Published for SISSA by Springer
Received: June 21, 2021
Accepted: July 18, 2021
Published: August 25, 2021

Masses and decay constants of the η and η' mesons from lattice QCD

RQCD

The RQCD collaboration

Gunnar S. Bali,[a,b] **Vladimir Braun,**[a] **Sara Collins,**[a] **Andreas Schäfer**[a] **and Jakob Simeth**[a]

[a] *Institut für Theoretische Physik, Universität Regensburg, D-93040 Regensburg, Germany*
[b] *Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400005, India*

E-mail: gunnar.bali@ur.de, vladimir.braun@ur.de, sara.collins@ur.de, andreas.schaefer@ur.de, jakob.simeth@ur.de

Abstract: We determine the masses, the singlet and octet decay constants as well as the anomalous matrix elements of the η and η' mesons in $N_f = 2 + 1$ QCD. The results are obtained using twenty-one CLS ensembles of non-perturbatively improved Wilson fermions that span four lattice spacings ranging from $a \approx 0.086\,\text{fm}$ down to $a \approx 0.050\,\text{fm}$. The pion masses vary from $M_\pi = 420\,\text{MeV}$ to $126\,\text{MeV}$ and the spatial lattice extents L_s are such that $L_s M_\pi \gtrsim 4$, avoiding significant finite volume effects. The quark mass dependence of the data is tightly constrained by employing two trajectories in the quark mass plane, enabling a thorough investigation of U(3) large-N_c chiral perturbation theory (ChPT). The continuum limit extrapolated data turn out to be reasonably well described by the next-to-leading order ChPT parametrization and the respective low energy constants are determined. The data are shown to be consistent with the singlet axial Ward identity and, for the first time, also the matrix elements with the topological charge density are computed. We also derive the corresponding next-to-leading order large-N_c ChPT formulae. We find $F^8 = 115.0(2.8)$ MeV, $\theta_8 = -25.8(2.3)^\circ$, $\theta_0 = -8.1(1.8)^\circ$ and, in the $\overline{\text{MS}}$ scheme for $N_f = 3$, $F^0(\mu = 2\,\text{GeV}) = 100.1(3.0)$ MeV, where the decay constants read $F_\eta^8 = F^8 \cos\theta_8$, $F_{\eta'}^8 = F^8 \sin\theta_8$, $F_\eta^0 = -F^0 \sin\theta_0$ and $F_{\eta'}^0 = F^0 \cos\theta_0$. For the gluonic matrix elements, we obtain $a_\eta(\mu = 2\,\text{GeV}) = 0.0170(10)\,\text{GeV}^3$ and $a_{\eta'}(\mu = 2\,\text{GeV}) = 0.0381(84)\,\text{GeV}^3$, where statistical and all systematic errors are added in quadrature.

Keywords: Lattice QCD, Chiral Lagrangians, $1/N$ Expansion, Anomalies in Field and String Theories

ArXiv ePrint: 2106.05398

https://doi.org/10.1007/JHEP08(2021)137

Contents

JHEP08(2021)137

1 Introduction

The physics of the pseudoscalar η and η' mesons is a fascinating area at the crossroads of many themes in hadron and particle physics. In the exact flavour SU(3) limit the η meson is part of the flavour-octet whereas the η' is a pure flavour-singlet state whose properties are intimately related to the celebrated axial anomaly [1, 2]. However, it is known empirically that the SU(3) breaking effects are large and have a non-trivial structure. These effects are usually described in terms of a mixing scheme that considers the physical η and η' mesons as superpositions of fundamental (e.g., flavour-singlet and -octet) fields in a low energy effective theory. Modern phenomenological analyses of η-η' mixing are largely based on large-N_c chiral perturbation theory (ChPT) [3–6] which allows for a unified treatment of the $\eta^{(\prime)}$-mesons together with the pseudo-Goldstone octet of the lightest pseudoscalars. 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., [7] and references therein.

Flavour-singlet pseudoscalar mesons are a very active area of research. Chiral dynamics has been very successful in describing low energy pion and kaon reactions and it is natural to attempt to generalize this to include the $\eta^{(\prime)}$ sector. Theoretical developments as well as new high-precision experimental measurements are needed to advance this agenda. The study of η and η' mesons also provides an interesting window to beyond-the-standard-model (BSM) physics. BSM searches in $\eta^{(\prime)}$ decays have initially been related to flavour-conserving tests of discrete symmetries, however, other interesting searches have been proposed [7], e.g., for axion-like particles. Corresponding efforts are ongoing or planned in many experimental facilities. A less well explored area is the production of η and η' in hard processes, e.g., in B-meson decays or in two-photon reactions $\gamma^*\gamma \to \eta^{(\prime)}$, which constitute part of the Belle II research programme [8]. It is not obvious whether and to what extent the approaches based on low energy effective field theory provide an adequate description of such processes, that are dominated by meson wave functions at small transverse separations, referred to as light-cone distribution amplitudes (LCDAs). However, this is usually taken as a working hypothesis in phenomenological applications, see, e.g., [9–12]. One important issue in this context is that $\eta^{(\prime)}$ mesons, in contrast to the pion, can contain a significant admixture of a two-gluon component at low scales, i.e. a comparably large two-gluon LCDA. Several

JHEP08(2021)137

different reactions were considered in an effort to extract or at least to constrain these contributions, see, e.g., [10, 13, 14]. However, no definite conclusion can be drawn as yet.

Lattice simulations of properties of flavour non-singlet pseudoscalar mesons are quite advanced. Recently, continuum limit results at physical quark masses of the first two Gegenbauer moments of the twist-two pion and kaon LCDAs [15] were obtained, pion transition form factors calculated [16] and exploratory studies of higher twist LCDA parameters undertaken [17]. However, only a few investigations of matrix elements involving the η and η' mesons exist to-date [18, 19]. These 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 [20–22], and continuing with $N_f = 2$ mass-degenerate dynamical light quarks [21, 23–30]. 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 ($N_f = 2+1$) [18, 31–33] (see also [34] for a different attempt using correlators of the topological charge density). More recently, $N_f = 2+1+1$ results [19, 35, 36] employing the twisted-mass fermion formulation, using several ensembles and lattice spacings, enabled a physical point extrapolation. In [37] the η' mass was calculated at non-zero temperature from topological charge density correlators. In [19] also pseudoscalar matrix elements were determined. Relating these to the four decay constants of the η/η' system enabled the first lattice determination 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 calculation of the semileptonic decay form factors $D_s \to \eta, \eta'$, albeit only on two ensembles at a single lattice spacing [18].

Here, we compute the masses, decay constants and gluonic anomaly matrix elements of the η and η' mesons in $N_f = 2+1$ QCD. The simulations are carried out on twenty-one ensembles generated by the CLS (Coordinated Lattice Simulations) initiative [38, 39], employing non-perturbatively improved Sheikholeslami-Wilson fermions. Most of the ensembles have open boundary conditions in time, ensuring that the topological sectors are sampled uniformly. We employ pion masses that range from the SU(3) symmetric point at $M_\pi \approx 420\,\text{MeV}$ down to just below the physical pion mass. The corresponding kaon masses are tuned so that the ensembles fall onto two distinct trajectories — one at a constant average quark mass and the other at an approximately constant strange quark mass. Both lines intersect close to the physical point, aiding the chiral interpolation. The continuum extrapolation is carried out utilizing four lattice spacings ranging from $a \approx 0.086\,\text{fm}$ down to $a \approx 0.050\,\text{fm}$.

Using a combination of all-to-all propagator methods and various interpolating operators, we obtain a correlation matrix between pairs of interpolators as a function of the Euclidean time separation. In addition, for each of three local currents, axialvector, pseu-

JHEP08(2021)137

doscalar and gluonic, we compute the vector of correlators with the interpolating operators. From these the meson masses and matrix elements are extracted via a fit, utilizing a generalized effective mass method that we introduce. A comparison is made with the result of the usual generalized eigenvalue problem (GEVP) method. The matrix elements are renormalized and partially order a improved. Remaining order a terms with (as yet) unknown coefficients as well as order a^2 corrections are included in parametrizations of the lattice spacing and quark mass dependence. Regarding the continuum limit, we are able to simultaneously parameterize all data in terms of the six low energy constants (LECs) of large-N_c U(3) ChPT at next-to-leading order (NLO). Systematic errors are estimated by carrying out a multitude of fits and also by excluding data points at large average quark masses.

The masses of the mesons are found to be in agreement with experiment and we determine the two decay constants (singlet and octet) for each meson as well as the LECs of large-N_c ChPT. Due to the axial anomaly, some LECs depend on the QCD renormalization scale, as do the singlet decay constants, and we present our results in the $\overline{\text{MS}}$ scheme at different scales. The gluonic matrix elements of the η and η' mesons and the topological susceptibility are found to be affected by sizeable lattice cut-off effects. In the continuum limit the topological susceptibility is well described by the leading order (LO) ChPT expectation. This only depends on the pion decay constant in the chiral limit, which we obtain from our global fits to the axial matrix elements. The continuum limit η and η' matrix elements satisfy the flavour-octet and flavour-singlet axial Ward identities (AWIs) and the pseudoscalar fermionic matrix elements are determined too. We address implications on the phenomenology of hard processes for the example of the $\gamma\gamma^* \to \eta^{(\prime)}$ transition form factors as well as radiative decays of the J/ψ to an $\eta^{(\prime)}$ meson.

The conventions and main results of this article can be found in the following places. In section 2 we detail the flavour mixing schemes and define our normalization conventions and some of our notations. The basic ChPT formulae can be found in section 5.3, and figure 5 illustrates the main results on the masses and decay constants. Section 6 details the determination of the gluonic matrix elements. The continuum limit results are collected and discussed in section 7. Our main results on the masses, decay constants and gluonic matrix elements are summarized in section 8.

The remainder of the article is organized as follows. In section 3 we discuss the simulation parameters, the lattice observables and the computational techniques used. We then move on to introduce the generalized effective mass method, that we employ to extract the masses in section 4. In that section we also explain the lattice evaluation of the necessary matrix elements and our statistical analysis. In section 5 we discuss the renormalization and improvement of the lattice results and the parametrizations of the quark mass and lattice spacing dependence. We then determine the meson masses, decay constants and LECs, and estimate their systematics. As mentioned above, in section 6 we determine the quark mass dependence of the topological susceptibility and the gluonic matrix elements of the η and η' mesons. The results are parameterized in terms of NLO large-N_c U(3) ChPT. In section 7, apart from discussing the continuum limit results, we also address implications on the phenomenology of hard processes for the example of the $\gamma\gamma^* \to \eta^{(\prime)}$ transition form factors, see section 7.5.

JHEP08(2021)137

The article is augmented by several appendices: in appendix A, we present the parametrization of the pseudoscalar loop contributions that appear at next-to-next-to-leading order (NNLO). In appendix B we derive the dependence of the gluonic and pseudoscalar fermionic matrix elements on the pion and kaon masses in terms of the six NLO large-N_c U(3) LECs. In appendix C we show the result of a LO fit to the η and η' masses that we omitted from the main body of the paper for brevity. In appendix D we collect the values of the parameters accompanying the lattice spacing effects for seventeen different fit forms. Finally, in appendix E we list our results for the decay constants and mixing angles in different flavour bases and at different scales, both in units of the gradient flow scale t_0 and in physical units. A corresponding table with gluonic matrix elements can be found in section 6.

2 Definitions, conventions and octet/singlet mixing

The couplings between axialvector currents and pseudoscalar states, that are also known as meson decay constants, play a crucial role in the description of low energy physics with η and η' mesons. Different normalization conventions are used throughout the literature. Here we introduce the conventions for the local currents, decay constants and interpolating operators that we employ in this article. We also briefly address what is often referred to as η-η' mixing or pseudoscalar octet/singlet mixing.

Everywhere we will assume $N_f = 3$ quark flavours with the masses m_u, m_d and m_s, where we ignore the mass difference between the up and the down quark as well as electromagnetic effects, i.e. we set $m_\ell = m_u = m_d$. For convenience, sometimes we write out the dependence on N_f and N_c. Formulae without these factors always refer to the case $N_f = N_c = 3$.

We define the U(N_f) generators t^a with $t^0 = \mathbb{1}/\sqrt{2N_f}$, where in the $N_f = 3$ case $t^a = \lambda^a/2$ for $a \neq 0$ and λ^a are the eight Gell-Mann matrices. This normalization of the generators corresponds to $\mathrm{tr}\left(t^a t^b\right) = \frac{1}{2}\delta^{ab}$. Using $\overline{\psi} = (\bar{u}, \bar{d}, \bar{s})$, we can introduce local currents as

$$J^a = \overline{\psi} t^a \Gamma_J \psi, \tag{2.1}$$

where the Dirac matrix structure Γ_J defines the current J, for instance $A^a_\mu = \overline{\psi} t^a \gamma_\mu \gamma_5 \psi$ and $P^a = \overline{\psi} t^a \gamma_5 \psi$. It is also useful to define currents for individual quark flavours:

$$J^q = \bar{q} \Gamma_J q, \quad J^\ell = \frac{1}{\sqrt{2}}\left(J^u + J^d\right), \tag{2.2}$$

where $q \in \{u, d, s\}$. The flavour-diagonal singlet, triplet and octet currents J^0, J^3 and J^8 can be written as linear combinations of the above flavour basis currents:

$$J^3 = \frac{1}{2}\left(J^u - J^d\right), \tag{2.3}$$

$$J^8 = \frac{1}{\sqrt{12}}\left(J^u + J^d - 2J^s\right) = \frac{1}{\sqrt{6}} J^\ell - \frac{1}{\sqrt{3}} J^s, \tag{2.4}$$

$$J^0 = \frac{1}{\sqrt{6}}\left(J^u + J^d + J^s\right) = \frac{1}{\sqrt{3}} J^\ell + \frac{1}{\sqrt{6}} J^s. \tag{2.5}$$

JHEP08(2021)137

We define the decay constants $F^a_{\mathcal{M}}$ of a pseudoscalar meson $\mathcal{M}$:

$$\left\langle \Omega \left| A^a_\mu \right| \mathcal{M}(p) \right\rangle = i F^a_{\mathcal{M}} p_\mu, \tag{2.6}$$

where $|\Omega\rangle$ denotes the vacuum and $|\mathcal{M}(p)\rangle$ a meson state with four-momentum p. Below we will often refer to the latter as $|n\rangle$ where $n \in \mathbb{N}_0$ labels the η meson ($n = 0$), the η' meson ($n = 1$) and their excitations ($n \geq 2$). Note that in the above normalization, at the physical point, $F^3_{\pi^0} \approx 92\,\text{MeV}$. For $m_u = m_d = m_s$, these conventions correspond to $f_\pi = \sqrt{2}F^3_{\pi^0} = \sqrt{2}F^8_{\eta^8} = 2F^1_{\pi^+} = -2iF^2_{\pi^+}$, where $if_\pi p_\mu = \langle 0|\bar{d}\gamma_\mu\gamma_5 u|\pi^+(p)\rangle$. The usual normalization for the analogous decay constants in the light/strange flavour basis reads:

$$\left\langle \Omega \left| A^q_\mu \right| \mathcal{M}(p) \right\rangle = i\sqrt{2} F^q_{\mathcal{M}} p_\mu. \tag{2.7}$$

Since we only consider the isospin symmetric limit, the triplet couplings $F^3_{\mathcal{M}}$ vanish identically for the η and the η' mesons. Note that the singlet couplings $F^0_{\mathcal{M}}$ in the standard $\overline{\text{MS}}$ scheme acquire an anomalous dimension due to the axial U(1) anomaly [40],

$$\mu^2 \frac{\mathrm{d}}{\mathrm{d}\mu^2} F^0_{\mathcal{M}}(\mu) = -\frac{N_f}{2}\frac{\alpha_s^2}{\pi^2} F^0_{\mathcal{M}}(\mu) + \mathcal{O}(\alpha_s^3), \tag{2.8}$$

whereas the octet couplings $F^8_{\mathcal{M}}$ are scale independent. This simplifies the scale evolution in the octet/singlet basis relative to that in the light/strange flavour basis.

