Pion Resonance Near 550-MeV". In: *Phys. Rev. Lett.* 7 (1961), pp. 421–423. DOI: 10.1103/PhysRevLett.7.421.

[4] V. E. Barnes et al. "Confirmation of the existence of the Ω^- hyperon". In: *Phys. Lett.* 12 (1964), pp. 134–136. DOI: 10.1016/0031-9163(64)91137-0.

[5] Susumu Okubo. "Note on unitary symmetry in strong interactions". In: *Prog. Theor. Phys.* 27 (1962), pp. 949–966. DOI: 10.1143/PTP.27.949.

[6] S. Okubo. "Note on Unitary Symmetry in Strong Interaction. II Excited States of Baryons". In: *Prog. Theor. Phys.* 28 (1962), pp. 24–32. DOI: 10.1143/PTP.28.24.

[7] George R. Kalbfleisch et al. "Observation of a Nonstrange Meson of Mass 959 MeV". In: *Phys. Rev. Lett.* 12 (1964), pp. 527–530. DOI: 10.1103/PhysRevLett.12.527.

[8] Stephen L. Adler. "Axial vector vertex in spinor electrodynamics". In: *Phys. Rev.* 177 (1969), pp. 2426–2438. DOI: 10.1103/PhysRev.177.2426.

[9] J.S. Bell and R. Jackiw. "A PCAC puzzle: $\pi^0 \to \gamma\gamma$ in the σ model". In: *Nuovo Cim. A* 60 (1969), pp. 47–61. DOI: 10.1007/BF02823296.

[10] Alexander M. Polyakov. "Compact Gauge Fields and the Infrared Catastrophe". In: *Phys. Lett. B* 59 (1975). Ed. by J. C. Taylor, pp. 82–84. DOI: 10.1016/0370-2693(75)90162-8.

[11] A. A. Belavin et al. "Pseudoparticle Solutions of the Yang-Mills Equations". In: *Phys. Lett. B* 59 (1975). Ed. by J. C. Taylor, pp. 85–87. DOI: 10.1016/0370-2693(75)90163-X.

[12] Edward Witten. "Instantons, the quark model, and the $1/N$ expansion". In: *Nucl. Phys. B* 149 (1979), pp. 285–320. DOI: 10.1016/0550-3213(79)90243-8.

[13] Edward Witten. "Current Algebra Theorems for the U(1) Goldstone Boson". In: *Nucl. Phys. B* 156 (1979), pp. 269–283. DOI: 10.1016/0550-3213(79)90031-2.

[14] G. Veneziano. "U(1) without instantons". In: *Nucl. Phys. B* 159 (1979), pp. 213–224. DOI: 10.1016/0550-3213(79)90332-8.

[15] A. Anastasi et al. "Precision measurement of the $\eta \to \pi^+\pi^-\pi^0$ Dalitz plot distribution with the KLOE detector". In: *JHEP* 05 (2016), p. 019. DOI: 10.1007/JHEP05(2016)019. arXiv: 1601.06985 [hep-ex].

[16] Liping Gan et al. "Precision tests of fundamental physics with η and η' mesons". In: (July 2020). arXiv: 2007.00664 [hep-ph].

[17] G. W. Bennett et al. "Final Report of the Muon E821 Anomalous Magnetic Moment Measurement at BNL". In: *Phys. Rev. D* 73 (2006), p. 072003. DOI: 10.1103/PhysRevD.73.072003. arXiv: hep-ex/0602035.

[18] B. Abi et al. "Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm". In: *Phys. Rev. Lett.* 126.14 (2021), p. 141801. DOI: 10.1103/PhysRevLett.126.141801. arXiv: 2104.03281 [hep-ex].

[19] P. A. Zyla et al. "Review of Particle Physics". In: *PTEP* 2020.8 (2020), p. 083C01. DOI: 10.1093/ptep/ptaa104.

[20] Sz. Borsanyi et al. "Leading hadronic contribution to the muon magnetic moment from lattice QCD". In: *Nature* 593.7857 (2021), pp. 51–55. DOI: 10.1038/s41586-021-03418-1. arXiv: 2002.12347 [hep-lat].

[21] Bernard Aubert et al. "Measurement of the η and η' transition form-factors at $q^2 = 112$-GeV^2". In: *Phys. Rev. D* 74 (2006), p. 012002. DOI: 10.1103/PhysRevD.74.012002. arXiv: hep-ex/0605018.

[22] P. del Amo Sanchez et al. "Measurement of the $\gamma\gamma^* \to \eta$ and $\gamma\gamma^* \to \eta'$ transition form factors". In: *Phys. Rev. D* 84 (2011), p. 052001. DOI: 10.1103/PhysRevD.84.052001. arXiv: 1101.1142 [hep-ex].

[23] A. J. Bevan et al. "The Physics of the B Factories". In: *Eur. Phys. J. C* 74 (2014), p. 3026. DOI: 10.1140/epjc/s10052-014-3026-9. arXiv: 1406.6311 [hep-ex].

[24] Xiaolin Kang. "Study of η-meson Decays at KLOE/KLOE-2". In: *J. Phys. Conf. Ser.* 1586.1 (2020). Ed. by Eberhard Widmann et al., p. 012035. DOI: 10.1088/1742-6596/1586/1/012035. arXiv: 1904.12034 [hep-ex].

[25] Yi Fang. "Charm semileptonic physics at BESIII". In: *PoS* FPCP2017 (2017), p. 001. DOI: 10.22323/1.304.0001.

[26] Tomoyuki Konno. "Status and prospects of the Belle II experiment". In: *J. Phys. Conf. Ser.* 627.1 (2015). Ed. by Trinh Xuan Hoang et al., p. 012009. DOI: 10.1088/1742-6596/627/1/012009.

[27] W. Altmannshofer et al. "The Belle II Physics Book". In: *PTEP* 2019.12 (2019). Ed. by E. Kou and P. Urquijo. [Erratum: PTEP 2020, 029201 (2020)], p. 123C01. DOI: 10.1093/ptep/ptz106. arXiv: 1808.10567 [hep-ex].

[28] M. Dugger et al. *Eta Decays with Emphasis on Rare Neutral Modes: The JLab Eta Factory (JEF) Experiment*. Tech. rep. PR12-14-004. Jefferson Lab, 2014. URL: www.jlab.org/exp_prog/proposals/14/PR12-14-004.pdf.

[29] P. Di Vecchia and G. Veneziano. "Chiral dynamics in the large N limit". In: *Nucl. Phys. B* 171 (1980), pp. 253–272. DOI: 10.1016/0550-3213(80)90370-3.

[30] Ken Kawarabayashi and Nobuyoshi Ohta. "The problem of η in the large N limit: Effective Lagrangian approach". In: *Nucl. Phys. B* 175 (1980), pp. 477–492. DOI: 10.1016/0550-3213(80)90024-3.

[31] P. Di Vecchia et al. "Large n, chiral approach to pseudoscalar masses, mixings and decays". In: *Nucl. Phys. B* 181 (1981), pp. 318–334. DOI: 10.1016/0550-3213(81)90356-4.

[32] H. Leutwyler. "On the $1/N$ expansion in chiral perturbation theory". In: *Nucl. Phys. B Proc. Suppl.* 64 (1998). Ed. by Stephan Narison, pp. 223–231. DOI: 10.1016/S0920-5632(97)01065-7. arXiv: hep-ph/9709408.

[33] Roland Kaiser and H. Leutwyler. "Large N_c in chiral perturbation theory". In: *Eur. Phys. J. C* 17 (2000), pp. 623–649. DOI: 10.1007/s100520000499. arXiv: hep-ph/0007101 [hep-ph].

[34] Y. Kuramashi et al. "η' meson mass in lattice QCD". In: *Phys. Rev. Lett.* 72 (1994), pp. 3448–3451. DOI: 10.1103/PhysRevLett.72.3448.

[35] L. Venkataraman and G. Kilcup. "The η' meson with staggered fermions". In: (Nov. 1997). arXiv: hep-lat/9711006.

[36] William A. Bardeen et al. "Anomalous chiral behavior in quenched lattice QCD". In: *Phys. Rev. D* 62 (2000), p. 114505. DOI: 10.1103/PhysRevD.62.114505. arXiv: hep-lat/0007010.

[37] T. Struckmann et al. "Flavor singlet pseudoscalar masses in $N_f = 2$ QCD". In: *Phys. Rev. D* 63 (2001), p. 074503. DOI: 10.1103/PhysRevD.63.074503. arXiv: hep-lat/0010005 [hep-lat].