We introduce pseudoscalar interpolating operators $\mathcal{P}^a_{\vec{p}}$ that have the same flavour structure as the local currents P^a. However, these can be non-local in space (due to quark smearing) and are projected onto a definite spatial momentum $\vec{p}$ in order to destroy physical states with matching quantum numbers. For instance $\mathcal{P}^{3\dagger}_{\mathbf{0}}|\Omega\rangle$ gives a linear combination of the $|\pi^0\rangle$ and its excitations (at rest). The relation between the interpolators $\mathcal{P}^a$ and their flavour basis counterparts $\mathcal{P}^q$ are so as to preserve the normalization of the respective quark model wave functions, resulting in normalizations that differ by factors $\sqrt{2}$ from the relations between P^a and P^q of eqs. (2.3)–(2.5). For the pseudoscalar singlet and octet interpolators this means that

$$\mathcal{P}^8 = \frac{1}{\sqrt{6}}\left(\mathcal{P}^u + \mathcal{P}^d - 2\mathcal{P}^s\right) = \frac{1}{\sqrt{3}}\mathcal{P}^\ell - \sqrt{\frac{2}{3}}\mathcal{P}^s, \tag{2.9}$$

$$\mathcal{P}^0 = \frac{1}{\sqrt{3}}\left(\mathcal{P}^u + \mathcal{P}^d + \mathcal{P}^s\right) = \sqrt{\frac{2}{3}}\mathcal{P}^\ell + \frac{1}{\sqrt{3}}\mathcal{P}^s. \tag{2.10}$$

The factor $\sqrt{2}$ within eq. (2.7) stems from enforcing the above relations also between the decay constants in the octet/singlet basis and the light/strange flavour basis:

$$\begin{pmatrix} F^8_{\mathcal{M}} \\ F^0_{\mathcal{M}} \end{pmatrix} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & -\sqrt{2} \\ \sqrt{2} & 1 \end{pmatrix} \begin{pmatrix} F^\ell_{\mathcal{M}} \\ F^s_{\mathcal{M}} \end{pmatrix}. \tag{2.11}$$

An interesting limit, that we also simulate here, is that of exact SU(3) flavour symmetry ($m_s = m_\ell$). In this limit the η meson is part of a flavour-octet, $\eta = \eta_8$, and η' is a flavour-singlet, $\eta' = \eta_0$. This means that the interpolator $\mathcal{P}^0$ has no overlap with the η meson,

JHEP08(2021)137

i.e. $\langle\eta(p)|\mathcal{P}^{0\dagger}_{\vec{p}}|\Omega\rangle = 0$, and the interpolator $\mathcal{P}^8$ cannot create an η' meson. In this limit, in terms of the decay constants, $F^8_{\eta'} = F^0_\eta = 0$. However, SU(3) breaking corrections are known to be large and phenomenologically significant.

In the chiral effective theory the η' meson can be included in the framework of the $1/N_c$ expansion [3–6]. In this approach the axial anomaly contributes an effective mass term to the η-η' system. This affects not only the flavour-singlet sector but also the flavour-diagonal octet if SU(3) symmetry is explicitly broken by a quark mass difference. An additional off-diagonal contribution to the kinetic term $\partial_\mu\eta_8\partial^\mu\eta_0$ shows up at the loop-level [6]. As a result, the relation of the physical η and η' states to the octet and singlet fields η_8 and η_0 in the chiral Lagrangian becomes more involved, see, e.g., [7, 41, 42].

In general, there are four decay constants, one octet and one singlet decay constant each for the η and for the η'. One can always parameterize these in terms of two fundamental decay constants F^8 and F^0 and two mixing angles θ_8 and θ_0:

$$\begin{pmatrix} F^8_\eta & F^0_\eta \\ F^8_{\eta'} & F^0_{\eta'} \end{pmatrix} = \begin{pmatrix} F^8\cos\theta_8 & -F^0\sin\theta_0 \\ F^8\sin\theta_8 & F^0\cos\theta_0 \end{pmatrix}. \tag{2.12}$$

An analogous parametrization can also be introduced in the flavour basis:

$$\begin{pmatrix} F^\ell_\eta & F^s_\eta \\ F^\ell_{\eta'} & F^s_{\eta'} \end{pmatrix} = \begin{pmatrix} F^\ell\cos\phi_\ell & -F^s\sin\phi_s \\ F^\ell\sin\phi_\ell & F^s\cos\phi_s \end{pmatrix}. \tag{2.13}$$

Obviously, in the SU(3) limit $F^8 = F^8_\eta$, $F^0 = F^0_{\eta'}$ and $\theta_0 = \theta_8 = 0$ while no such simplification exists in the flavour basis. Note that in the standard $\overline{\text{MS}}$ scheme, within the right hand sides of the above equations, only F^8, θ_8 and θ_0 are scale independent, whereas F^0, F^ℓ, F^s, ϕ_ℓ and ϕ_s all will depend on the QCD renormalization scale.

The above two choices of basis are essentially equivalent and the rationale for the popularity of the flavour scheme is that the difference between ϕ_ℓ and ϕ_s (which is formally a $1/N_c$ effect) is small and compatible with zero in phenomenological extractions from experimental data [41, 43]. This feature may be related to the observation that the vector mesons ω and ϕ are to a very good approximation pure $\bar{u}u + \bar{d}d$ and $\bar{s}s$ states, respectively, and the same holds for the tensor mesons $f_2(1270)$ and $f_2'(1525)$. The smallness of flavour mixing in these cases is a manifestation of the phenomenologically very successful Okubo-Zweig-Iizuka (OZI) rule. If the axial U(1) anomaly was the only new effect in the pseudoscalar channels, it may be natural to assume that physical states are related to the flavour states by an orthogonal transformation with a single mixing angle $\phi = \phi_\ell = \phi_s$ [43]. In this approximation the relation between the two schemes simplifies to [7, 41]

$$(F^8)^2 = \frac{1}{3}(F^\ell)^2 + \frac{2}{3}(F^s)^2, \qquad (F^0)^2 = \frac{2}{3}(F^\ell)^2 + \frac{1}{3}(F^s)^2, \tag{2.14}$$

$$\theta_8 = \phi - \arctan\left(\frac{\sqrt{2}F^s}{F^\ell}\right), \qquad \theta_0 = \phi - \arctan\left(\frac{\sqrt{2}F^\ell}{F^s}\right). \tag{2.15}$$

Within QCD, obviously, the above relations cannot hold at arbitrary renormalization scales. Moreover, other matrix elements, e.g., $\langle\Omega|P^\ell|\eta^{(\prime)}\rangle$ and $\langle\Omega|P^s|\eta^{(\prime)}\rangle$, are not necessarily related by the same angles. Our lattice QCD calculation will enable us to check the

JHEP08(2021)137

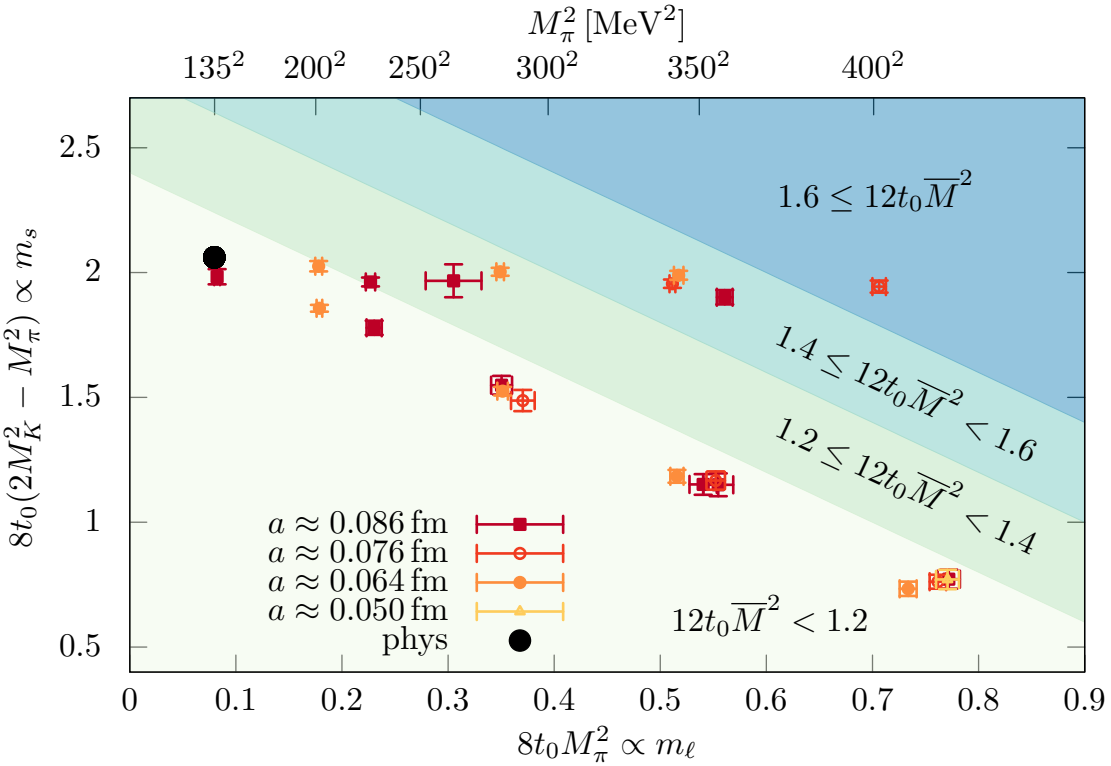


Figure 1. The positions of the analysed CLS ensembles in the quark mass plane. Two mass trajectories were realized that intersect approximately at the physical point (black circle). Along one trajectory the average quark mass is held fixed, along the other trajectory the strange quark mass is kept approximately constant. The symbols encode the four lattice spacings while the shaded areas indicate the values of the average squared pseudoscalar mass $\overline{M}^2$, see eq. (3.4). At the physical point, $12t_0^{\rm ph}\overline{M}^2 \approx 1.11$, where $(8t_0^{\rm ph})^{-1/2} \approx 0.475\,\text{GeV}$ [44].

extent of the validity of the simple mixing picture and the range of applicability of large-N_c ChPT.

3 Lattice computation

In this section we give details of the lattice setup and gauge ensembles and outline the construction of correlation functions, using the local currents and interpolators defined above. We describe the methods for the efficient evaluation of the resulting connected and disconnected quark line diagrams.

3.1 Gauge ensembles

We analyse gauge ensembles with $N_f = 2+1$ non-perturbatively improved Wilson fermions on a Lüscher-Weisz gauge background that were generated within the CLS initiative [38, 39]. To avoid topological freezing at fine lattice spacings, most of the ensembles employ open boundary conditions in time [45]. This breaks translational invariance in that direction and introduces boundary effects, such that measurements must be taken in the bulk of

JHEP08(2021)137

the lattice. The fermion action ensures that hadron masses are free of discretization effects that are linear in the lattice spacing, however, the operators also need to be $\mathcal{O}(a)$ improved. For the currents relevant for this study, we perform the substitutions [46] ($a = 1, \ldots, 8$):

$$A^a_\mu \mapsto A^a_\mu + ac_A \partial_\mu P^a, \tag{3.1}$$

$$A^0_\mu \mapsto A^0_\mu + ac^s_A \partial_\mu P^0, \tag{3.2}$$

$$P^0 \mapsto P^0 + ag_P \,\mathrm{tr}\, F_{\mu\nu}\widetilde{F}_{\mu\nu} = P^0 + ac^s_P \partial_\mu A^0_\mu, \tag{3.3}$$

where we re-expressed the last equation using the singlet AWI in the massless case.[1] In the chiral limit these replacements remove all $\mathcal{O}(a)$ effects. The non-singlet pseudoscalar currents, P^a, are already $\mathcal{O}(a)$ improved in this case. For non-vanishing quark masses, additional mass dependent $\mathcal{O}(a)$ improvement terms appear. These terms and the values of the associated improvement coefficients as well as those of c_A and the (unknown) singlet coefficient c^s_A and c^s_P will be addressed in sections 4.4 and 5.4 and used in the fits of section 6.2.

We carry out our analysis on 21 distinct CLS ensembles that differ in terms of the quark masses, volumes and lattice spacings, see table 1. This enables us to control all sources of systematic error. In the table we give dimensionless combinations involving the average $\overline{M}^2$ and the difference δM^2 of the squared non-singlet pseudoscalar masses,

$$12t_0\overline{M}^2 = 4t_0\left(2M_K^2 + M_\pi^2\right), \quad 8t_0\delta M^2 = 16t_0(M_K^2 - M_\pi^2), \tag{3.4}$$

where t_0 denotes the gradient flow scale, introduced in [48]. The combinations $\overline{M}^2$ and δM^2 will be used in the expressions for the quark mass dependence of the masses and decay constants of the η and η' mesons in section 5.3.

The quark masses of the ensembles considered in this study follow two distinct trajectories, see figure 1: along one trajectory the average quark mass is kept constant [38], starting from the $N_f = 3$ symmetric point ($m_s = m_\ell$), while along the other the renormalized strange quark mass is held close to its physical value [39]. The two trajectories intersect close to the physical point where $12t_0\overline{M}^2 = 12t_0^{\mathrm{ph}}\overline{M}^{\mathrm{ph}\,2} = 1.11$. The pion masses, listed in table 1 along with the kaon masses, range from 422 MeV down to slightly below the physical mass.

We employ three lattice spacings, $a = 0.0859(12)\,\mathrm{fm}$, $a = 0.0761(10)\,\mathrm{fm}$ and $a = 0.0643(9)\,\mathrm{fm}$, with multiple pion and kaon masses, complemented by an additional ensemble at a finer lattice spacing $a = 0.0497(7)\,\mathrm{fm}$ [47]. Extrapolations to the physical point are performed using dimensionless combinations with t_0, determined on the same ensemble. After extrapolation, values for t_0 at the physical point (t_0^{ph}) and, for determinations of LECs, in the SU(3) chiral limit (t_0^χ) are required. Using the pion and kaon decay constants as input, [49] obtain

$$\left(8t_0^{\mathrm{ph}}\right)^{-1/2} = 475(6)\,\mathrm{MeV}. \tag{3.5}$$

[1] Note that $c^s_P = -16\pi^2 g_P/\sqrt{6}$ to leading order in g^2. This replacement will also affect the definitions of the mass dependent improvement coefficients $\tilde{d}_P$ and d_P, relative to [46].

JHEP08(2021)137

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 1. Details of the CLS ensembles analysed in this study. 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 [47], while $12t_0\overline{M}^2$ and $8t_0\delta M^2$ have been determined within this study (on a subset of the configurations analysed in [47]). Ensembles H102a and H102b were generated with the same quark masses and lattice coupling but different simulation parameters and are therefore analysed separately.

β	t_0^*/a^2	a/fm	a^{-1}/GeV
3.4	2.888(8)	0.0859(12)	2.296(33)
3.46	3.686(11)	0.0761(11)	2.594(38)
3.55	5.157(15)	0.0643(9)	3.068(44)
3.7	8.617(22)	0.0497(7)	3.966(57)

Table 2. The values of t_0^*/a^2 and the lattice spacings employed in this study, as determined in [47]. The lattice spacings were obtained using $(8t_0^{\mathrm{ph}})^{-1/2} \approx 0.475(6)\,\mathrm{GeV}$ [44].

For $t_0^\chi = 1.036(4)t_0^*$ [47], using $(8t_0^*)^{-1/2} = 0.478(7)\,\mathrm{GeV}$ [49, 50], we find

$$(8t_0^\chi)^{-1/2} = 470(7)\,\mathrm{MeV}. \tag{3.6}$$

The line along which we keep the sum of the quark masses constant intersects both the physical point and a point where $m_s = m_\ell$. Along this quark mass trajectory (in the continuum limit) $\overline{M}^2$ as well as t_0 are constant [51] to NLO in SU(3) ChPT. This motivates the definition of another scale [49], t_0^*, equating $12t_0^*\overline{M}^{*2} = 12t_0^{\mathrm{ph}}\overline{M}^{\mathrm{ph}\,2} = 1.11$. Naturally, this implies that $t_0^* \approx t_0^{\mathrm{ph}}$, however, determining t_0^* at each lattice spacing does not require an extrapolation to the physical point but just a small interpolation. We extracted values for t_0^*/a^2 (see table 2) from a global fit to a large number of CLS ensembles in [47] and we use these values to set the relative scale between the different lattice spacings.

All spatial volumes are considerably larger than $(2\,\mathrm{fm})^3$ and for most of the ensembles the dimensionless combination of the spatial lattice extent L_s and the mass of the pion M_π, L_sM_π, is larger than four.[2] On these ensembles only very mild finite volume effects have been observed for the non-singlet pseudoscalar masses and decay constants [47, 49]. Given the larger errors in the η-η' system compared to the pion, we are confident that for our volumes such effects can be neglected.

3.2 Wick contractions

The momentum projected pseudoscalar interpolators, introduced in section 2, are obtained by Fourier transforming spatially extended operators $\mathcal{P}^a(t,\vec{x})$ that have the same flavour structure as the local pseudoscalar currents P^a and are centred about $\vec{x}$:

$$\mathcal{P}^a_{\vec{p}}(t) = a^3 \sum_{\vec{x}} e^{-i\vec{p}\cdot\vec{x}}\mathcal{P}^a(t,\vec{x}). \tag{3.7}$$

Any linear combination of interpolators $\mathcal{P}^{a\dagger}_{\vec{p}}$ and $\mathcal{P}^{q\dagger}_{\vec{p}}$ can be applied to the vacuum $|\Omega\rangle$ to probe the physical eigenstates. We shall label such linear combinations as $\mathcal{B}_i^\dagger(-\vec{p},t)$. From these one can define matrices of correlation functions

$$C_{ij}(\vec{p},t) = \frac{1}{N_{t_{\mathrm{in}}}} \sum_{t_{\mathrm{in}}} \left\langle \Omega \left| \mathcal{B}_i(\vec{p},t+t_{\mathrm{in}})\mathcal{B}_j^\dagger(-\vec{p},t_{\mathrm{in}}) \right| \Omega \right\rangle, \tag{3.8}$$

where $N_{t_{\mathrm{in}}}$ denotes the number of source time slices that we average over.

[2] The only exceptions are H106 ($L_sM_\pi \approx 3.79$) and D150 ($L_sM_\pi \approx 3.51$).

JHEP08(2021)137

Performing the Wick contractions between the quark bilinears at the source and the sink leads to both connected and disconnected quark line contributions $\tilde{C}$ and $\tilde{D}$, respectively,

$$\overbrace{q^{f_1}\Gamma_1\bar{q}^{f_2}}\,\overbrace{\bar{q}^{f_3}\Gamma_2 q^{f_4}} = \delta_{f_1,f_3}\delta_{f_2,f_4}\tilde{C}^{\Gamma_1\Gamma_2}_{f_1,f_2} - \delta_{f_1,f_2}\delta_{f_3,f_4}\tilde{D}^{\Gamma_1\Gamma_2}_{f_1,f_3}, \tag{3.9}$$

where for brevity we only display the flavour indices $f_i \in \{u, d, s\}$ and the Dirac structures $\Gamma_i \in \{\gamma_5, \gamma_\mu\gamma_5\}$. The disconnected terms only contribute when contracting flavour-diagonal quark bilinears and are particularly challenging to compute.

3.3 Stochastic measurement of disconnected loops

The basic building blocks of disconnected correlation functions are quark loops (one-point functions)

$$L_s^{\Gamma,f}(\vec{p},t) = a^3 \sum_{\vec{x},\vec{y},\vec{z}} \mathrm{tr}\left(e^{-i\vec{p}\cdot\vec{x}}\phi^s(x,y)\Gamma D_f^{-1}(y,z)\phi^s(z,x)\right), \tag{3.10}$$

where $f \in \{\ell, s\}$ labels the flavour of the Dirac operator D_f. After performing the Wick contractions no distinction between the mass degenerate u and d quarks needs to be made. 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 $z = (t, \vec{z})$. To ensure ground state dominance, we apply the Wuppertal smearing kernel [52]

$$\phi(x,y) = \frac{1}{1+6\delta}\left(\delta(x,y) + \delta \sum_{j=\pm1}^{\pm3} U_j(x)\delta(x+a\hat{j},y)\right) \tag{3.11}$$

to the source and the sink of the quark propagators. Above, $U_j(x)$ is a (spatially APE smeared [53]) gauge link at x, pointing in the spatial direction $\hat{j}$ and $U_{-j}(x) = U_j^\dagger(x - a\hat{j})$. The number of smearing iterations s as well as the parameter $\delta = 0.25$ determine the smearing radius.

Since D_f^{-1} is a very large matrix, the trace in eq. (3.10) cannot be computed exactly but must be estimated stochastically. To do so, we start by constructing time-partitioned (also referred to as "diluted" in the literature) sources [54]:

$$\rho_{\tau,i}(x,\alpha,a) = \begin{cases} \frac{r_i(x,\alpha,a)}{\sqrt{2}} & \mathrm{mod}(t,\Delta_t) = \tau \text{ and } b \le t < L_t - b, \\ 0 & \text{otherwise} \end{cases}, \tag{3.12}$$

where L_t 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. These random sources span a space in the bulk of the lattice,

$$\sum_{\tau=0}^{\Delta_t/a-1}\sum_{i=0}^{N_{\mathrm{stoch}}-1} |\rho_{\tau,i}\rangle\langle\rho_{\tau,i}| = N_{\mathrm{stoch}}\ \mathrm{diag}(\underbrace{0,\ldots,0}_{b/a},\underbrace{1,\ldots,1}_{L_t/a-2b/a},\underbrace{0,\ldots,0}_{b/a}) \otimes \mathbb{1}_{12V_3/a^3} + \mathcal{O}(1/\sqrt{N_{\mathrm{stoch}}}), \tag{3.13}$$

$$a^4\langle\rho_{\tau,i}|\rho_{\tau',j}\rangle = (L_t - 2b)V_3\,\delta_{i,j}\delta_{\tau,\tau'}, \tag{3.14}$$

JHEP08(2021)137

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

$$D_f \sigma^f_{\tau,i} = \rho_{\tau,i} \tag{3.15}$$

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 $\sigma^f_{\tau,i}$.

Summing over the dilution index τ, we obtain the estimate of the loop computed on the i-th stochastic source,

$$L^{\Gamma,f}_{i,s}(\vec{p},t) = a^3 \sum_{\vec{x}} \sum_{\tau=0}^{\Delta_t/a-1} \operatorname{tr}\left(e^{-i\vec{p}\cdot\vec{x}}(\rho_{\tau,i}\phi^s)(x)\Gamma(\sigma^f_{\tau,i}\phi^s)(x)\right). \tag{3.16}$$

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

$$L^{\Gamma,f}_{s}(\vec{p},t) = \frac{1}{N_{\text{stoch}}} \sum_{i=0}^{N_{\text{stoch}}-1} L^{\Gamma,f}_{i,s}(\vec{p},t) + \mathcal{O}\left(\frac{1}{\sqrt{N_{\text{stoch}}}}\right). \tag{3.17}$$

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 in eq. (3.16).

The inverse of the Wilson Dirac operator $D_f = \frac{1}{2\kappa_f}(\mathbb{1} - \kappa_f H)$ within the trace of eq. (3.10) 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 [55–57],

$$\operatorname{tr}\left(\Gamma D_f^{-1}\right) = 2\kappa_f \sum_{i=0}^{\infty} \kappa_f^i \operatorname{tr}\left(\Gamma H^i\right) = 2\kappa_f \sum_{i=0}^{n-1} \kappa_f^i \operatorname{tr}(\Gamma H^i) + \kappa_f^n \operatorname{tr}\left(\Gamma H^n D_f^{-1}\right). \tag{3.18}$$

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 $\operatorname{tr}\Gamma H^i = 0$ and a remainder. Note that the value of n depends on Γ and the fermion action employed. 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 only contributes to the noise. Hence, we can obtain an improved estimate of the trace, by applying the Dirac operator n times to the solution, replacing

$$\Gamma\sigma^f_{\tau,i} \mapsto \Gamma(1 - 2\kappa_f D_f)^{n(\Gamma)}\sigma^f_{\tau,i} \tag{3.19}$$

in eq. (3.16).

In summary, we apply two noise reduction techniques, time partitioning which eliminates noise from neighbouring time slices at the expense of additional inversions and the hopping parameter expansion that also reduces short-distance noise by exploiting the locality of the Dirac operator. The latter is only applied to unsmeared loops and is most

JHEP08(2021)137

effective for small values of the hopping parameter (corresponding to a large quark mass). We remark that due to the use of a highly efficient multigrid solver [58–60], we do not benefit from the truncated solver method [57] within our setup. This method involves computing many (computationally cheap) approximate solves, each of which needs to be contracted according to eq. (3.16), also applying smearing. The latter would dominate the cost even though the implementation 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.

3.4 Measurements of connected and disconnected correlation functions

In order to estimate the disconnected two-point function appearing in eq. (3.9), we correlate and average two of the loops, defined in eq. (3.16):

$$\tilde{D}^{\Gamma_1\Gamma_2}_{\substack{f_1,f_2\\s_1,s_2}}(\vec{p},t) = \frac{1}{N_{\text{stoch}}(N_{\text{stoch}}-1)}\frac{a}{L_t}\sum_{\substack{i,j=0\\i\neq j}}^{N_{\text{stoch}}-1}\sum_{t_{\text{in}}=0}^{L_t-a}\left\langle L^{\Gamma_1,f_1}_{i,s_1}(\vec{p},t_{\text{in}}+t)L^{\Gamma_2,f_2}_{j,s_2}(-\vec{p},t_{\text{in}})\right\rangle . \quad (3.20)$$

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 (3.20) 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:

$$\overline{D}^{\Gamma_1\Gamma_2}_{\substack{f_1,f_2\\s_1,s_2}}(\vec{p},t) = \frac{1}{2}\left(\tilde{D}^{\Gamma_1\Gamma_2}_{\substack{f_1,f_2\\s_1,s_2}}(\vec{p},t) + \text{sgn}(\Gamma_1,\Gamma_2)\tilde{D}^{\Gamma_2\Gamma_1}_{\substack{f_2,f_1\\s_2,s_1}}(\vec{p},t)\right), \quad (3.21)$$

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 expression (3.9) also contains the connected correlation function matrix $\tilde{C}$. This can be computed by exploiting the γ_5-Hermiticity of the Dirac operator and inverting on a smeared point source $P^s_{\alpha,a}(x) = \phi^s(x,y)\delta(y,x_{\text{in}})\delta(\alpha,\alpha_{\text{in}})\delta(a,a_0)$. This yields the point-to-all propagator $M^s_f(x) = D^{-1}_f(x,y)P^s(y)$, which is a matrix in Dirac and colour space and can be used to construct connected correlation functions, starting at a fixed source position x_{in}:

$$\tilde{C}^{\Gamma_1,\Gamma_2}_{\substack{f_1,f_2\\s_1,s_2}}(\vec{p},t) = V_3a^3\sum_{\vec{x}} e^{-i\vec{p}\cdot(\vec{x}-\vec{x}_{\text{in}})}\left\langle \Gamma_1\gamma_5\phi^{s_2}M^{s_1}_{f_1}(x)\gamma_5\Gamma_2\phi^{s_2}M^{s_1}_{f_2}(x)\right\rangle . \quad (3.22)$$

In our case, $f_1 = f_2$. Due to the reduced error compared to the disconnected correlation functions, it is sufficient to employ only a single source position per configuration at $x_{\text{in}} =$

JHEP08(2021)137

$(b, \vec{0})$, 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. (3.21).

Based on previous experience with the pion correlation function regarding the boundary effects [47], we fix b such that these effects are smaller than the statistical error of the pion correlation function at the corresponding timeslice b/a. This is a conservative choice, given the comparably large errors of the disconnected contributions.

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. In the measurement of the disconnected loops, we choose a time separation of $\Delta_t = 4a$ in the stochastic dilution, except for the finest lattice spacing where we used $\Delta_t = 6a$. These choices are listed in table 3, as well as the number of configurations analysed and the distance between consecutive measurements in Hybrid Monte-Carlo molecular dynamics time units.