[38] C. McNeile and Christopher Michael. "The η and η' mesons in QCD". In: *Phys. Lett. B* 491 (2000). [Erratum: Phys. Lett.B552,280(2003)], pp. 123–129. DOI: 10.1016/S0370-2693(02)03005-8, 10.1016/S0370-2693(00)01010-8. arXiv: hep-lat/0006020 [hep-lat].

[39] Gunnar S. Bali et al. "Quark mass effects on the topological susceptibility in QCD". In: *Phys. Rev. D* 64 (2001), p. 054502. DOI: 10.1103/PhysRevD.64.054502. arXiv: hep-lat/0102002.

[40] V. I. Lesk et al. "Flavor singlet meson mass in the continuum limit in two flavor lattice QCD". In: *Phys. Rev. D* 67 (2003), p. 074503. DOI: 10.1103/PhysRevD.67.074503. arXiv: hep-lat/0211040 [hep-lat].

[41] Koichi Hashimoto and Taku Izubuchi. "η' Meson from Two Flavor Dynamical Domain Wall Fermions". In: *Prog. Theor. Phys.* 119 (2008), pp. 599–641. DOI: 10.1143/PTP.119.599. arXiv: 0803.0186 [hep-lat].

[42] K. Jansen, Christopher Michael, and C. Urbach. "The η' meson from lattice QCD". In: *Eur. Phys. J. C* 58 (2008), pp. 261–269. DOI: 10.1140/epjc/s10052-008-0764-6. arXiv: 0804.3871 [hep-lat].

[43] Wei Sun et al. "Glueball spectrum from $N_f = 2$ lattice QCD study on anisotropic lattices". In: *Chin. Phys. C* 42.9 (2018), p. 093103. DOI: 10.1088/1674-1137/42/9/093103. arXiv: 1702.08174 [hep-lat].

[44] Petros Dimopoulos et al. "Topological susceptibility and η' meson mass from $N_f = 2$ lattice QCD at the physical point". In: *Phys. Rev. D* 99.3 (2019), p. 034511. DOI: 10.1103/PhysRevD.99.034511. arXiv: 1812.08787 [hep-lat].

[45] N. H. Christ et al. "The η and η' mesons from Lattice QCD". In: *Phys. Rev. Lett.* 105 (2010), p. 241601. DOI: 10.1103/PhysRevLett.105.241601. arXiv: 1002.2999 [hep-lat].

[46] Jozef J. Dudek et al. "Isoscalar meson spectroscopy from lattice QCD". In: *Phys. Rev. D* 83 (2011), p. 111502. DOI: 10.1103/PhysRevD.83.111502. arXiv: 1102.4299 [hep-lat].

[47] Eric B. Gregory et al. "A study of the η and η' mesons with improved staggered fermions". In: *Phys. Rev. D* 86 (2012), p. 014504. DOI: 10.1103/PhysRevD.86.014504. arXiv: 1112.4384 [hep-lat].

[48] Gunnar S. Bali et al. "$D_s \to \eta, \eta'$ semileptonic decay form factors with disconnected quark loop contributions". In: *Phys. Rev. D* 91.1 (2015), p. 014503. DOI: 10.1103/PhysRevD.91.014503. arXiv: 1406.5449 [hep-lat].

[49] H. Fukaya et al. "η' meson mass from topological charge density correlator in QCD". In: *Phys. Rev. D* 92.11 (2015), p. 111501. DOI: 10.1103/PhysRevD.92.111501. arXiv: 1509.00944 [hep-lat].

[50] Konstantin Ottnad et al. "η and η' mesons from $N_f = 2 + 1 + 1$ twisted mass lattice QCD". In: *JHEP* 11 (2012), p. 048. DOI: 10.1007/JHEP11(2012)048. arXiv: 1206.6719 [hep-lat].

[51] Chris Michael, Konstantin Ottnad, and Carsten Urbach. "η and η' mixing from Lattice QCD". In: *Phys. Rev. Lett.* 111.18 (2013), p. 181602. DOI: 10.1103/PhysRevLett.111.181602. arXiv: 1310.1207 [hep-lat].

[52] Konstantin Ottnad and Carsten Urbach. "Flavor-singlet meson decay constants from $N_f = 2 + 1 + 1$ twisted mass lattice QCD". In: *Phys. Rev. D* 97.5 (2018), p. 054508. DOI: 10.1103/PhysRevD.97.054508. arXiv: 1710.07986 [hep-lat].

[53] Andrey Yu. Kotov, Maria Paola Lombardo, and Anton M. Trunin. "Fate of the η' in the quark gluon plasma". In: *Phys. Lett. B* 794 (2019), pp. 83–88. DOI: 10.1016/j.physletb.2019.05.035. arXiv: 1903.05633 [hep-lat].

[54] Z. R. Kordov et al. "State mixing and masses of the π0, η and η' mesons from nf=1+1+1 lattice QCD+QED". In: *Phys. Rev. D* 104.11 (2021), p. 114514. DOI: 10.1103/PhysRevD.104.114514. arXiv: 2110.11533 [hep-lat].

[55] T. P. Cheng and L. F. Li. *Gauge theory of elementary particle physics*. 1984. ISBN: 9780198519614.

[56] M.E. Peskin and D.V. Schroeder. *An Introduction To Quantum Field Theory*. Frontiers in physics. Westview Press, 1995. ISBN: 9780813345437.

[57] Heinz J. Rothe. *Lattice Gauge Theories : An Introduction (Fourth Edition)*. Vol. 43. World Scientific Publishing Company, 2012. ISBN: 9789814365857. DOI: 10.1142/8229.

[58] J. Smit. *Introduction to quantum fields on a lattice: A robust mate*. Vol. 15. Cambridge University Press, Jan. 2011. ISBN: 9780521890519.

[59] C. Gattringer and C.B. Lang. *Quantum Chromodynamics on the Lattice: An Introductory Presentation*. Lecture Notes in Physics. Springer, 2009. ISBN: 9783642018497.

[60] Gunnar S. Bali et al. "Masses and decay constants of the η and η' mesons from lattice QCD". In: *JHEP* 08 (2021), p. 137. DOI: 10.1007/JHEP08(2021)137. arXiv: 2106.05398 [hep-lat].

[61] C. A. Baker et al. "An Improved experimental limit on the electric dipole moment of the neutron". In: *Phys. Rev. Lett.* 97 (2006), p. 131801. DOI: 10.1103/PhysRevLett.97.131801. arXiv: hep-ex/0602020.

[62] Mattia Bruno et al. "Effects of Heavy Sea Quarks at Low Energies". In: *Phys. Rev. Lett.* 114.10 (2015), p. 102001. DOI: 10.1103/PhysRevLett.114.102001. arXiv: 1410.8374 [hep-lat].

[63] M. Bruno et al. "The Λ-parameter in 3-flavour QCD and $\alpha_s(m_Z)$ by the ALPHA collaboration". In: *PoS* LATTICE2016 (2016), p. 197. DOI: 10.22323/1.256.0197. arXiv: 1701.03075 [hep-lat].

[64] D. J. Gross and F. Wilczek. "Ultraviolet Behavior of Nonabelian Gauge Theories". In: *Phys.Rev.Lett.* 30 (1973), pp. 1343–1346. DOI: 10.1103/PhysRevLett.30.1343.

[65] H. D. Politzer. "Reliable Perturbative Results for Strong Interactions?" In: *Phys.Rev.Lett.* 30 (1973), pp. 1346–1349. DOI: 10.1103/PhysRevLett.30.1346.

[66] A. Zee. "Study of the renormalization group for small coupling constants". In: *Phys. Rev. D* 7 (1973), pp. 3630–3636. DOI: 10.1103/PhysRevD.7.3630.

[67] Sidney R. Coleman and David J. Gross. "Price of asymptotic freedom". In: *Phys. Rev. Lett.* 31 (1973), pp. 851–854. DOI: 10.1103/PhysRevLett.31.851.

[68] Nathan Isgur and Jack E. Paton. "A Flux Tube Model for Hadrons in QCD". In: *Phys. Rev. D* 31 (1985), p. 2910. DOI: 10.1103/PhysRevD.31.2910.

[69] Gunnar S. Bali et al. "Observation of string breaking in QCD". In: *Phys. Rev. D* 71 (2005), p. 114513. DOI: 10.1103/PhysRevD.71.114513. arXiv: hep-lat/0505012.