4 Extraction of mass eigenstates and local matrix elements

The masses of the η and η' are extracted from the matrix of correlation functions $C(t)$, eq. (3.8). Usually, this is done by solving a generalized eigenvalue problem (GEVP) and fitting to the resulting eigenvalues [61, 62]. 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. 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. In addition, 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.

4.1 Fitting to matrices of correlation functions

In the limit of infinite statistics, C (eq. (3.8)) is a real symmetric positive-definite $M \times M$ matrix. The spectral decomposition gives

$$C_{ij}(t) = \sum_{n=0}^{\infty} \frac{1}{2E_n V_3} \left\langle \Omega \left| \mathcal{B}_i(t) \right| n \right\rangle \left\langle n \left| \mathcal{B}_j^\dagger(0) \right| \Omega \right\rangle , \tag{4.1}$$

$$= \sum_{n=0}^{\infty} \frac{1}{2E_n V_3} \exp(-E_n t) \left\langle \Omega \left| \mathcal{B}_i(0) \right| n \right\rangle \left\langle n \left| \mathcal{B}_j^\dagger(0) \right| \Omega \right\rangle , \tag{4.2}$$

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 (4.1) can be written as

$$C(t) = \widehat{Z} \widehat{D}(t) \widehat{Z}^\intercal , \tag{4.3}$$

JHEP08(2021)137

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 3. 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.

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

$$\widehat{Z}_{\text{in}} = \frac{1}{\sqrt{2E_n V_3}} \langle \Omega | \mathcal{B}_i(0) | n \rangle \tag{4.4}$$

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,

$$C(t) = Z D(t) Z^{\intercal} + \mathcal{O}(\exp(-E_N t)), \tag{4.5}$$

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

JHEP08(2021)137

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 $C(t)$. It is therefore important to include at least a third state, $N \geq 3$, in our analysis to provide an effective parametrization of the contributions of these 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 % [63]). 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.

One can use eq. (4.5) 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 table 4. 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 section 3.4. This allows the simple ansatz $D(t) = \text{diag}(\exp(-E_n t))$ for the time dependent matrix in eq. (4.5). 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:

$$D(t) = \text{diag}\left[2\exp\left(-E_n\frac{L_t}{2}\right)\cosh\left(-E_n\left(t-\frac{L_t}{2}\right)\right)\right]. \tag{4.6}$$

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

$$aE_n(\vec{p}) = \sqrt{a^2M_n^2 + a^2\vec{p}^2}, \tag{4.7}$$

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. In figure 2 we show examples of these fits (see also figure 3 of [64]). In addition, we plot the naive lattice dispersion relation for a free scalar particle,

$$aE_n(\vec{p}) = \text{arccosh}\left(\cosh(aM_n) + \sum_j 2\sin^2(ap_j/2)\right). \tag{4.8}$$

Since within the relevant momentum range the differences between the two curves (4.7) and (4.8) are much smaller than the errors of the data, we conclude that assuming eq. (4.7)

JHEP08(2021)137

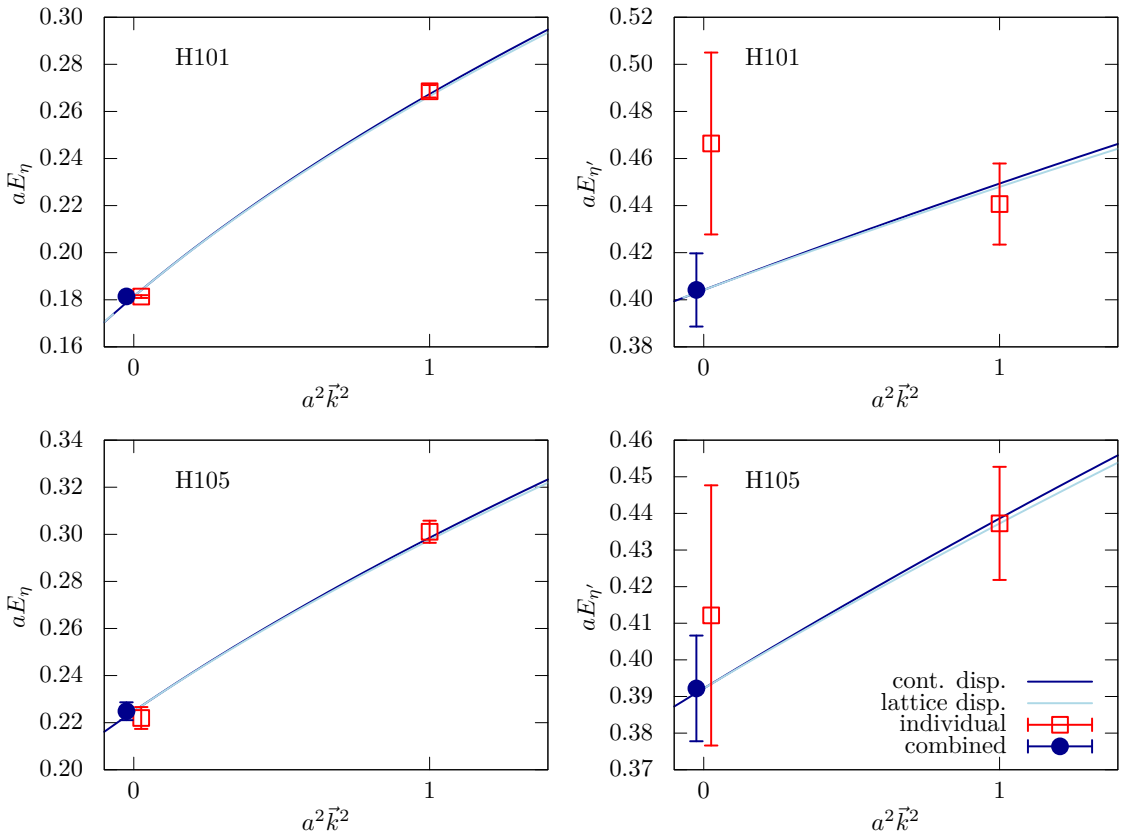


Figure 2. Energies of the η (left panels) and η' (right panels) mesons determined on ensembles H101 (top) and H105 (bottom). The red squares display the energies with $a^2\vec{k}^2 = 0$ and $a^2\vec{k}^2 = 1$ extracted from individual fits, while the blue filled symbols show the masses determined from 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.

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 [18]. 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.

4.2 The generalized effective mass method

As suggested in [65, 66], it is advantageous to reduce the correlations between time slices by fitting to the temporal derivative of the correlation functions. The fit form, eq. (4.3), is

JHEP08(2021)137

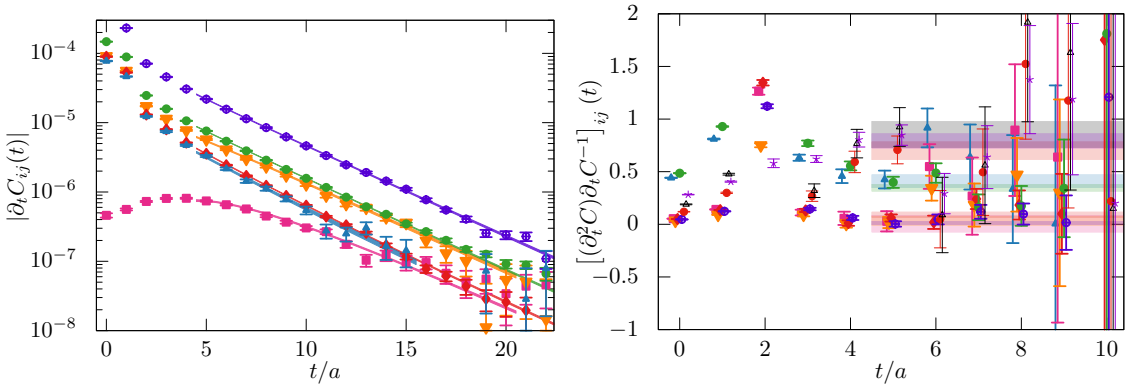


Figure 3. (Left) The moduli of $C_{ij}(t)$ determined on ensemble H105 at momentum $a^2\vec{k}^2 = 1$. (Right) The elements of the generalized effective mass matrix, eq. (4.15). The shaded regions in both panels correspond to the results of a simultaneous fit to eqs. (4.9) and (4.15). The widths of the regions indicate the fit ranges.

modified to

$$\partial_t C(t) \sim Z\left(\partial_t D(t)\right) Z^\intercal, \tag{4.9}$$

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

$$\partial_t D(t) = -\operatorname{diag}\left[E_n \exp\left(-E_n t\right)\right] \qquad \text{(open boundaries)}, \tag{4.10}$$

$$\partial_t D(t) = -2\operatorname{diag}\left[E_n \exp\left(-E_n \frac{L_T}{2}\right) \sinh\left(-E_n\left(t - \frac{L_T}{2}\right)\right)\right] \qquad \text{(periodic boundaries)}. \tag{4.11}$$

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., [18, 19]) are automatically removed. Although we do not encounter any significant shifts within our data, we observe that utilizing eq. (4.9) leads to decreased correlations and more stable fit results.

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 $N = M$),

$$\partial_t \log C(t) = \left(\partial_t C(t)\right) C^{-1}(t) \tag{4.12}$$

$$= \left(\widehat{Z}\partial_t \widehat{D}(t)\widehat{Z}^\intercal\right)\left(\widehat{Z}\widehat{D}(t)\widehat{Z}^\intercal\right)^{-1} \tag{4.13}$$

$$= -Z\ \operatorname{diag}_{n=0}^{N-1}(E_n)\, Z^{-1} + \mathcal{O}\left[\exp\left(-\left(E_N - E_{N-1}\right)t\right)\right], \tag{4.14}$$

JHEP08(2021)137

which is constant in time (up to excited states corrections and statistical noise).[3] 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

$$(\partial_t^2 C)(\partial_t C)^{-1} = -Z\ \mathrm{diag}(E_n)\, Z^{-1} + \mathcal{O}\left[\exp\left(-\left(E_N - E_{N-1}\right) t\right)\right]. \tag{4.15}$$

This alters the contributions from higher excited states but leaves the large-time behaviour 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 $N = M$.[4] In this case, simultaneous fits are performed to eqs. (4.9) and (4.15), where the latter enables the fit to unambiguously resolve the spectrum of states.

To summarize our fitting strategy: we simultaneously fit the correlation functions with two momenta $a^2\vec{k}^2 = 0$ and $a^2\vec{k}^2 = 1$ to eqs. (4.9) and (4.15). 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 figure 3. The resulting η and η' masses for all the ensembles are collected in table 4, along with the χ^2/N_{df} of the fits, where in most cases we achieve $\chi^2/N_{\mathrm{df}} \approx 1$.

4.3 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 [61, 62],

$$C(t)V(t,t_0) = C(t_0)V(t,t_0)\Lambda(t,t_0), \tag{4.16}$$

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 [67] 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 figure 4. 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.

[3] 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.

[4] 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. (4.9) and (4.15).

JHEP08(2021)137

id	basis	$a\,M_\eta$	$\sqrt{8t_0}M_\eta$	$a\,M_{\eta'}$	$\sqrt{8t_0}M_{\eta'}$	$\chi^2/N_{\rm df}$
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. Masses of the η and η' mesons obtained from fits to eqs. (4.9) and (4.15) in lattice units and in units of the gradient flow scale, $\sqrt{8t_0}$ (determined on the same ensemble). See table 1 for the corresponding pion and kaon masses and table 2 for the lattice spacings. We also give the smearing bases used in the construction of the matrix of correlation functions, eq. (4.1), 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 table 3. The resulting $\chi^2/N_{\rm df}$ 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 $\chi^2/N_{\rm df}$ 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.

JHEP08(2021)137

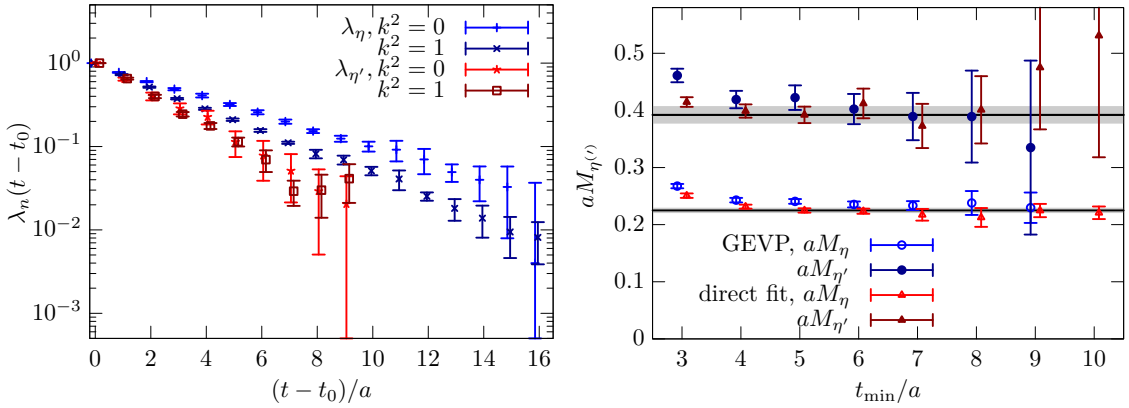


Figure 4. (Left) The η and η' eigenvalue functions for two momenta, $a^2\vec{k}^2 = 0$ and $a^2\vec{k}^2 = 1$ determined on ensemble H105, obtained by solving the GEVP, with the reference time slice $t_0 = 5a$. For better visibility, some points are shifted slightly and the data for the η' are omitted for $(t - t_0)/a \geq 10$ due to the large errors. (Right) Results for the masses determined by fitting to the lowest two eigenvalues of the GEVP and from direct fits (eqs. (4.9) and (4.15)) 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, $t_0/a = t_{\min}/a - 1$.