[70] Roel Aaij et al. "Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays". In: *Phys. Rev. Lett.* 115 (2015), p. 072001. DOI: 10.1103/PhysRevLett.115.072001. arXiv: 1507.03414 [hep-ex].

[71] Roel Aaij et al. "Observation of $J/\psi\phi$ structures consistent with exotic states from amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays". In: *Phys. Rev. Lett.* 118.2 (2017), p. 022003. DOI: 10.1103/PhysRevLett.118.022003. arXiv: 1606.07895 [hep-ex].

[72] Christopher Michael and M. Teper. "The Glueball Spectrum and Scaling in SU(3) Lattice Gauge Theory". In: *Phys. Lett. B* 206 (1988), pp. 299–308. DOI: 10.1016/0370-2693(88)91510-9.

[73] G. S. Bali et al. "A Comprehensive lattice study of SU(3) glueballs". In: *Phys. Lett. B* 309 (1993), pp. 378–384. DOI: 10.1016/0370-2693(93)90948-H. arXiv: hep-lat/9304012.

[74] Colin J. Morningstar and Mike J. Peardon. "The Glueball spectrum from an anisotropic lattice study". In: *Phys. Rev. D* 60 (1999), p. 034509. DOI: 10.1103/PhysRevD.60.034509. arXiv: hep-lat/9901004.

[75] Wolfgang Ochs. "The Status of Glueballs". In: *J. Phys. G* 40 (2013), p. 043001. DOI: 10.1088/0954-3899/40/4/043001. arXiv: 1301.5183 [hep-ph].

[76] T. Csörgő et al. "Evidence of Odderon-exchange from scaling properties of elastic scattering at TeV energies". In: *Eur. Phys. J. C* 81.2 (2021), p. 180. DOI: 10.1140/epjc/s10052-021-08867-6. arXiv: 1912.11968 [hep-ph].

[77] Jiro Kodaira. "QCD Higher Order Effects in Polarized Electroproduction: Flavor Singlet Coefficient Functions". In: *Nucl. Phys. B* 165 (1980), pp. 129–140. DOI: 10.1016/0550-3213(80)90310-7.

[78] Stephen L. Adler and William A. Bardeen. "Absence of higher order corrections in the anomalous axial vector divergence equation". In: *Phys. Rev.* 182 (1969), pp. 1517–1536. DOI: 10.1103/PhysRev.182.1517.

[79] B. L. Ioffe. "Axial anomaly: The Modern status". In: *Int. J. Mod. Phys. A* 21 (2006), pp. 6249–6266. DOI: 10.1142/S0217751X06035051. arXiv: hep-ph/0611026.

[80] Murray Gell-Mann. "A Schematic Model of Baryons and Mesons". In: *Phys. Lett.* 8 (1964), pp. 214–215. DOI: 10.1016/S0031-9163(64)92001-3.

[81] K. G. Wilson. "Confinement of Quarks". In: *Phys.Rev.* D10 (1974), pp. 2445–2459. DOI: 10.1103/PhysRevD.10.2445.

[82] P. A.M. Dirac. "The Lagrangian in quantum mechanics". In: *Phys.Z.Sowjetunion* 3 (1933), pp. 64–72.

[83] R. P. Feynman. "The principle of least action in quantum mechanics". In: (1942).

[84] M. Lüscher and P. Weisz. "On-Shell Improved Lattice Gauge Theories". In: *Commun. Math. Phys.* 97 (1985). [Erratum: Commun.Math.Phys. 98, 433 (1985)], p. 59. DOI: 10.1007/BF01206178.

[85] P. Weisz. "Continuum Limit Improved Lattice Action for Pure Yang-Mills Theory. I." In: *Nucl. Phys. B* 212 (1983), pp. 1–17. DOI: 10.1016/0550-3213(83)90595-3.

[86] Y. Iwasaki. "Renormalization Group Analysis of Lattice Theories and Improved Lattice Action. II. Four-dimensional non-Abelian SU(N) gauge model". In: (Dec. 1983). arXiv: 1111.7054 [hep-lat].

[87] B. Sheikholeslami and R. Wohlert. "Improved Continuum Limit Lattice Action for QCD with Wilson Fermions". In: *Nucl. Phys. B* 259 (1985), p. 572. DOI: 10.1016/0550-3213(85)90002-1.

[88] John Bulava and Stefan Schaefer. "Improvement of $N_f = 3$ lattice QCD with Wilson fermions and tree-level improved gauge action". In: *Nucl. Phys. B* 874 (2013), pp. 188–197. DOI: 10.1016/j.nuclphysb.2013.05.019. arXiv: 1304.7093 [hep-lat].

[89] Gunnar S. Bali et al. "Lattice simulations with $N_f = 2 + 1$ improved Wilson fermions at a fixed strange quark mass". In: *Phys. Rev. D* 94.7 (2016), p. 074501. DOI: 10.1103/PhysRevD.94.074501. arXiv: 1606.09039 [hep-lat].

[90] Gunnar S. Bali et al. "Scale setting and the light hadron spectrum in $N_f = 2 + 1$ QCD with Wilson fermions". In: (in preparation).

[91] Holger Bech Nielsen and M. Ninomiya. "No Go Theorem for Regularizing Chiral Fermions". In: *Phys. Lett. B* 105 (1981), pp. 219–223. DOI: 10.1016/0370-2693(81)91026-1.

[92] Luuk H. Karsten and Jan Smit. "Lattice Fermions: Species Doubling, Chiral Invariance, and the Triangle Anomaly". In: *Nucl. Phys. B* 183 (1981). Ed. by J. Julve and M. Ramón-Medrano, p. 103. DOI: 10.1016/0550-3213(81)90549-6.

[93] Paul H. Ginsparg and Kenneth G. Wilson. "A Remnant of Chiral Symmetry on the Lattice". In: *Phys. Rev. D* 25 (1982), p. 2649. DOI: 10.1103/PhysRevD.25.2649.

[94] Herbert Neuberger. "Exactly massless quarks on the lattice". In: *Phys. Lett. B* 417 (1998), pp. 141–144. DOI: 10.1016/S0370-2693(97)01368-3. arXiv: hep-lat/9707022.

[95] David B. Kaplan. "A Method for simulating chiral fermions on the lattice". In: *Phys. Lett. B* 288 (1992), pp. 342–347. DOI: 10.1016/0370-2693(92)91112-M. arXiv: hep-lat/9206013.

[96] Martin Lüscher. "Computational Strategies in Lattice QCD". In: *Les Houches Summer School: Session 93: Modern perspectives in lattice QCD: Quantum field theory and high performance computing*. Feb. 2010, pp. 331–399. arXiv: 1002.4232 [hep-lat].

[97] *openQCD — Simulation programs for lattice QCD*. Tech. rep. URL: https://luscher.web.cern.ch/luscher/openQCD/.

[98] M. Creutz. "Monte Carlo Study of Quantized SU(2) Gauge Theory". In: *Phys. Rev. D* 21 (1980), pp. 2308–2315. DOI: 10.1103/PhysRevD.21.2308.

[99] P. T. Matthews and Abdus Salam. "The Green's functions of quantized fields". In: *Nuovo Cim.* 12 (1954), pp. 563–565. DOI: 10.1007/BF02781302.

[100] P. T. Matthews and Abdus Salam. "Propagators of quantized field". In: *Nuovo Cim.* 2 (1955), pp. 120–134. DOI: 10.1007/BF02856011.

[101] S. Duane et al. "Hybrid Monte Carlo". In: *Phys. Lett. B* 195 (1987), pp. 216–222. DOI: 10.1016/0370-2693(87)91197-X.

[102] J. C. Sexton and D. H. Weingarten. "Hamiltonian evolution for the hybrid Monte Carlo algorithm". In: *Nucl. Phys. B* 380 (1992), pp. 665–677. DOI: 10.1016/0550-3213(92)90263-B.

[103] M. Hasenbusch and K. Jansen. "Speeding up lattice QCD simulations with clover improved Wilson fermions". In: *Nucl. Phys. B* 659 (2003), pp. 299–320. DOI: 10.1016/S0550-3213(03)00227-X. arXiv: hep-lat/0211042.

[104] Martin Lüscher and Filippo Palombi. "Fluctuations and reweighting of the quark determinant on large lattices". In: *PoS* LATTICE2008 (2008). Ed. by Christopher Aubin et al., p. 049. DOI: 10.22323/1.066.0049. arXiv: 0810.0946 [hep-lat].

[105] Martin Lüscher. "Local coherence and deflation of the low quark modes in lattice QCD". In: *JHEP* 07 (2007), p. 081. DOI: 10.1088/1126-6708/2007/07/081. arXiv: 0706.2298 [hep-lat].