4.4 Decay constants and local matrix elements

The decay constants of pseudoscalar mesons are defined in eq. (2.6). The singlet ($a = 0$) and octet ($a = 8$) decay constants of the η ($n = 0$) and η' ($n = 1$) mesons can be obtained via

$$\left\langle \Omega \left| A_\mu^a \right| n \right\rangle = i F_n^a p_\mu, \tag{4.17}$$

where the local currents $A_\mu^a = \overline{\psi} t^a \gamma_\mu \gamma_5 \psi$ are introduced in eqs. (2.2)–(2.5). In addition to the singlet/octet basis, one can also define decay constants in the flavour basis according to eq. (2.7). The two sets of decay constants are related via a SO(2) rotation (see eq. (2.11)):

$$\begin{pmatrix} F_n^\ell \\ F_n^s \end{pmatrix} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & \sqrt{2} \\ -\sqrt{2} & 1 \end{pmatrix} \begin{pmatrix} F_n^8 \\ F_n^0 \end{pmatrix}. \tag{4.18}$$

It is also useful to define the pseudoscalar matrix elements,

$$\langle \Omega | P^a | n \rangle = H_n^a. \tag{4.19}$$

We remark that for the two-point functions one can write (in the limit of large times)

$$\begin{aligned} \langle \Omega \left| \partial_t P^a(t) \mathcal{B}_i(0) \right| \Omega \rangle &= \frac{1}{2a} \langle \Omega \left| (P(t+a) - P(t-a)) \, \mathcal{B}_i(0) \right| \Omega \rangle \\ &= -\frac{\sinh(aE_n)}{a} \langle \Omega \left| P^a(t) \mathcal{B}_i(0) \right| \Omega \rangle \\ &\approx -E_n \langle \Omega \left| P^a(t) \mathcal{B}_i(0) \right| \Omega \rangle , \end{aligned} \tag{4.20}$$

JHEP08(2021)137

where E_n is the energy of the lowest state with a non-vanishing overlap $Z_{\text{in}} = \langle\Omega|\mathcal{B}_i|n\rangle \neq 0$ (as well as $\langle\Omega|P^a|n\rangle \neq 0$). Using appropriate linear combinations of the interpolators $\mathcal{B}_i$ as described above, it is possible to project onto $n = 0$ (η) as well as onto $n = 1$ (η'). In this sense, we can substitute $\langle\Omega|\partial_t P^a|n\rangle$ for $-E_n H_n^a$, up to order a^2 corrections.

We define partially $\mathcal{O}(a)$ improved, unrenormalized lattice decay constants:

$$\tilde{F}_n^a E_n = \langle\Omega|A_0^a + a c_A^a \partial_t P^a|n\rangle \tag{4.21}$$

with the improvement coefficients $c_A^a = c_A$ for $a \neq 0$ and $c_A^a = c_A^s$ for $a = 0$. For the moment being, mass dependent order a terms are ignored. While the non-singlet improvement coefficient c_A has been determined non-perturbatively [68], its singlet equivalent c_A^s is unknown and we parameterize it as $c_A^s = c_A + \delta c_A$, where δc_A is of order g^4. For the singlet case we can rewrite eq. (4.21) as

$$\tilde{F}_n^0 = \check{F}_n^0 + a\,\delta c_A \check{H}_n^0, \tag{4.22}$$

$$\check{F}_n^0 E_n = \langle\Omega|A_0^0 + a c_A \partial_t P^0|n\rangle, \tag{4.23}$$

$$\check{H}_n^0 = \langle\Omega|P^0|n\rangle, \tag{4.24}$$

where even in the chiral limit $\check{F}_n^0$ is only partially order a-improved since we have neglected the difference δc_A between the singlet and the non-singlet improvement coefficients. We also introduced the singlet pseudoscalar matrix element $\check{H}_n^0$. We note that $\check{H}_n^8$ is already order a improved in the chiral limit ($\check{H}_n^8 = \tilde{H}_n^8$), however, this is not the case for the singlet pseudoscalar current [46], see eq. (3.3), where we have to distinguish between $\tilde{H}_n^0$ and $\check{H}_n^0$:

$$\tilde{H}_n^0 = \check{H}_n^0 + a c_P^s \left\langle \Omega \left| \partial_\mu A_\mu^0 \right| n \right\rangle. \tag{4.25}$$

Note that the above $\mathcal{O}(a)$ difference between $\tilde{H}_n^0$ and $\check{H}_n^0$ does not affect eq. (4.22), where only the unimproved matrix element is needed.

4.5 Determination of the decay constants

In order to extract matrix elements with a local current J, we start from a vector of M correlation functions ($i = 1, \ldots, M$):

$$C_i^J(t) = \left\langle \Omega \left| J(t) \mathcal{B}_i^\dagger(0) \right| \Omega \right\rangle, \tag{4.26}$$

where $\mathcal{B}_i(0)$ is an interpolator with the momentum $\vec{p}$ inserted at the time $t_{\text{in}} = 0$. 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 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. (3.16) and (3.20), however, without smearing at the sink and with the additional normalization factor $1/V_3$.

JHEP08(2021)137

We carry out a spectral decomposition similar to that of eq. (4.1):

$$C_i^J(t) \approx \sum_{n=0}^{N-1} \frac{1}{2E_n V_3} \exp(-E_n t) \left\langle \Omega \left| J(0) \right| n \right\rangle \left\langle n \left| \mathcal{B}_i^\dagger(0) \right| \Omega \right\rangle$$
$$= \sum_{n=0}^{N-1} Z_{\text{in}} D_{nn}(t) j_n, \tag{4.27}$$

where

$$j_n = \frac{1}{\sqrt{2E_n V_3}} \langle \Omega | J | n \rangle, \tag{4.28}$$

Z is the overlap matrix $\widehat{Z}$, defined in eq. (4.4), truncated to N columns and $D(t) = \text{diag}(\exp(-E_n t))$ is a diagonal $N \times N$ matrix. 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 $\tilde{F}_n^8$ and $\check{F}_n^0$ for $n \in \{\eta, \eta'\}$ as well as the corresponding pseudoscalar matrix elements $\tilde{H}_n^8$ and $\check{H}_n^0$. Below we will discuss the remaining improvement and renormalization steps and we will add any missing improvement coefficients, e.g., δc_A, as fit parameters in the continuum limit extrapolation.

5 Physical point and continuum extrapolation

In this section we motivate the parametrizations of the quark mass and lattice spacing dependence of our data and present continuum limit results on the η and η' meson masses and their respective decay constants. We give the physical point values as well as the values of the NLO large-N_c ChPT LECs, and provide a detailed study of their statistical and systematic uncertainties.

First, in section 5.1 we detail the renormalization and $\mathcal{O}(a)$ improvement of the octet and singlet decay constants. This affects the functional form of our continuum limit extrapolation since not all the improvement coefficients are known. Different possibilities exist regarding the renormalization scheme for the singlet decay constants. This will be discussed in section 5.2. In the continuum limit, large-N_c U(3) ChPT will be used to describe the mass dependence of the data. We summarize the relevant LO and NLO ChPT expectations in section 5.3. We combine these continuum limit functions with a parametrization of the remaining $\mathcal{O}(a)$ and $\mathcal{O}(a^2)$ lattice spacing effects in section 5.4, while in section 5.5 we carry out several fits to our data in order to quantify the various systematic uncertainties. A central fit is used to predict the values for the masses and decay constants at the physical point and the systematic errors are estimated by varying the fit form. The results are presented in section 5.6. The parameters that describe the continuum limit behaviour correspond to the LECs of NLO large-N_c ChPT. Our estimates of their values are given in section 5.7.

JHEP08(2021)137

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 5. Renormalized and partially improved octet and singlet decay constants of the η and η' mesons obtained from fits to eq. (4.27) 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.

5.1 Renormalization and $\mathcal{O}(a)$ improvement

The $\mathcal{O}(a)$ improvement of quark bilinears has been worked out for Wilson fermions in [46]. It turns out (see eq. (15) of this reference) that even for the improvement of the octet axialvector current, singlet currents are required. The renormalized $\mathcal{O}(a)$ improved octet decay constant reads

$$F^8_{\eta^{(\prime)}} = Z_A\Big[(1+3a\tilde{b}_A\overline{m})\tilde{F}^8_{\eta^{(\prime)}} + \frac{a}{\sqrt{3}}b_A(m_\ell\check{F}^\ell_{\eta^{(\prime)}} - \sqrt{2}m_s\check{F}^s_{\eta^{(\prime)}}) - \sqrt{2}af_A\delta m\check{F}^0_{\eta^{(\prime)}}\Big] + \mathcal{O}(a^2), \tag{5.1}$$

JHEP08(2021)137

where Z_A is the renormalization factor for non-singlet currents and $\tilde{b}_A$,[5] b_A and f_A are coefficients of mass dependent improvement terms. 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 section 5.4. 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 (4.18). The bare quark masses ($f = \ell, s$) and their average and difference are given by

$$am_f = \frac{1}{2}\left(\frac{1}{\kappa_f} - \frac{1}{\kappa_{\rm cr}}\right), \quad a\overline{m} = \frac{a}{3}\left(2m_\ell + m_s\right), \quad a\delta m = am_s - am_\ell. \tag{5.2}$$

The critical hopping parameter $\kappa_{\rm cr}$ was determined for our action and lattice spacings in [39, 47].

We use the non-perturbatively determined values of Z_A that can be found in the $Z^l_{A,\rm sub}$ column of table 6 of [71]. The improvement coefficients b_A and $\tilde{b}_A$ have been determined non-perturbatively in [69, 70]. The sea quark coefficients, $\tilde{b}_A$ for our β values are

$$\begin{aligned} \tilde{b}_A(\beta = 3.4) &= -0.11(13), \qquad \tilde{b}_A(\beta = 3.46) = 0.10(11), \\ \tilde{b}_A(\beta = 3.55) &= -0.04(12), \qquad \tilde{b}_A(\beta = 3.70) = -0.05(8). \end{aligned} \tag{5.3}$$

For the valence quark coefficient b_A we use the parametrization [70]

$$b_A(g^2) = 1 + 0.0881\, C_F g^2 + b\, g^4, \quad \text{where} \quad b = 0.0113(44) \tag{5.4}$$

and $g^2 = 6/\beta$ and $C_F = 4/3$. The improvement term in eq. (5.1) 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 $\mathcal{O}(g^6)$ [46, 72]. 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 section 5.4.

Regarding the improvement of the singlet decay constants, utilizing eq. (23) of [46], we obtain

$$F^0_{\eta^{(\prime)}} = Z^s_A \left[\left(1 + 3a\tilde{d}_A\overline{m}\right)\tilde{F}^0_{\eta^{(\prime)}} + \frac{1}{\sqrt{3}} a d_A \left(\sqrt{2} m_l \check{F}^{\ell}_{\eta^{(\prime)}} + m_s \check{F}^{s}_{\eta^{(\prime)}}\right)\right]. \tag{5.5}$$

Again, we replaced the lattice decay constants within the $\mathcal{O}(a)$ terms by partially improved ones. The renormalization factor $Z^s_A \neq Z_A$ is discussed in the next subsection. 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 eqs. (4.22)–(4.24)). The results for the renormalized but only partially improved η and η' singlet and octet decay constants (derived from the fits presented in section 4.5) are given in table 5.

[5] We simulate at constant values of the unimproved lattice coupling parameter g^2. The difference with respect to keeping the $\mathcal{O}(a)$ improved coupling fixed amounts to replacing the improvement coefficients $\bar{b}_J$ and $\bar{d}_J$ of [46] by $\tilde{b}_J$ and $\tilde{d}_J$. The relation between these sets of parameters is detailed in [69] and the differences turn out to be tiny at our lattice spacings [70].

JHEP08(2021)137

5.2 Renormalization of the singlet axialvector current

The renormalization factor of the singlet axialvector current Z_A^s in the standard $\overline{\rm MS}$ scheme [40, 73] depends on the renormalization scale. For definiteness, we detail our conventions for the QCD β-function and the anomalous dimension of a current J:

$$\beta(a_s) = \mu^2 \frac{{\rm d}a_s}{{\rm d}\mu^2} = -\sum_{n\geq 0} \beta_n a_s^{n+2}, \tag{5.6}$$

$$\gamma_J(a_s) = \mu^2 \frac{{\rm d}\ln J(a_s,\mu)}{\mu^2} = \mu^2 \frac{{\rm d}Z_J}{{\rm d}\mu^2} = -\sum_{n\geq 0} \gamma_n a_s^{n+1}, \tag{5.7}$$

where $a_s = \alpha_s/\pi = g_s^2/(4\pi^2)$ and the renormalized and bare currents $\widehat{J}$ and J_0, respectively, are related by $\widehat{J}(\mu) = Z_J(\mu) J_0$. Using these normalizations, the first two β-function coefficients read

$$\beta_0 = \frac{1}{4}\left(\frac{11}{3}C_A - \frac{2}{3}N_f\right), \tag{5.8}$$

$$\beta_1 = \frac{1}{16}\left[\frac{34}{3}C_A^2 - \left(\frac{10}{3}C_A - \frac{5}{2}C_F\right)N_f\right], \tag{5.9}$$

while the first three γ-function coefficients for the singlet axialvector current are [40, 73][6]

$$\gamma_{A0}^s = 0, \tag{5.10}$$

$$\gamma_{A1}^s = \frac{3}{8}C_F N_f, \tag{5.11}$$

$$\gamma_{A2}^s = \frac{1}{64}\left[\left(\frac{142}{3}C_F C_A - 18 C_F^2\right)N_f - \frac{4}{3}C_F N_f^2\right]. \tag{5.12}$$

In QCD $C_A = 3$ and $C_F = 4/3$. Note that γ_A^s vanishes for $N_f = 0$ since the anomalous dimension is a sea quark effect.

From the β- and γ-functions one can easily derive the scale evolution of local currents:

$$Z_J\left(a_s(\mu_1),\mu_1\right) = Z_J\left(a_s(\mu_0),\mu_0\right) \exp\left(\int_{a_s(\mu_0)}^{a_s(\mu_1)} {\rm d}a\, \frac{\gamma_J(a)}{\beta(a)}\right). \tag{5.13}$$

Normally, to leading non-trivial order, the evolution factor is given by $(a_s(\mu_1)/a_s(\mu_0))^{\gamma_{J0}/\beta_0}$, which diverges if one of the scales is sent to infinity. In our case, however, $\gamma_{A0}^s = 0$, leading to a finite renormalization group evolution

$$\frac{Z_A^s\left(a_s(\mu_1),\mu_1\right)}{Z_A^s\left(a_s(\mu_0),\mu_0\right)} = \exp\Bigg\{\frac{\gamma_{A1}^s}{\beta_0}\Bigg[\left(a_s(\mu_1) - a_s(\mu_0)\right) + \frac{1}{2}\left(\frac{\gamma_{A2}^s}{\gamma_{A1}^s} - \frac{\beta_1}{\beta_0}\right)\left(a_s^2(\mu_1) - a_s^2(\mu_0)\right) + \ldots\Bigg]\Bigg\}. \tag{5.14}$$

This suggests a modified scheme (see, e.g., [75]), where the renormalization group running is absorbed into the renormalization constant:

$$\widehat{A}'_\mu = \left[Z_A^s(\infty)/Z_A^s(\mu)\right]\widehat{A}_\mu = Z_A^{s\prime} A_{0\mu}, \tag{5.15}$$

[6] Recently, γ_{A3}^s has been computed too [74].