[106] R. Babich et al. "Adaptive multigrid algorithm for the lattice Wilson-Dirac operator". In: *Phys. Rev. Lett.* 105 (2010), p. 201602. DOI: 10.1103/PhysRevLett.105.201602. arXiv: 1005.3043 [hep-lat].

[107] Andreas Frommer et al. "An Adaptive Aggregation Based Domain Decomposition Multilevel Method for the Lattice Wilson Dirac Operator: Multilevel Results". In: *arXiv:1307.6101* (2013). arXiv: 1307.6101 [hep-lat].

[108] Martin Lüscher. "Schwarz-preconditioned HMC algorithm for two-flavour lattice QCD". In: *Comput. Phys. Commun.* 165 (2005), pp. 199–220. DOI: 10.1016/j.cpc.2004.10.004. arXiv: hep-lat/0409106.

[109] Martin Lüscher. "Deflation acceleration of lattice QCD simulations". In: *JHEP* 12 (2007), p. 011. DOI: 10.1088/1126-6708/2007/12/011. arXiv: 0710.5417 [hep-lat].

[110] Daniel Mohler and Stefan Schaefer. "Remarks on strange-quark simulations with Wilson fermions". In: *Phys. Rev. D* 102.7 (2020), p. 074506. DOI: 10.1103/PhysRevD.102.074506. arXiv: 2003.13359 [hep-lat].

[111] S. Güsken et al. "Nonsinglet axial vector couplings of the baryon octet in lattice QCD". In: *Phys. Lett. B* 227 (1989), pp. 266–269. DOI: 10.1016/S0370-2693(89)80034-6.

[112] M. Falcioni et al. "Again on SU(3) glueball mass". In: *Nucl. Phys. B* 251 (1985), pp. 624–632. DOI: 10.1016/0550-3213(85)90280-9.

[113] M. Lüscher. "Construction of a Selfadjoint, Strictly Positive Transfer Matrix for Euclidean Lattice Gauge Theories". In: *Commun. Math. Phys.* 54 (1977), p. 283. DOI: 10.1007/BF01614090.

[114] J. Schechter, A. Subbaraman, and H. Weigel. "Effective hadron dynamics: From meson masses to the proton spin puzzle". In: *Phys. Rev. D* 48 (1993), pp. 339–355. DOI: 10.1103/PhysRevD.48.339. arXiv: hep-ph/9211239.

[115] T. Feldmann, P. Kroll, and B. Stech. "Mixing and decay constants of pseudoscalar mesons". In: *Phys. Rev. D* 58 (1998), p. 114006. DOI: 10.1103/PhysRevD.58.114006. arXiv: hep-ph/9802409 [hep-ph].

[116] Karl Jansen et al. "Nonperturbative renormalization of lattice QCD at all scales". In: *Phys. Lett. B* 372 (1996), p. 275. DOI: 10.1016/0370-2693(96)00075-5. arXiv: hep-lat/9512009 [hep-lat].

[117] Martin Lüscher et al. "Chiral symmetry and $O(a)$ improvement in Lattice QCD". In: *Nucl. Phys. B* 478 (1996), p. 365. DOI: 10.1016/0550-3213(96)00378-1. arXiv: hep-lat/9605038 [hep-lat].

[118] Tanmoy Bhattacharya et al. "Improved bilinears in lattice QCD with non-degenerate quarks". In: *Phys. Rev. D* 73 (2006), p. 034504. DOI: 10.1103/PhysRevD.73.034504. arXiv: hep-lat/0511014 [hep-lat].

[119] John Bulava et al. "Non-perturbative improvement of the axial current in N_f=3 lattice QCD with Wilson fermions and tree-level improved gauge action". In: *Nucl. Phys. B* 896 (2015), pp. 555–568. DOI: 10.1016/j.nuclphysb.2015.05.003. arXiv: 1502.04999 [hep-lat].

[120] Piotr Korcyl and Gunnar S. Bali. "Non-perturbative determination of improvement coefficients using coordinate space correlators in $N_f = 2 + 1$ lattice QCD". In: *Phys. Rev. D* 95.1 (2017), p. 014505. DOI: 10.1103/PhysRevD.95.014505. arXiv: 1607.07090 [hep-lat].

[121] Gunnar S. Bali et al. "Non-perturbative determination of quark-mass independent improvement coefficients in $n_f = 2 + 1$ lattice QCD". In: (in preparation).

[122] Antoine Gérardin, Tim Harris, and Harvey B. Meyer. "Nonperturbative renormalization and $O(a)$-improvement of the nonsinglet vector current with $N_f = 2 + 1$ Wilson fermions and tree-level Symanzik improved gauge action". In: *Phys. Rev. D* 99.1 (2019), p. 014519. DOI: 10.1103/PhysRevD.99.014519. arXiv: 1811.08209 [hep-lat].

[123] Mattia Dalla Brida et al. "High precision renormalization of the flavour non-singlet Noether currents in lattice QCD with Wilson quarks". In: *Eur. Phys. J. C* 79.1 (2019), p. 23. DOI: 10.1140/epjc/s10052-018-6514-5. arXiv: 1808.09236 [hep-lat].

[124] S. A. Larin. "The Renormalization of the axial anomaly in dimensional regularization". In: *Phys. Lett. B* 303 (1993), pp. 113–118. DOI: 10.1016/0370-2693(93)90053-K. arXiv: hep-ph/9302240 [hep-ph].

[125] M. Constantinou et al. "Singlet versus nonsinglet perturbative renormalization of fermion bilinears". In: *Phys. Rev. D* 94.11 (2016), p. 114513. DOI: 10.1103/PhysRevD.94.114513. arXiv: 1610.06744 [hep-lat].

[126] Taushif Ahmed, Long Chen, and Michał Czakon. "Renormalization of the flavor-singlet axial-vector current and its anomaly in dimensional regularization". In: *JHEP* 05 (2021), p. 087. DOI: 10.1007/JHEP05(2021)087. arXiv: 2101.09479 [hep-ph].

[127] M. F. Zoller. "OPE of the pseudoscalar gluonium correlator in massless QCD to three-loop order". In: *JHEP* 07 (2013), p. 040. DOI: 10.1007/JHEP07(2013)040. arXiv: 1304.2232 [hep-ph].

[128] Mattia Bruno et al. "QCD coupling from a nonperturbative determination of the three-flavor Λ parameter". In: *Phys. Rev. Lett.* 119.10 (2017), p. 102001. DOI: 10.1103/PhysRevLett.119.102001. arXiv: 1706.03821 [hep-lat].

[129] P. A. Baikov, K. G. Chetyrkin, and J. H. Kühn. "Quark mass and field anomalous dimensions to $\mathcal{O}(\alpha_s^5)$". In: *JHEP* 10 (2014), p. 076. DOI: 10.1007/JHEP10(2014)076. arXiv: 1402.6611 [hep-ph].

[130] Florian Herren and Matthias Steinhauser. "Version 3 of RunDec and CRunDec". In: *Comput. Phys. Commun.* 224 (2018), pp. 333–345. DOI: 10.1016/j.cpc.2017.11.014. arXiv: 1703.03751 [hep-ph].

[131] K. G. Chetyrkin, Johann H. Kühn, and M. Steinhauser. "RunDec: A Mathematica package for running and decoupling of the strong coupling and quark masses". In: *Comput. Phys. Commun.* 133 (2000), pp. 43–65. DOI: 10.1016/S0010-4655(00)00155-7. arXiv: hep-ph/0004189 [hep-ph].

[132] Martin Lüscher and Stefan Schaefer. "Lattice QCD with open boundary conditions and twisted-mass reweighting". In: *Comput. Phys. Commun.* 184 (2013), pp. 519–528. DOI: 10.1016/j.cpc.2012.10.003. arXiv: 1206.2809 [hep-lat].

[133] Mattia Bruno et al. "Simulation of QCD with $N_f = 2 + 1$ flavors of non-perturbatively improved Wilson fermions". In: *JHEP* 02 (2015), p. 043. DOI: 10.1007/JHEP02(2015)043. arXiv: 1411.3982 [hep-lat].

[134] Daniel Mohler, Stefan Schaefer, and Jakob Simeth. "CLS 2+1 flavor simulations at physical light- and strange-quark masses". In: *EPJ Web Conf.* 175 (2018), p. 02010. DOI: 10.1051/epjconf/201817502010. arXiv: 1712.04884 [hep-lat].

[135] Mattia Bruno, Tomasz Korzec, and Stefan Schaefer. "Setting the scale for the CLS $2 + 1$ flavor ensembles". In: *Phys. Rev. D* 95.7 (2017), p. 074504. DOI: 10.1103/PhysRevD.95.074504. arXiv: 1608.08900 [hep-lat].

[136] S. Aoki et al. "Review of lattice results concerning low-energy particle physics". In: *Eur. Phys. J. C* 77.2 (2017), p. 112. DOI: 10.1140/epjc/s10052-016-4509-7. arXiv: 1607.00299 [hep-lat].

[137] Wolfgang Bietenholz et al. "Tuning the strange quark mass in lattice simulations". In: *Phys. Lett. B* 690 (2010), p. 436. DOI: 10.1016/j.physletb.2010.05.067. arXiv: 1003.1114 [hep-lat].

[138] Christian Hoelbling. "Lattice QCD: concepts, techniques and some results". In: *Acta Phys. Polon. B* 45.12 (2014). Ed. by Michal Praszalowicz, p. 2143. DOI: 10.5506/APhysPolB.45.2143. arXiv: 1410.3403 [hep-lat].

[139] Jacob Finkenrath, Francesco Knechtli, and Björn Leder. "One flavor mass reweighting in lattice QCD". In: *Nucl. Phys. B* 877 (2013). [Erratum: Nucl.Phys.B 880, 574–575 (2014)], pp. 441–456. DOI: 10.1016/j.nuclphysb.2013.10.019. arXiv: 1306.3962 [hep-lat].

[140] Björn Leder and Jacob Finkenrath. "Tuning of the strange quark mass with optimal reweighting". In: *PoS* LATTICE2014 (2015), p. 040. DOI: 10.22323/1.214.0040. arXiv: 1501.06617 [hep-lat].

[141] Gilberto Colangelo, Stephan Durr, and Christoph Haefeli. "Finite volume effects for meson masses and decay constants". In: *Nucl. Phys. B* 721 (2005), pp. 136–174. DOI: 10.1016/j.nuclphysb.2005.05.015. arXiv: hep-lat/0503014.

[142] Stefan Schaefer, Rainer Sommer, and Francesco Virotta. "Critical slowing down and error analysis in lattice QCD simulations". In: *Nucl. Phys. B* 845 (2011), pp. 93–119. DOI: 10.1016/j.nuclphysb.2010.11.020. arXiv: 1009.5228 [hep-lat].

[143] Martin Lüscher and Stefan Schaefer. "Lattice QCD without topology barriers". In: *JHEP* 07 (2011), p. 036. DOI: 10.1007/JHEP07(2011)036. arXiv: 1105.4749 [hep-lat].

[144] Martin Lüscher. "Properties and uses of the Wilson flow in lattice QCD". In: *JHEP* 08 (2010). [Erratum: JHEP03,092(2014)], p. 071. DOI: 10.1007/JHEP08(2010)071,10.1007/JHEP03(2014)092. arXiv: 1006.4518 [hep-lat].

[145] Oliver Bär and Maarten Golterman. "Chiral perturbation theory for gradient flow observables". In: *Phys. Rev. D* 89.3 (2014). [Erratum: Phys.Rev.D 89, 099905 (2014)], p. 034505. DOI: 10.1103/PhysRevD.89.034505. arXiv: 1312.4999 [hep-lat].

[146] Szabolcs Borsanyi et al. "High-precision scale setting in lattice QCD". In: *JHEP* 09 (2012), p. 010. DOI: 10.1007/JHEP09(2012)010. arXiv: 1203.4469 [hep-lat].

[147] T. Blum et al. "Domain wall QCD with physical quark masses". In: *Phys. Rev. D* 93.7 (2016), p. 074505. DOI: 10.1103/PhysRevD.93.074505. arXiv: 1411.7017 [hep-lat].

[148] V. G. Bornyakov et al. "Wilson flow and scale setting from lattice QCD". In: (Aug. 2015). arXiv: 1508.05916 [hep-lat].

[149] Ben Straßberger et al. "Scale Setting for CLS 2+1 Simulations". In: *38th International Symposium on Lattice Field Theory*. Dec. 2021. arXiv: 2112.06696 [hep-lat].

[150] Gunnar S. Bali et al. "Lattice simulations with $N_f = 2 + 1$ improved Wilson fermions at a fixed strange quark mass". In: *Phys. Rev.* D94.7 (2016), p. 074501. DOI: 10.1103/PhysRevD.94.074501. arXiv: 1606.09039 [hep-lat].

[151] Apostolos Skouroupathis and Haralambos Panagopoulos. "Λ-parameter of lattice QCD with Symanzik improved gluon actions". In: *Phys. Rev. D* 76 (2007), p. 114514. DOI: 10.1103/PhysRevD.76.114514. arXiv: 0709.3239 [hep-lat].

[152] S. Aoki et al. "FLAG Review 2019". In: *Eur. Phys. J. C* 80.2 (2020), p. 113. DOI: 10.1140/epjc/s10052-019-7354-7. arXiv: 1902.08191 [hep-lat].

[153] A. Bazavov et al. "Results for light pseudoscalar mesons". In: *PoS* LATTICE2010 (2010). Ed. by Giancarlo Rossi, p. 074. DOI: 10.22323/1.105.0074. arXiv: 1012.0868 [hep-lat].

[154] S. R. Beane et al. "SU(2) Low-Energy Constants from Mixed-Action Lattice QCD". In: *Phys. Rev. D* 86 (2012), p. 094509. DOI: 10.1103/PhysRevD.86.094509. arXiv: 1108.1380 [hep-lat].

[155] Szabolcs Borsanyi et al. "SU(2) chiral perturbation theory low-energy constants from 2+1 flavor staggered lattice simulations". In: *Phys. Rev. D* 88 (2013), p. 014513. DOI: 10.1103/PhysRevD.88.014513. arXiv: 1205.0788 [hep-lat].

[156] Stephan Dürr et al. "Lattice QCD at the physical point meets SU(2) chiral perturbation theory". In: *Phys. Rev. D* 90.11 (2014), p. 114504. DOI: 10.1103/PhysRevD.90.114504. arXiv: 1310.3626 [hep-lat].

[157] P. A. Boyle et al. "Low energy constants of SU(2) partially quenched chiral perturbation theory from N_f=2+1 domain wall QCD". In: *Phys. Rev. D* 93.5 (2016), p. 054502. DOI: 10.1103/PhysRevD.93.054502. arXiv: 1511.01950 [hep-lat].

[158] Michael F Hutchinson. "A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines". In: *Communications in Statistics-Simulation and Computation* 18.3 (1989), pp. 1059–1076.

[159] S. Bernardson, P. McCarty, and C. Thron. "Monte Carlo methods for estimating linear combinations of inverse matrix entries in lattice QCD". In: *Comput. Phys. Commun.* 78 (1993), pp. 256–264. DOI: 10.1016/0010-4655(94)90004-3.

[160] J. Viehoff et al. "Improving stochastic estimator techniques for disconnected diagrams". In: *Nucl. Phys. Proc. Suppl.* 63 (1998), pp. 269–271. DOI: 10.1016/S0920-5632(97)00742-1. arXiv: hep-lat/9710050 [hep-lat].

[161] Alan Ó Cais et al. "Improving algorithms to compute all elements of the lattice quark propagator". In: *Nucl. Phys. B Proc. Suppl.* 140 (2005). Ed. by Geoffrey T. Bodwin et al., pp. 844–849. DOI: 10.1016/j.nuclphysbps.2004.11.286. arXiv: hep-lat/0409069.

[162] Heather Switzer et al. "Probing for the Trace Estimation of a Permuted Matrix Inverse Corresponding to a Lattice Displacement". In: (June 2021). arXiv: 2106.01275 [hep-lat].

[163] Ronald B. Morgan and Walter Wilcox. "Deflated iterative methods for linear equations with multiple right-hand sides". In: (May 2004). arXiv: math-ph/0405053.

[164] Justin Foley et al. "Practical all-to-all propagators for lattice QCD". In: *Comput. Phys. Commun.* 172 (2005), pp. 145–162. DOI: 10.1016/j.cpc.2005.06.008. arXiv: hep-lat/0505023 [hep-lat].

[165] Gunnar Bali et al. "(Approximate) Low-Mode Averaging with a new Multigrid Eigensolver". In: *PoS* LATTICE2015 (2015), p. 350. arXiv: 1509.06865 [hep-lat].