JHEP08(2021)137

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 6. Conversion factors $Z_A^s(\mu)/Z_A^{s\prime} = Z_A^s(\mu)/Z_A^s(\infty)$, computed according to eq. (5.14) 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 [44].

and

$$Z_A^{s\prime} = Z_A^s(\mu = \infty) = \left[1 - \frac{\gamma_{A1}^s}{\beta_0} a_s(\mu) + \frac{\gamma_{A1}^s}{2\beta_0}\left(\frac{\gamma_{A1}^s}{\beta_0} + \frac{\beta_1}{\beta_0} - \frac{\gamma_{A2}^s}{\gamma_{A1}^s}\right) a_s^2(\mu) + \cdots\right] Z_A^s(\mu). \tag{5.16}$$

Similar to the renormalization group invariant (RGI) scheme, in the above $\overline{\text{MS}}'$ scheme the current is scale independent and the corresponding γ-function is trivial: $\gamma_A^{s\prime} = 0$. However, there are two differences: it remains a variant of the $\overline{\text{MS}}$ scheme and there is no multiplicative ambiguity in the definition of $Z_A^{s\prime}$. Renormalizing the singlet axialvector current in this way corresponds to the usual convention, setting $\mu = \infty$.

At present, the difference $Z_A^{s\prime} - Z_A$ has only been computed in perturbation theory. Setting c_{SW} to its leading order value $c_{\text{SW}} = 1$ within eq. (32) of [76], we obtain for our action

$$Z_A^s(\mu) = Z_A - a_s^2(a^{-1})\left[\gamma_{A1}\ln(\mu^2 a^2) + 2.834(11)\right], \tag{5.17}$$

where again we use the non-perturbative Z_A values of [71]. 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 $\overline{\text{MS}}$ 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 $\mu = a^{-1}$ as our central value. The results are then run via eq. (5.14) (not eq. (5.16)) to $\mu = \infty$ to obtain the scale independent $\overline{\text{MS}}'$ result. This is carried out using the three-loop γ_A^s-function and, for the running of $a_s(\mu)$, starting from the value determined in [50], the five-loop β-function [77] (as implemented in version 3 of the RunDec package for Mathematica [78, 79]). For convenience we also quote results in the more commonly used scale dependent prescription at the scales $\mu = 10\,\text{GeV}$, $\mu = 2\,\text{GeV}$ and $\mu = 1\,\text{GeV}$ in QCD with $N_f = 3$ active quark flavours. The corresponding conversion factors are listed in table 6.

5.3 Fit form for the chiral extrapolation

We summarize the results of large-N_c ChPT that will be used to parameterize 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.

JHEP08(2021)137

Conventional SU(3) ChPT entails expansions in the masses of the octet mesons (the pions, the kaons and the octet η) — the Goldstone bosons of the spontaneous breaking of $SU(3)_A$ in the QCD vacuum. To include the singlet η, one extends the symmetry group to U(3) and expands simultaneously around the limit $N_c \to \infty$, in which the axial anomaly vanishes. 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 [42, 80], where the power counting is as follows:

$$p = \mathcal{O}(\sqrt{\delta}), \qquad m = \mathcal{O}(\delta), \qquad 1/N_c = \mathcal{O}(\delta) \tag{5.18}$$

with p being the momentum and m a quark mass.

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 Gell-Mann-Oakes-Renner (GMOR) relations, $\mu_{ab}^2 = 2B_0\,\mathrm{tr}[t^a\,\mathrm{diag}(m_\ell, m_\ell, m_s)t^b] = \delta^{ab}\mu_a^2$, 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. However, at this order in δ one also has to add the $\mathcal{O}(1/N_c)$ Witten-Veneziano contribution [1, 2] to the singlet mass, $M_0^2 = 2N_f\tau_0/F^2$, where τ_0 denotes the quenched topological susceptibility. Moreover, for non-degenerate quark masses, the singlet ($a = 0$) and the octet ($a = 8$) pseudoscalar mesons will mix and the corresponding non-diagonal part of the pseudoscalar mass matrix reads, see, e.g., [42, 80]:

$$\mu^2 = \begin{pmatrix} \mu_8^2 & \mu_{80}^2 \\ \mu_{80}^2 & \mu_0^2 \end{pmatrix}. \tag{5.19}$$

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

$$R\mu^2 R^\intercal = \begin{pmatrix} M_\eta^2 & 0 \\ 0 & M_{\eta'}^2 \end{pmatrix}. \tag{5.20}$$

R is an orthogonal transformation

$$R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{5.21}$$

which defines the so-called mass mixing angle θ. One can easily read off the relations

$$\mu_8^2 = M_\eta^2\cos^2\theta + M_{\eta'}^2\sin^2\theta, \tag{5.22}$$

$$\mu_0^2 = M_\eta^2\sin^2\theta + M_{\eta'}^2\cos^2\theta, \tag{5.23}$$

$$\mu_{80}^2 = (M_{\eta'}^2 - M_\eta^2)\sin\theta\cos\theta, \tag{5.24}$$

$$\theta = \frac{1}{2}\arcsin\left(\frac{2\mu_{80}^2}{\sqrt{(\mu_8^2-\mu_0^2)^2+4\mu_{80}^4}}\right) = \frac{1}{2}\arcsin\left(\frac{2\mu_{80}^2}{M_-^2}\right), \tag{5.25}$$

where

$$M_-^2 = M_{\eta'}^2 - M_\eta^2 = \sqrt{(\mu_8^2-\mu_0^2)^2+4\mu_{80}^4}, \tag{5.26}$$

$$M_+^2 = M_{\eta'}^2 + M_\eta^2 = \mu_8^2 + \mu_0^2. \tag{5.27}$$

JHEP08(2021)137

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

$$M_\eta = \sqrt{\frac{1}{2}\left(M_+^2 - M_-^2\right)} \quad \text{and} \quad M_{\eta'} = \sqrt{\frac{1}{2}\left(M_+^2 + M_-^2\right)} \tag{5.28}$$

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. We simulate with quark masses chosen to follow two distinct quark mass trajectories to the physical point, along one of which the average quark mass is held constant. A more convenient parametrization is in terms of the average and difference of the squared pion and kaon masses,

$$\overline{M}^2 = \frac{1}{3}\left(2M_K^2 + M_\pi^2\right) \approx 2B_0\overline{m}, \tag{5.29}$$

$$\delta M^2 = 2(M_K^2 - M_\pi^2) \approx 2B_0\delta m. \tag{5.30}$$

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 [41, 42] (see eq. (2.12)),

$$\begin{pmatrix} F_\eta^8 & F_\eta^0 \\ F_{\eta'}^8 & F_{\eta'}^0 \end{pmatrix} = \begin{pmatrix} F^8\cos\theta_8 & -F^0\sin\theta_0 \\ F^8\sin\theta_8 & F^0\cos\theta_0 \end{pmatrix}, \tag{5.31}$$

leading to

$$F^8 = \sqrt{(F_\eta^8)^2 + (F_{\eta'}^8)^2}, \qquad F^0 = \sqrt{(F_\eta^0)^2 + (F_{\eta'}^0)^2}, \tag{5.32}$$

$$\tan\theta_8 = \frac{F_{\eta'}^8}{F_\eta^8}, \qquad \tan\theta_0 = -\frac{F_\eta^0}{F_{\eta'}^0}. \tag{5.33}$$

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

$$F^s = \sqrt{(F_\eta^s)^2 + (F_{\eta'}^s)^2}, \qquad F^\ell = \sqrt{(F_\eta^\ell)^2 + (F_{\eta'}^\ell)^2}, \tag{5.34}$$

$$\tan\phi_s = -\frac{F_{\eta'}^s}{F_\eta^s}, \qquad \tan\phi_\ell = \frac{F_\eta^\ell}{F_{\eta'}^\ell}. \tag{5.35}$$

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 [41].

5.3.1 LO large-N_c ChPT

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

JHEP08(2021)137

meson masses via the LO GMOR relations

$$(\mu_8^{\rm LO})^2 = \frac{2}{3}B_0\,(m_\ell + 2m_s) = \overline{M}^2 + \frac{1}{3}\delta M^2, \tag{5.36}$$

$$(\mu_0^{\rm LO})^2 = \frac{2}{3}B_0\,(2m_\ell + m_s) + M_0^2 = \overline{M}^2 + M_0^2, \tag{5.37}$$

$$(\mu_{80}^{\rm LO})^2 = -\frac{2\sqrt{2}}{3}B_0\,(m_s - m_\ell) = -\frac{\sqrt{2}}{3}\delta M^2, \tag{5.38}$$

where the anomalous contribution $M_0^2 = 6\tau_0/F^2$ is proportional to the quenched topological susceptibility τ_0 [1, 2] 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. (5.25):

$$F_\eta^8 = F_{\eta'}^0 = F\cos\theta, \qquad -F_\eta^0 = F_{\eta'}^8 = F\sin\ \theta, \tag{5.39}$$

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 [41, 42] and also the results of the present study clearly show $F_\eta^8 \neq F_{\eta'}^0$ and $F_\eta^0 \neq -F_{\eta'}^0$.

5.3.2 NLO large-N_c ChPT

The large-N_c ChPT expansion for the masses and decay constants has been worked out to NNLO in [81]. Here, we use the results of [42] 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

$$(\mu_8^{\rm NLO})^2 = (\mu_8^{\rm LO})^2 + \frac{8}{3F^2}\,(2L_8 - L_5)\,\delta M^4, \tag{5.40}$$

$$(\mu_0^{\rm NLO})^2 = (\mu_0^{\rm LO})^2 + \frac{4}{3F^2}\,(2L_8 - L_5)\,\delta M^4 - \frac{8}{F^2}L_5\overline{M}^2M_0^2 - \tilde{\Lambda}\overline{M}^2 - \Lambda_1 M_0^2, \tag{5.41}$$

$$(\mu_{80}^{\rm NLO})^2 = (\mu_{80}^{\rm LO})^2 - \frac{4\sqrt{2}}{3F^2}\,(2L_8 - L_5)\,\delta M^4 + \frac{4\sqrt{2}}{3F^2}L_5M_0^2\delta M^2 + \frac{\sqrt{2}}{6}\tilde{\Lambda}\delta M^2, \tag{5.42}$$

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 [42]. 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.

JHEP08(2021)137

The decay constants are given by

$$F_\eta^8 = F\left[\cos\theta + \frac{4L_5}{3F^2}\left(3\cos\theta\overline{M}^2 + (\sqrt{2}\sin\theta + \cos\theta)\delta M^2\right)\right], \tag{5.43}$$

$$F_{\eta'}^8 = F\left[\sin\theta + \frac{4L_5}{3F^2}\left(3\sin\theta\overline{M}^2 + (\sin\theta - \sqrt{2}\cos\theta)\delta M^2\right)\right], \tag{5.44}$$

$$F_\eta^0 = -F\left[\sin\theta\left(1 + \frac{\Lambda_1}{2}\right) + \frac{4L_5}{3F^2}\left(3\sin\theta\overline{M}^2 + \sqrt{2}\cos\theta\delta M^2\right)\right], \tag{5.45}$$

$$F_{\eta'}^0 = F\left[\cos\theta\left(1 + \frac{\Lambda_1}{2}\right) + \frac{4L_5}{3F^2}\left(3\cos\theta\overline{M}^2 - \sqrt{2}\sin\theta\delta M^2\right)\right], \tag{5.46}$$

where θ is the mass mixing angle defined in eq. (5.25), evaluated with the entries of the NLO mass matrix, eqs. (5.40)–(5.42). Note that in the standard $\overline{\text{MS}}$ 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.

5.3.3 Impact of the mass dependence of t_0 on the NLO parametrization

In order to eliminate some of the lattice spacing effects, in section 5.5 we carry out our fits after forming dimensionless combinations $\sqrt{8t_0}M_\mathcal{M}$ and $\sqrt{8t_0}F_\mathcal{M}^a$. However, t_0 depends on the pseudoscalar masses too. In the continuum limit, to leading order, this can be parameterized as [51]

$$t_0(\overline{M}, \delta M) = t_0^\chi\left(1 + k\,8t_0\overline{M}^2\right), \tag{5.47}$$

where $k = -0.0466(62)$ [47]. This dependence somewhat alters the functional form of the ChPT expectations for the rescaled quantities. Within the NLO parametrization of the squared mass matrix eqs. (5.40)–(5.42), only the element μ_0^2 is affected. In this case one has to add a term $-kM_0^2\overline{M}^2 8t_0$ to the parametrization, which we do. This is due to the fact that M_0^2, which is defined in the chiral limit, appears at leading order and $8t_0^\chi M_0^2 = 8t_0 M_0^2(1 - k\,8t_0\overline{M}^2)$. Regarding the decay constants, rewriting $\sqrt{8t_0^\chi}F = \sqrt{8t_0}F(1 - \frac{k}{2}8t_0\overline{M}^2)$ means that the term $-(k/2)F\cos\theta\overline{M}^2 8t_0$ needs to be added to eqs. (5.43) and (5.46) while the term $-(k/2)F\sin\theta\overline{M}^2 8t_0$ has to be added to eq. (5.44) and subtracted from eq. (5.45).

The impact of $k \neq 0$ on the physical point masses and decay constants turns out to be marginal. 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 $1.4\cdot 10^{-4}$.

Note that the parametrizations of the anomalous matrix elements eqs. (B.9) and (B.11) contain leading order terms $\propto FM_0^2$, $\propto F\overline{M}^2$ and $\propto F\delta M^2$. This amounts to adding terms $\propto -(3k/2)FM_0^2\overline{M}^2 8t_0$, $\propto -(k/2)F\overline{M}^4 8t_0$ and $\propto -(k/2)F\delta M^2\overline{M}^2 8t_0$ in our analysis of section 6.2.

5.4 Parameterizing lattice spacing effects

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 $f_O^{\text{cont}}(\overline{M}^2, \delta M^2|\ldots)$ where the ellipses represent the fit parameters (i.e. the LECs) and O can be either of the two masses or four decay constants.