[166] Gunnar S. Bali, Sara Collins, and Andreas Schäfer. "Effective noise reduction techniques for disconnected loops in Lattice QCD". In: *Comput. Phys. Commun.* 181 (2010), pp. 1570–1583. DOI: 10.1016/j.cpc.2010.05.008. arXiv: 0910.3970 [hep-lat].

[167] Thomas Blum, Taku Izubuchi, and Eigo Shintani. "New class of variance-reduction techniques using lattice symmetries". In: *Phys. Rev. D* 88.9 (2013), p. 094503. DOI: 10.1103/PhysRevD.88.094503. arXiv: 1208.4349 [hep-lat].

[168] Thomas Blum, Taku Izubuchi, and Eigo Shintani. "Error reduction technique using covariant approximation and application to nucleon form factor". In: *PoS* LATTICE2012 (2012). Ed. by Derek Leinweber et al., p. 262. DOI: 10.22323/1.164.0262. arXiv: 1212.5542 [hep-lat].

[169] Eigo Shintani et al. "Covariant approximation averaging". In: *Phys. Rev. D* 91.11 (2015), p. 114511. DOI: 10.1103/PhysRevD.91.114511. arXiv: 1402.0244 [hep-lat].

[170] C. Thron et al. "Padé - Z_2 estimator of determinants". In: *Phys. Rev. D* 57 (1998), pp. 1642–1653. DOI: 10.1103/PhysRevD.57.1642. arXiv: hep-lat/9707001 [hep-lat].

[171] Andreas Frommer et al. "Adaptive aggregation based Domain Decomposition Multigrid for the lattice Wilson Dirac operator". In: *SIAM J. Sci. Comput.* 36 (2014), A1581–A1608. DOI: 10.1137/130919507. arXiv: 1303.1377 [hep-lat].

[172] Simon Heybrock et al. "Lattice QCD with Domain Decomposition on Intel Xeon Phi Co-Processors". In: *The International Conference for High Performance Computing, Networking, Storage, and Analysis: SC14: HPC matters (SC) New Orleans, LA, USA, November 16–21, 2014*. 2014. DOI: 10.1109/SC.2014.11. arXiv: 1412.2629 [hep-lat]. URL: https://misportal.jlab.org/ul/publications/view_pub.cfm?pub_id=13603.

[173] Peter Georg, Daniel Richtmann, and Tilo Wettig. "DD-αAMG on QPACE 3". In: *EPJ Web Conf.* 175 (2018), p. 02007. DOI: 10.1051/epjconf/201817502007. arXiv: 1710.07041 [hep-lat].

[174] Xu Feng, Karl Jansen, and Dru B. Renner. "The $\pi^+\pi^+$ scattering length from maximally twisted mass lattice QCD". In: *Phys. Lett. B* 684 (2010), pp. 268–274. DOI: 10.1016/j.physletb.2010.01.018. arXiv: 0909.3255 [hep-lat].

[175] Takashi Umeda. "A constant contribution in meson correlators at finite temperature". In: *Phys. Rev. D* 75 (2007), p. 094502. DOI: 10.1103/PhysRevD.75.094502. arXiv: hep-lat/0701005 [hep-lat].

[176] Christopher Michael. "Adjoint sources in lattice gauge theory". In: *Nucl. Phys. B* 259 (1985), pp. 58–76. DOI: 10.1016/0550-3213(85)90297-4.

[177] Martin Lüscher and Ulli Wolff. "How to calculate the elastic scattering matrix in two-dimensional quantum field theories by numerical simulation". In: *Nucl. Phys. B* 339 (1990), pp. 222–252. DOI: 10.1016/0550-3213(90)90540-T.

[178] Gunnar Bali, Sara Collins, and Jakob Simeth. "η and η' masses and decay constants". In: *EPJ Web Conf.* 175 (2018), p. 05028. DOI: 10.1051/epjconf/201817505028. arXiv: 1710.06733 [hep-lat].

[179] Benoit Blossier et al. "On the generalized eigenvalue method for energies and matrix elements in lattice field theory". In: *JHEP* 04 (2009), p. 094. DOI: 10.1088/1126-6708/2009/04/094. arXiv: 0902.1265 [hep-lat].

[180] J. Gasser and H. Leutwyler. "Chiral Perturbation Theory: Expansions in the mass of the strange quark". In: *Nucl. Phys. B* 250 (1985), pp. 465–516. DOI: 10.1016/0550-3213(85)90492-4.

[181] P. Bickert, P. Masjuan, and S. Scherer. "η-η' Mixing in Large-N_c Chiral Perturbation Theory". In: *Phys. Rev. D* 95.5 (2017), p. 054023. DOI: 10.1103/PhysRevD.95.054023. arXiv: 1612.05473 [hep-ph].

[182] Xu-Kun Guo et al. "Scrutinizing the η-η' mixing, masses and pseudoscalar decay constants in the framework of U(3) chiral effective field theory". In: *JHEP* 06 (2015), p. 175. DOI: 10.1007/JHEP06(2015)175. arXiv: 1503.02248 [hep-ph].

[183] T. Feldmann, P. Kroll, and B. Stech. "Mixing and decay constants of pseudoscalar mesons: The sequel". In: *Phys. Lett. B* 449 (1999), pp. 339–346. DOI: 10.1016/S0370-2693(99)00085-4. arXiv: hep-ph/9812269 [hep-ph].

[184] Thorsten Feldmann. "Quark structure of pseudoscalar mesons". In: *Int. J. Mod. Phys. A* 15 (2000), pp. 159–207. DOI: 10.1142/S0217751X00000082. arXiv: hep-ph/9907491 [hep-ph].

[185] Rafel Escribano et al. "η' transition form factor from space- and timelike experimental data". In: *Phys. Rev. D* 94.5 (2016), p. 054033. DOI: 10.1103/PhysRevD.94.054033. arXiv: 1512.07520 [hep-ph].

[186] Roland Kaiser and H. Leutwyler. "Pseudoscalar decay constants at large N_c". In: *Proceedings of the Workshop "Nonperturbative methods in quantum field theory", Adelaide, Australia, February 2–13, 1998*. Ed. by A. W. Schreiber, A. G. Willoams, and A. W. Thomas. World Scientific, Singapore, June 1998, pp. 15–29. DOI: 10.1142/3951. arXiv: hep-ph/9806336.

[187] Zhi-Hui Guo and J. A. Oller. "Resonances from meson-meson scattering in U(3) CHPT". In: *Phys. Rev. D* 84 (2011), p. 034005. DOI: 10.1103/PhysRevD.84.034005. arXiv: 1104.2849 [hep-ph].

[188] Maurice Benayoun, L. DelBuono, and Heath Bland O'Connell. "VMD, the WZW Lagrangian and ChPT: The third mixing angle". In: *Eur. Phys. J. C* 17 (2000), pp. 593–610. DOI: 10.1007/s100520000497. arXiv: hep-ph/9905350.

[189] Rafel Escribano and Jean-Marie Frère. "Study of the η-η' system in the two mixing angle scheme". In: *JHEP* 06 (2005), p. 029. DOI: 10.1088/1126-6708/2005/06/029. arXiv: hep-ph/0501072 [hep-ph].

[190] R. Escribano, P. Masjuan, and P. Sanchez-Puertas. "η and η' transition form factors from rational approximants". In: *Phys. Rev. D* 89.3 (2014), p. 034014. DOI: 10.1103/PhysRevD.89.034014. arXiv: 1307.2061 [hep-ph].

[191] Yun-Hua Chen, Zhi-Hui Guo, and Bing-Song Zou. "Unified study of $J/\psi \to PV$, $P\gamma^{(*)}$ and light hadron radiative processes". In: *Phys. Rev. D* 91 (2015), p. 014010. DOI: 10.1103/PhysRevD.91.014010. arXiv: 1411.1159 [hep-ph].

[192] R. Escribano, P. Masjuan, and P. Sanchez-Puertas. "The η transition form factor from space- and time-like experimental data". In: *Eur. Phys. J. C* 75.9 (2015), p. 414. DOI: 10.1140/epjc/s10052-015-3642-z. arXiv: 1504.07742 [hep-ph].

[193] Minghui Ding et al. "$\gamma^*\gamma \to \eta, \eta'$ transition form factors". In: *Phys. Rev. D* 99.1 (2019), p. 014014. DOI: 10.1103/PhysRevD.99.014014. arXiv: 1810.12313 [nucl-th].