JHEP08(2021)137

We remind the reader that for the decay constants not all the $\mathcal{O}(a)$ improvement coefficients are known. Therefore, we start from the following ansatz

$$f_O(a,\overline{M}^2,\delta M^2) = f_O^{\rm cont}(\overline{M}^2,\delta M^2|\ldots)h_O^{(1)}(a,am_\ell,am_s|\ldots)h_O^{(2)}(a^2,a^2\overline{M}^2,a^2\delta M^2|\ldots), \tag{5.48}$$

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. (4.21) and (4.23), respectively) 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 these functions contain the known parameters b_A, $\tilde{b}_A$ and the free parameter f_A. In the case of the singlet decay constants, within $h_O^{(1)}$ the unknown parameters d_A, $\tilde{d}_A$ and δc_A appear, the latter multiplying the term $E_n\check{H}_n^0$ (see eqs. (4.22) and (4.24)). 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 section 3.1 and table 2). 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

$$12t_0^{\rm ph}\overline{M}^{\rm ph\,2} = 1.110 \quad \text{and} \quad 8t_0^{\rm ph}\delta M^{\rm ph\,2} = 1.902, \tag{5.49}$$

see section 3.1. For the linear lattice effects on the octet and singlet decay constants, we combine the results of eqs. (5.1), (5.5) and (4.22)–(4.24), to obtain the functions

$$h^{(1)}_{F_n^8}(a|f_A^l) = 1 - 3a\tilde{b}_A\overline{m} - ab_A\frac{Z_A}{\sqrt{3}}\frac{m_\ell\check{F}_n^\ell - \sqrt{2}m_s\check{F}_n^s}{f^{\rm cont}_{F_n^8}(\overline{M}^2,\delta M^2)} + \sqrt{2}af_AZ_A\frac{\delta m\check{F}_n^0}{f^{\rm cont}_{F_n^8}(\overline{M}^2,\delta M^2)}, \tag{5.50}$$

$$h^{(1)}_{F_n^0}(a|d_A^l,\tilde{d}_A^l,\delta c_A^l) = 1 - 3a\tilde{d}_A\overline{m} - ad_A\frac{Z_A^s}{\sqrt{3}}\frac{\sqrt{2}m_\ell\check{F}_n^\ell + m_s\check{F}_n^s}{f^{\rm cont}_{F_n^0}(\overline{M}^2,\delta M^2)} - a\delta c_AZ_A^s\frac{\check{H}_n^0}{f^{\rm cont}_{F_n^0}(\overline{M}^2,\delta M^2)}, \tag{5.51}$$

where $n = \eta, \eta'$. We have substituted the data on the decay constants $\tilde{F}_n^8$ and $\check{F}_n^0$ by the fitted continuum limit parametrizations $f^{\rm cont}_{F_n^a}(\overline{M}^2,\delta M^2)$, 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 (5.48) $h_O^{(1)}$ is multiplied by $f^{\rm cont}_{F_n^a}(\overline{M}^2,\delta M^2)$. 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 parameterize these as follows

$$f_A(g^2) = f_A^lg^6, \quad d_A(g^2) = b_A(g^2) + d_A^lg^4, \quad \tilde{d}_A(g^2) = \tilde{d}_A^lg^4, \quad \delta c_A(g^2) = \delta c_A^lg^4, \tag{5.52}$$

JHEP08(2021)137

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. (5.50) and (5.51). The above powers of g^2 correspond to the first non-trivial orders of the perturbative expansions.

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

$$h_O^{(2)}(a^2, t_0\overline{M}^2, t_0\delta M^2 | l_O, m_O, n_O) = 1 + a^2\left(l_O + m_O\overline{M}^2 + n_O\delta M^2\right). \tag{5.53}$$

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 [49].

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 cutoff effects and $6 \times 3 = 18$ more coefficients to parameterize 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.

5.5 Continuum and chiral extrapolation: fits and error estimates

We describe how we determine which fit parameters are most relevant and how we estimate the systematic uncertainty associated with the chosen set of fit forms. 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, the latter by employing Orear's method [82]. 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.

The systematics associated with the continuum and quark mass extrapolations need to be quantified. Since lattice spatial extents of $L_s M_\pi \gtrsim 4$ are realized, finite volume effects can safely be neglected. 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 $\chi^2/N_{\rm df} \approx 220/126 \approx 1.75$. Additional discretization terms are subsequently included and those fits for which the coefficients can be resolved with reasonable precision are given in table 7. The reference fit has the id "1" in the table. The LECs extracted from these fits are collected in table 8 and the results for the masses and decay constants at the physical

JHEP08(2021)137

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 7. 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 section 5.4) 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. (5.53). The values of these coefficients can be found in appendix D. In all the cases, the NLO large-N_c expressions are used for the continuum part of the fit function.

point are detailed in table 9. The coefficients of the discretization terms are provided in appendix D.

Among the linear improvement terms (see eqs. (5.50) and (5.51)) 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 table 7 have the lowest and very similar χ^2 values, see table 9. 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. (5.53)). In what follows we take fit 7 with $\chi^2/N_{\rm df} \approx 179/122 \approx 1.47$ as our main fit. This was selected from the fits with $1.46 \leq \chi^2/N_{\rm df} \leq 1.49$ since the resulting parameter values are in the centre of the scatter between the different fit forms, see tables 10–12. 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 $\chi^2/N_{\rm df} \approx 155/122 \approx 1.27$. 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.

JHEP08(2021)137

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 8. Results for the LECs obtained when employing the fit forms detailed in table 7. The dimensionful LECs are given in units of the gradient flow scale in the chiral limit.

The $\chi^2/N_{\rm df}$ 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 cutoffs $12t_0\overline{M}^2 < c$, where $c = 1.6, 1.4, 1.2$. These values correspond to $\overline{M} \approx 493\,\text{MeV}$, $462\,\text{MeV}$ and $427\,\text{MeV}$, 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 table 1 and figure 1. 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 on this trajectory. We perform a fit for each cut using fit form 7. The results are listed in tables 13 and 14. 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.

JHEP08(2021)137

id	$\chi^2/N_{\rm df}$	$\sqrt{8t_0^{\rm ph}}M_\eta$	$\sqrt{8t_0^{\rm ph}}M_{\eta'}$	$\sqrt{8t_0^{\rm ph}}F_\eta^8$	$\sqrt{8t_0^{\rm ph}}F_{\eta'}^8$	$\sqrt{8t_0^{\rm ph}}F_\eta^0$	$\sqrt{8t_0^{\rm ph}}F_{\eta'}^0$
1	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 9. 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 table 7.

In principle, large-N_c ChPT expressions for the masses and decay constants to NNLO are available [42, 81], 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 appendix A for details on the parametrization and the resulting LECs. However, fits to this functional form did not improve the description of the 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.

5.6 Fit results for the masses and decay constants

LO. For completeness, we perform a fit to the η and η' masses employing the LO large-N_c ChPT expressions (see section 5.3.1). 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

JHEP08(2021)137

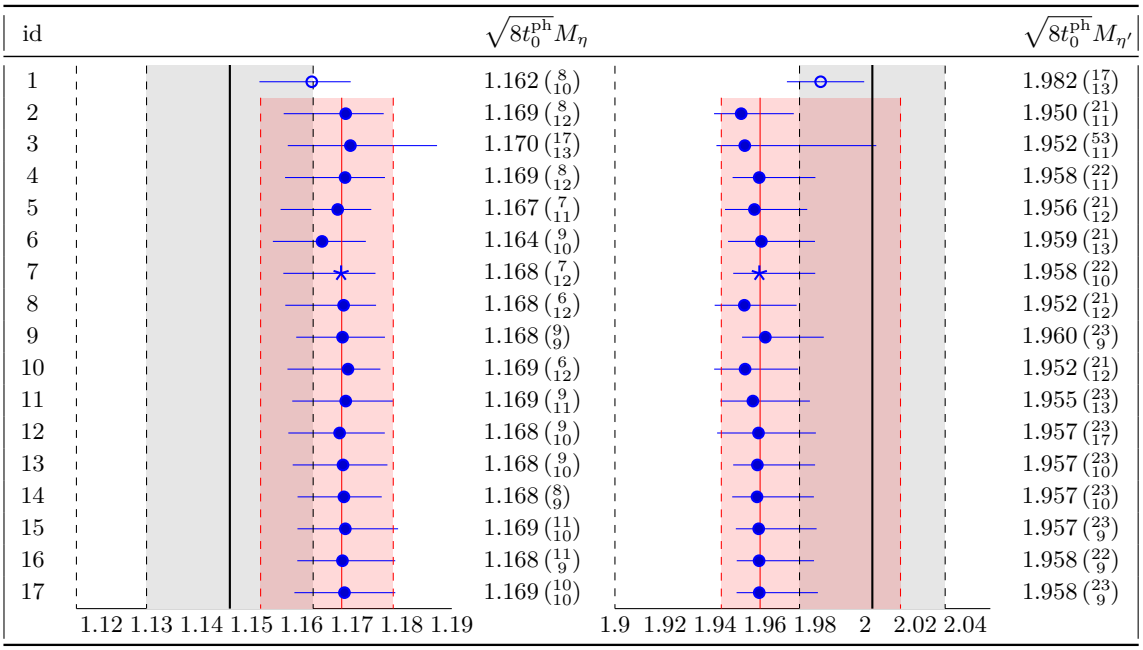


Table 10. Results for the masses of the η and η' at the physical point in units of the gradient flow scale. The black line marks the experimental result converted using $(8t_0^{\rm ph})^{-1/2} = 475(6)\,\text{MeV}$ [49] and the grey shaded region marks the uncertainty due to the error on $t_0^{\rm ph}$. The red line indicates the central values predicted by fit 7 (see table 7). The red shaded region represents the total uncertainty of our final results where all errors are added in quadrature (see section 5.6). The first fit does not sufficiently parameterize the lattice spacing effects and is not included in the determination of the associated systematic error.

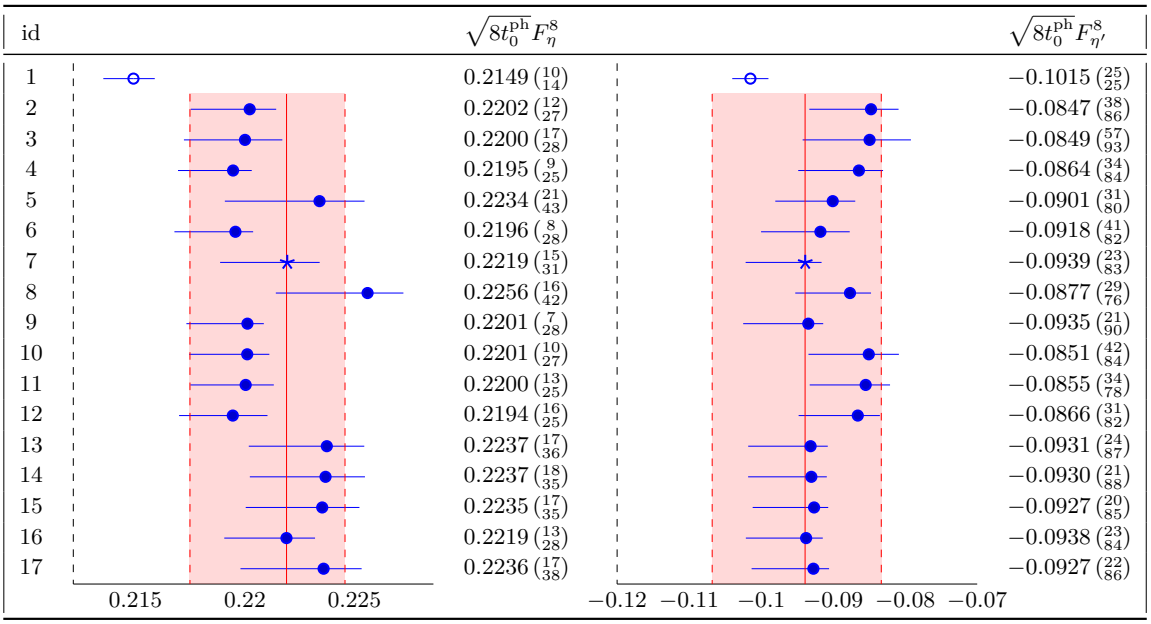


Table 11. Octet decay constants of the η and η' at the physical point in units of the gradient flow scale, displayed as in table 10.

JHEP08(2021)137

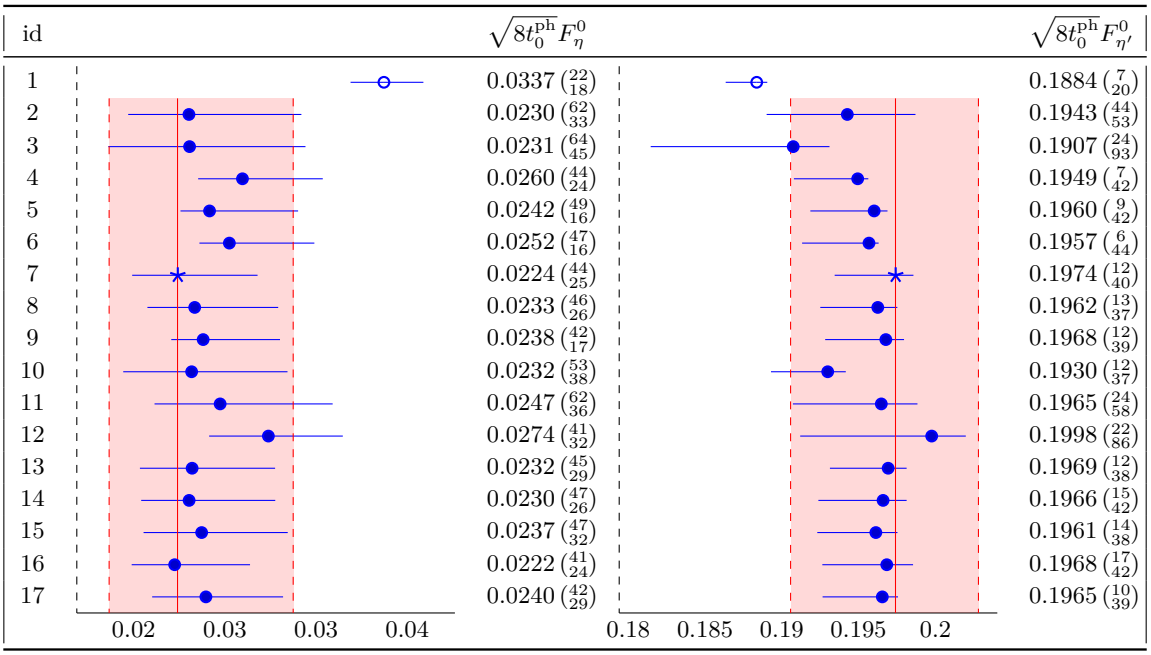


Table 12. Singlet decay constants of the η and η' at the physical point in units of the gradient flow scale, displayed as in table 10.

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 13. 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 table 7 and imposing cut-offs $12t_0\overline{M}^2 < c$, as well as including all the data (first row).

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 14. LECs obtained for the fits detailed in table 13. The dimensionful LECs are given in units of the gradient flow scale.