[194] Xiao-Wei Gu, Chun-Gui Duan, and Zhi-Hui Guo. "Updated study of the η-η' mixing and the thermal properties of light pseudoscalar mesons at low temperatures". In: *Phys. Rev. D* 98.3 (2018), p. 034007. DOI: 10.1103/PhysRevD.98.034007. arXiv: 1803.07284 [hep-ph].

[195] Maurice Benayoun et al. "Radiative decays, nonet symmetry and SU(3) breaking". In: *Phys. Rev. D* 59 (1999), p. 114027. DOI: 10.1103/PhysRevD.59.114027. arXiv: hep-ph/9902326.

[196] T. Feldmann and P. Kroll. "Mixing of pseudoscalar mesons". In: *Phys. Scripta* T99 (2002). Ed. by J. Bijnens, G. Faldt, and B. M. K. Nefkens, pp. 13–22. DOI: 10.1238/Physica.Topical.099a00013. arXiv: hep-ph/0201044.

[197] Pilar Hernández, Carlos Pena, and Fernando Romero-López. "Large N_c scaling of meson masses and decay constants". In: *Eur. Phys. J. C* 79.10 (2019), p. 865. DOI: 10.1140/epjc/s10052-019-7395-y. arXiv: 1907.11511 [hep-lat].

[198] Johan Bijnens and Gerhard Ecker. "Mesonic low-energy constants". In: *Ann. Rev. Nucl. Part. Sci.* 64 (2014), pp. 149–174. DOI: 10.1146/annurev-nucl-102313-025528. arXiv: 1405.6488 [hep-ph].

[199] Johan Bijnens and Ilaria Jemos. "A new global fit of the L_i^r at next-to-next-to-leading order in Chiral Perturbation Theory". In: *Nucl. Phys. B* 854 (2012), pp. 631–665. DOI: 10.1016/j.nuclphysb.2011.09.013. arXiv: 1103.5945 [hep-ph].

[200] P. Herrera-Siklódy et al. "Chiral effective Lagrangian in the large N_c limit: The nonet case". In: *Nucl. Phys. B* 497 (1997), pp. 345–386. DOI: 10.1016/S0550-3213(97)00260-5. arXiv: hep-ph/9610549.

[201] A. Bazavov et al. "Results for light pseudoscalar mesons". In: *PoS* LATTICE2010 (2010). Ed. by Giancarlo Rossi, p. 074. DOI: 10.22323/1.105.0074. arXiv: 1012.0868 [hep-lat].

[202] Peter Kroll and Kornelija Passek-Kumericki. "The two gluon components of the η and η' mesons to leading twist accuracy". In: *Phys. Rev. D* 67 (2003), p. 054017. DOI: 10.1103/PhysRevD.67.054017. arXiv: hep-ph/0210045.

[203] S. S. Agaev et al. "Transition form factors $\gamma^*\gamma \to \eta$ and $\gamma^*\gamma \to \eta'$ in QCD". In: *Phys. Rev. D* 90.7 (2014), p. 074019. DOI: 10.1103/PhysRevD.90.074019. arXiv: 1409.4311 [hep-ph].

[204] Antoine Gérardin, Harvey B. Meyer, and Andreas Nyffeler. "Lattice calculation of the pion transition form factor with $N_f = 2 + 1$ Wilson quarks". In: *Phys. Rev. D* 100.3 (2019), p. 034520. DOI: 10.1103/PhysRevD.100.034520. arXiv: 1903.09471 [hep-lat].

[205] Antoine Gérardin et al. "Pseudoscalar transition form factors and the hadronic light-by-light contribution to the muon $g-2$". In: *38th International Symposium on Lattice Field Theory*. Dec. 2021. arXiv: 2112.08101 [hep-lat].

[206] J. Gronberg et al. "Measurements of the meson - photon transition form-factors of light pseudoscalar mesons at large momentum transfer". In: *Phys. Rev. D* 57 (1998), pp. 33–54. DOI: 10.1103/PhysRevD.57.33. arXiv: hep-ex/9707031.

[207] D. de Florian et al. "Handbook of LHC Higgs Cross Sections: 4. Deciphering the nature of the Higgs sector". In: 2/2017 (Oct. 2016). DOI: 10.23731/CYRM-2017-002. arXiv: 1610.07922 [hep-ph].

[208] S. S. Agaev et al. "BELLE data on the $\pi^0\gamma*\gamma$ form factor: A game changer?" In: *Phys. Rev. D* 86 (2012), p. 077504. DOI: 10.1103/PhysRevD.86.077504. arXiv: 1206.3968 [hep-ph].

[209] Jian Liang et al. "Quark spins and Anomalous Ward Identity". In: *Phys. Rev. D* 98.7 (2018), p. 074505. DOI: 10.1103/PhysRevD.98.074505. arXiv: 1806.08366 [hep-ph].

[210] Isabel Campos et al. "Non-perturbative quark mass renormalisation and running in $N_f = 3$ QCD". In: *Eur. Phys. J. C* 78.5 (2018), p. 387. DOI: 10.1140/epjc/s10052-018-5870-5. arXiv: 1802.05243 [hep-lat].

[211] Leonardo Giusti et al. "The $U_A(1)$ problem on the lattice with Ginsparg-Wilson fermions". In: *Nucl. Phys. B* 628 (2002), pp. 234–252. DOI: 10.1016/S0550-3213(02)00093-7. arXiv: hep-lat/0108009 [hep-lat].

[212] Luigi Del Debbio and Claudio Pica. "Topological susceptibility from the overlap". In: *JHEP* 02 (2004), p. 003. DOI: 10.1088/1126-6708/2004/02/003. arXiv: hep-lat/0309145 [hep-lat].

[213] H. Leutwyler and Andrei V. Smilga. "Spectrum of Dirac operator and role of winding number in QCD". In: *Phys. Rev. D* 46 (1992), pp. 5607–5632. DOI: 10.1103/PhysRevD.46.5607.

[214] A. Bazavov et al. "Topological susceptibility with the asqtad action". In: *Phys. Rev. D* 81 (2010), p. 114501. DOI: 10.1103/PhysRevD.81.114501. arXiv: 1003.5695 [hep-lat].

[215] Abhishek Chowdhury et al. "Topological susceptibility in Lattice QCD with unimproved Wilson fermions". In: *Phys. Lett. B* 707 (2012), pp. 228–232. DOI: 10.1016/j.physletb.2011.12.034. arXiv: 1110.6013 [hep-lat].

[216] Krzysztof Cichy, Elena Garcia-Ramos, and Karl Jansen. "Topological susceptibility from the twisted mass Dirac operator spectrum". In: *JHEP* 02 (2014), p. 119. DOI: 10.1007/JHEP02(2014)119. arXiv: 1312.5161 [hep-lat].

[217] Mattia Bruno, Stefan Schaefer, and Rainer Sommer. "Topological susceptibility and the sampling of field space in $N_f = 2$ lattice QCD simulations". In: *JHEP* 08 (2014), p. 150. DOI: 10.1007/JHEP08(2014)150. arXiv: 1406.5363 [hep-lat].

[218] Claudio Bonati et al. "Axion phenomenology and θ-dependence from $N_f = 2 + 1$ lattice QCD". In: *JHEP* 03 (2016), p. 155. DOI: 10.1007/JHEP03(2016)155. arXiv: 1512.06746 [hep-lat].

[219] Sz. Borsanyi et al. "Calculation of the axion mass based on high-temperature lattice quantum chromodynamics". In: *Nature* 539.7627 (2016), pp. 69–71. DOI: 10.1038/nature20115. arXiv: 1606.07494 [hep-lat].

[220] Constantia Alexandrou et al. "Topological susceptibility from twisted mass fermions using spectral projectors and the gradient flow". In: *Phys. Rev. D* 97.7 (2018), p. 074503. DOI: 10.1103/PhysRevD.97.074503. arXiv: 1709.06596 [hep-lat].

[221] V. A. Novikov et al. "A theory of the $J/\psi \to \eta(\eta')\gamma$ decays". In: *Nucl. Phys. B* 165 (1980), pp. 55–66. DOI: 10.1016/0550-3213(80)90305-3.

[222] Janardan Prasad Singh. "Light quark pseudoscalar densities and anomaly matrix elements for η and η' mesons". In: *Phys. Rev. D* 88.9 (2013), p. 096005. DOI: 10.1103/PhysRevD.88.096005. arXiv: 1307.3311 [hep-ph].

[223] Martin Beneke and Matthias Neubert. "Flavor singlet B decay amplitudes in QCD factorization". In: *Nucl. Phys. B* 651 (2003), pp. 225–248. DOI: 10.1016/S0550-3213(02)01091-X. arXiv: hep-ph/0210085.

[224] Hai-Yang Cheng, Hsiang-nan Li, and Keh-Fei Liu. "Pseudoscalar glueball mass from $\eta - eta' - G$ mixing". In: *Phys. Rev. D* 79 (2009), p. 014024. DOI: 10.1103/PhysRevD.79.014024. arXiv: 0811.2577 [hep-ph].

[225] Wen Qin, Qiang Zhao, and Xian-Hui Zhong. "Revisiting the pseudoscalar meson and glueball mixing and key issues in the search for a pseudoscalar glueball state". In: *Phys. Rev. D* 97.9 (2018), p. 096002. DOI: 10.1103/PhysRevD.97.096002. arXiv: 1712.02550 [hep-ph].

[226] H. Goldberg. "The quark-glue structure of the η and η' mesons with application to $\psi \to \eta(\eta')\,\gamma$ and $\psi' \to \psi\eta$". In: *Phys. Rev. Lett.* 44 (1980), p. 363. DOI: 10.1103/PhysRevLett.44.363.

[227] Patricia Bickert and Stefan Scherer. "Two-photon decays and transition form factors of π^0, η, and η' in large-N_c chiral perturbation theory". In: *Phys. Rev. D* 102.7 (2020), p. 074019. DOI: 10.1103/PhysRevD.102.074019. arXiv: 2005.08550 [hep-ph].

[228] Robert G. Edwards and Balint Joó. "The Chroma software system for Lattice QCD". In: *Nucl. Phys. B Proc. Suppl.* 140 (2005), p. 832. DOI: 10.1016/j.nuclphysbps.2004.11.254. arXiv: hep-lat/0409003 [hep-lat].

[229] F. James and M. Roos. "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations". In: *Comput. Phys. Commun.* 10 (1975), pp. 343–367. DOI: 10.1016/0010-4655(75)90039-9.

[230] Jay Orear. "Least squares when both variables have uncertainties". In: *Am. J. Phys.* 50 (1982). [Erratum: Am. J. Phys. **52** (1984) 278], p. 912. DOI: 10.1119/1.12972.

[231] Jülich Supercomputing Centre. "JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre". In: *Journal of Large-Scale Research Facilities* 1 (2015), A1. DOI: 10.17815/jlsrf-1-18.

[232] Jülich Supercomputing Centre. "JUWELS: Modular Tier-0/1 Supercomputer at the Jülich Supercomputing Centre". In: *Journal of Large-Scale Research Facilities* 5 (2019), A135. DOI: 10.17815/jlsrf-5-171.

[233] Jülich Supercomputing Centre. "JURECA: Modular supercomputer at Jülich Supercomputing Centre". In: *Journal of Large-Scale Research Facilities* 4 (2018), A132. DOI: 10.17815/jlsrf-4-121-1.

[234] Paul Arts et al. "QPACE 2 and Domain Decomposition on the Intel Xeon Phi". In: *PoS* LATTICE2014 (2015), p. 021. DOI: 10.22323/1.214.0021. arXiv: 1502.04025 [cs.DC].

[235] Simon Heybrock et al. "Adaptive algebraic multigrid on SIMD architectures". In: *PoS* LATTICE2015 (2016), p. 036. DOI: 10.22323/1.251.0036. arXiv: 1512.04506 [physics.comp-ph].

[236] Daniel Richtmann, Simon Heybrock, and Tilo Wettig. "Multiple right-hand-side setup for the DD-αAMG". In: *PoS* LATTICE2015 (2016), p. 035. DOI: 10.22323/1.251.0035. arXiv: 1601.03184 [hep-lat].

[237] Peter Georg, Daniel Richtmann, and Tilo Wettig. "pMR: A high-performance communication library". In: *PoS* LATTICE2016 (2017), p. 361. DOI: 10.22323/1.256.0361. arXiv: 1701.08521 [hep-lat].

List of acronyms

AWI	axial Ward identity
BNL	Brookhaven National Laboratory
ChPT	chiral perturbation theory
CLS	coordinated lattice simulations
EFT	effective field theory
ETMC	European twisted mass collaboration
FKS	Feldmann–Kroll–Stech
GEVP	generalized eigenvalue problem
GMOR	Gell-Mann–Oakes–Renner
HMC	hybrid Monte-Carlo
HPE	hopping parameter expansion
LCDA	light-cone distribution amplitude
LCSR	light-cone sum rule
LEC	low energy constant
LO	leading order
$\overline{\text{MS}}$	modified minimal subtraction
NLO	next-to-leading order
NNLO	next-to-next-to-leading order
OZI	Okubo-Zweig-Iizuka
PDG	particle data group
PCAC	partially conserved axial current
QED	quantum electrodynamics
QCD	quantum chromodynamics
RG	renormalization group
RHMC	rational hybrid Monte-Carlo
TFF	transition form factor

Acknowledgements

It is my pleasure to thank all the people that were important during all stages of this work. First and foremost there is my advisor Prof. Dr. Gunnar Bali, and PD Dr. Sara Collins who was not less important for the successful completion of this work. Through their supervision and guidance they made this work possible, and their doors were always open in all situations. I enjoyed witnessing Gunnar's steady flow of ideas that was sometimes hard for me to keep up with. Sara's constant encouragement and supportive nature kept me going and helped me direct my work to what was important. I was lucky to have you as my mentors.

I also wish to thank Prof. Dr. Vladimir Braun and Prof. Dr. Andreas Schäfer. Their commitment in enabling research made this and many other projects possible, and their contributions to the paper were indispensable. I am particularly grateful to Volodya for adjusting and rerunning the transition form factor code and for many useful discussions in the course of this work.

Further, I would like to mention Dr. André Sternbeck who sparked my excitement for (lattice) field theory in the early days of my scientific journey. I will keep this time and the many discussions in good memory.

Monika Maschek and Heidi Decock with their open ears and welcoming nature made mastering some of the administrative tasks much more enjoyable.

I am particularly indebted to Marius Löffler and Dr. Peter Plößl for their comments and suggestions during proofreading parts of the manuscript.

My special thanks go to all the other colleagues and office mates who accompanied me during my long time in this department. Some of them became close friends over the years and I want to mention (in no particular order) Dr. Rudolf Rödl, Dr. Johannes Najjar, Dr. Benjamin Gläßle, Daniel Richtmann, Peter Georg, Simon Weishäupl, PD Dr. Wolfgang Söldner, Dr. Enno Scholz, Dr. Stefan Solbrig, Prof. Dr. Gergely Endrődy, Dr. Piotr Korcyl, Dr. Claudia Uebler, Dr. Maximilian Emmerich, Dr. Jacob Wellnhofer. In particular the time in the office with Peter and Max were both very enjoyable and productive.

I am also grateful to the colleagues in CLS for the joint effort in the generation of the gauge field ensembles. I particularly enjoyed working with Dr. Hubert Simma towards an error resilient data management of the CLS configurations.

Lastly, I want to express my gratitude to my family: To my parents for teaching me the persistence and excitement that is needed to embark on — and finish — such a project, and for supporting me in all possible ways. And last, but not least, to my beloved wife Caro for all the support she has provided and for her perseverance in bearing my from time to time absent mind.

This work was supported by the Deutsche Forschungsgemeinschaft through the collaborative research centre SFB/TRR-55 and the Research Unit FOR 2926 "Next Generation pQCD for Hadron Structure: Preparing for the EIC" and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 813942 (ITN EuroPLEx) and grant agreement no. 824093 (STRONG-2020). We also acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time through the John von Neumann Institute for Computing (NIC) on the the super-computers JUQUEEN [231], JUWELS [232] and in particular JURECA-Booster [233] at Jülich Supercomputing Centre (JSC). GCS is the alliance of the three national supercomputing centres HLRS (Universität Stuttgart), JSC (Forschungszentrum Jülich), and LRZ (Bayerische Akademie der Wissenschaften), funded by the German Federal Ministry of Education and Research (BMBF) and the German State Ministries for Research of Baden-Württemberg (MWK), Bayern (StMWFK) and Nordrhein-Westfalen (MIWF). Simulations were also performed on the Regensburg iDataCool and Athene2 clusters, and the SFB/TRR 55 QPACE 2 [234] and QPACE 3 machines. The multigrid solver of [172, 235–237] was used for the inversion of the Dirac operator and the `QDP++` framework in conjunction with `chroma` [228] was used for the measurement code.