JHEP08(2021)137

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

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 $n_{M_{\eta'}}$. This fit is displayed in figure 17 in appendix C. 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 $M_0^2 = 2.787(8t_0^\chi)^{-1} = (785\,\text{MeV})^2$. 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 section 5.5. 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 figure 5. We obtain for the masses at the physical point, in the continuum limit

$$\sqrt{8t_0^{\rm ph}}M_\eta = 1.168\,(^{8}_{14})_{\rm stat}\,(^{1}_{0})_a\,(^{5}_{6})_\chi \quad \text{and} \quad \sqrt{8t_0^{\rm ph}}M_{\eta'} = 1.958\,(^{27}_{13})_{\rm stat}\,(^{0}_{6})_a\,(^{48}_{3})_\chi\,, \tag{5.54}$$

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 section 5.5 for details. The results are converted to physical units in section 7. 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

$$\begin{aligned}
\sqrt{8t_0^{\rm ph}}F_\eta^8 &= 0.2219\,(^{18}_{37})_{\rm stat}\,(^{17}_{24})_a\,(^{10}_{2})_\chi\,,\\
\sqrt{8t_0^{\rm ph}}F_{\eta'}^8 &= -0.0939\,(^{28}_{100})_{\rm stat}\,(^{84}_{0})_a\,(^{58}_{82})_\chi\,,\\
\sqrt{8t_0^{\rm ph}}F_\eta^0(\mu=\infty) &= 0.0224\,(^{53}_{30})_{\rm stat}\,(^{28}_{0})_a\,(^{5}_{21})_\chi\,(^{20}_{8})_{\rm renorm}\,,\\
\sqrt{8t_0^{\rm ph}}F_{\eta'}^0(\mu=\infty) &= 0.1974\,(^{14}_{48})_{\rm stat}\,(^{0}_{31})_a\,(^{4}_{27})_\chi\,(^{52}_{26})_{\rm renorm}\,.
\end{aligned} \tag{5.55}$$

The singlet decay constants depend on the QCD scale. As detailed in section 5.2, prior to the fits we run our results from a scale $\mu_0 = a^{-1}$ up to $\mu = \infty$. To quantify the systematic

JHEP08(2021)137

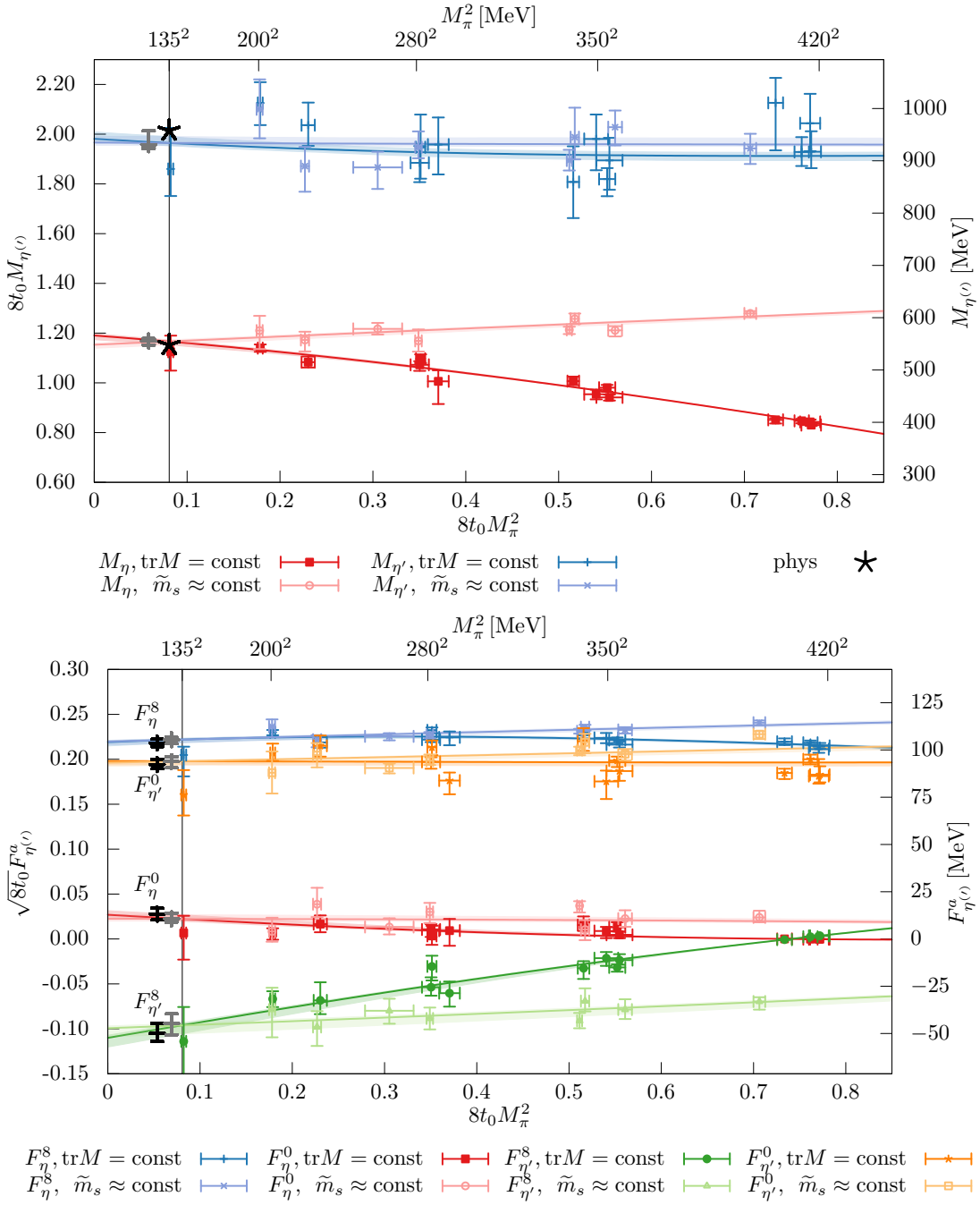


Figure 5. Simultaneous fit to the masses (top) and four decay constants (bottom) of the η and η'. The fit form incorporates the NLO large-N_c ChPT expressions and the discretization terms corresponding to fit 7 in table 7. The points have been shifted to compensate for lattice spacing effects and lie 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 section 5.7). All errors are added in quadrature.

JHEP08(2021)137

error from the matching to the $\overline{\text{MS}}$ scheme, we repeat the fits, setting $\mu_0 = a^{-1}/2$ and $\mu_0 = 2a^{-1}$, and add the range of results (see table 15) as an additional systematic error.[7] As one may expect, this error is dominated by the fit where we set $\mu = a^{-1}/2$, see table 15. The results for the decay constants can also be converted to the strange/light flavour basis (eq. (2.11)) and/or given in terms of two angles and two dimensionful decay constants, see eqs. (2.12) and (2.13). 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 table 24 in appendix E. We discuss the results and their scale dependence in detail in section 7.2.

5.7 Results for the large-N_c low energy constants

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

$$\begin{aligned}
L_5 &= 1.58\,\binom{17}{7}_{\text{stat}}\binom{0}{22}_{a}\binom{20}{9}_{\chi}\cdot 10^{-3},\\
L_8 &= 0.96\,\binom{15}{6}_{\text{stat}}\binom{0}{14}_{a}\binom{16}{7}_{\chi}\cdot 10^{-3},\\
M_0(\mu=\infty) &= 1.67\,\binom{2}{6}_{\text{stat}}\binom{1}{2}_{a}\binom{0}{1}_{\chi}\binom{3}{2}_{\text{renorm}}(8t_0^{\chi})^{-1/2},\\
F &= 0.1890\,\binom{27}{37}_{\text{stat}}\binom{36}{0}_{a}\binom{21}{38}_{\chi}(8t_0^{\chi})^{-1/2},\\
\Lambda_1(\mu=\infty) &= -0.22\,\binom{1}{5}_{\text{stat}}\binom{0}{3}_{a}\binom{0}{3}_{\chi}\binom{6}{3}_{\text{renorm}},\\
\tilde{\Lambda} &= -0.20\,\binom{5}{16}_{\text{stat}}\binom{19}{0}_{a}\binom{12}{29}_{\chi}\binom{3}{5}_{\text{renorm}}.
\end{aligned} \tag{5.56}$$

The combination $\tilde{\Lambda} = \Lambda_1 - 2\Lambda_2$ is scale invariant to this order in ChPT [42, 83], however, since its central value varies when changing μ_0, see table 15, we also assign a renormalization error in this case. The above results give

$$\Lambda_2(\mu=\infty) = -0.1\,\binom{8}{4}_{\text{stat}}\binom{0}{10}_{a}\binom{14}{8}_{\chi}\binom{5}{3}_{\text{renorm}}. \tag{5.57}$$

The fits on which these results are based give η and η' masses that are compatible, within errors, with experiment, see above and section 7.1. 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:

$$\begin{aligned}
\chi^2_{\text{priors}} = \chi^2 &+ \left[\frac{\sqrt{8t_0^{\text{ph}}}M_{\eta}^{\text{ph}} - f_{M_\eta}(a=0, 12t_0\overline{M}^2, 8t_0\delta M^2)}{\sigma\left(\sqrt{8t_0^{\text{ph}}}M_{\eta}^{\text{ph}}\right)}\right]^2\\
&+ \left[\frac{\sqrt{8t_0^{\text{ph}}}M_{\eta'}^{\text{ph}} - f_{M_{\eta'}}(a=0, 12t_0\overline{M}^2, 8t_0\delta M^2)}{\sigma\left(\sqrt{8t_0^{\text{ph}}}M_{\eta'}^{\text{ph}}\right)}\right]^2,
\end{aligned} \tag{5.58}$$

where we use the physical values from the Particle Data Group (PDG) [63] for M_η and $M_{\eta'}$, see eq. (7.3). These are converted to dimensionless numbers, using $(8t_0^{\text{ph}})^{-1/2} =$

[7] 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 $\tilde{\Lambda}$.

JHEP08(2021)137

475(6) MeV [49]. 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 $(8t_0^{\rm ph})^{1/2}$, rather than independently sampling the two dimensionless combinations $(8t_0^{\rm ph})^{1/2}M_\eta^{\rm ph}$ and $(8t_0^{\rm ph})^{1/2}M_{\eta'}^{\rm ph}$. If a more precise determination of $t_0^{\rm ph}$ 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. (5.56) and (5.57):

$$\begin{aligned}
L_5 &= 1.66\,\tbinom{12}{9}_{\rm stat}\,\tbinom{0}{26}_{a}\,\tbinom{13}{8}_{\chi}\cdot 10^{-3},\\
L_8 &= 1.08\,\tbinom{11}{6}_{\rm stat}\,\tbinom{0}{12}_{a}\,\tbinom{3}{10}_{\chi}\cdot 10^{-3},\\
M_0(\mu=\infty) &= 1.62\,\tbinom{2}{4}_{\rm stat}\,\tbinom{3}{1}_{a}\,\tbinom{3}{0}_{\chi}\,\tbinom{2}{1}_{\rm renorm}\,(8t_0^{\chi})^{-1/2},\\
F &= 0.1866\,\tbinom{26}{29}_{\rm stat}\,\tbinom{54}{0}_{a}\,\tbinom{19}{16}_{\chi}\,(8t_0^{\chi})^{-1/2},\\
\Lambda_1(\mu=\infty) &= -0.25\,\tbinom{1}{4}_{\rm stat}\,\tbinom{3}{1}_{a}\,\tbinom{1}{1}_{\chi}\,\tbinom{5}{2}_{\rm renorm},\\
\tilde{\Lambda} &= -0.46\,\tbinom{8}{10}_{\rm stat}\,\tbinom{21}{0}_{a}\,\tbinom{9}{10}_{\chi}\,\tbinom{1}{2}_{\rm renorm},\\
\Lambda_2(\mu=\infty) &= 0.11\,\tbinom{5}{5}_{\rm stat}\,\tbinom{0}{9}_{a}\,\tbinom{6}{5}_{\chi}\,\tbinom{3}{2}_{\rm renorm}.
\end{aligned} \tag{5.59}$$

In section 7.3 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 [42] and in section 7.3. 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 $\chi^2/N_{\rm df} \approx 312/122 \approx 2.56$. The functional form and the resulting LECs are detailed in appendix A.

6 Gluonic matrix elements and axial Ward identities

The AWIs are discussed and the octet AWI is tested against our data. We then proceed to construct the gluonic matrix elements of the η and η', using fermionic currents via the singlet AWI. After addressing the renormalization of pseudoscalar gluonic matrix elements, we compare the results obtained via the singlet AWI with a direct determination. The quark mass dependence of the topological susceptibility is also determined.

6.1 The axial Ward identities

The AWIs between renormalized operators (indicated by a hat) read

$$\partial_\mu \widehat{A}^a_\mu = \Big(\overline{\psi}\gamma_5\widehat{\{M,t^a\}}\psi\Big) + \sqrt{2N_f}\delta^{a0}\widehat{\omega}, \tag{6.1}$$

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

$$\omega(x) = -\frac{1}{16\pi^2}\,\mathrm{tr}\left[F_{\mu\nu}(x)\widetilde{F}_{\mu\nu}(x)\right] = -\frac{1}{32\pi^2}F^a_{\mu\nu}(x)\widetilde{F}^a_{\mu\nu}(x) = -\frac{1}{64\pi^2}\epsilon_{\mu\nu\rho\sigma}F^a_{\mu\nu}(x)F^a_{\rho\sigma}(x). \tag{6.2}$$

JHEP08(2021)137

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

$$\partial_\mu \widehat{A}^8_\mu = \frac{2}{3}\,(\widehat{m}_\ell + 2\widehat{m}_s)\,\widehat{P}^8 - \frac{2\sqrt{2}}{3}\delta\widehat{m}\widehat{P}^0, \tag{6.3}$$

and

$$\partial_\mu \widehat{A}^0_\mu = \frac{2}{3}\,(2\widehat{m}_\ell + \widehat{m}_s)\,\widehat{P}^0 - \frac{2\sqrt{2}}{3}\delta\widehat{m}\widehat{P}^8 + \sqrt{6}\,\widehat{\omega}, \tag{6.4}$$

respectively. In the octet/singlet basis only the singlet AWI receives a contribution from the anomaly. The corresponding AWIs in the flavour basis read

$$\partial_\mu \widehat{A}^s_\mu = 2\widehat{m}_s\widehat{P}^s + 2\,\widehat{\omega}, \quad \partial_\mu \widehat{A}^\ell_\mu = 2\widehat{m}_\ell\widehat{P}^\ell + 2\sqrt{2}\,\widehat{\omega}. \tag{6.5}$$

These are somewhat simpler because the quark flavours decouple, up to the anomaly contribution which now enters both AWIs.

We determine our quark masses, using the AWIs for $a = 1$ and $a = 4$ in the lattice scheme:[8]

$$\partial_\mu \langle \Omega | \bar{d}\gamma_\mu\gamma_5 u | \pi^+ \rangle = 2\widetilde{m}_\ell \langle \Omega | \bar{d}\gamma_5 u | \pi^+ \rangle, \tag{6.6}$$

$$\partial_\mu \langle \Omega | \bar{s}\gamma_\mu\gamma_5 u | K^+ \rangle = (\widetilde{m}_\ell + \widetilde{m}_s) \langle \Omega | \bar{s}\gamma_5 u | K^+ \rangle. \tag{6.7}$$

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. The lattice AWI quark masses are related to the continuum masses via

$$\widehat{m}_q(\mu) = \frac{Z_A}{Z_P(\mu)}\widetilde{m}_q. \tag{6.8}$$

Again Z_A/Z_P is known non-perturbatively in the RGI scheme [84] and can be related to the $\overline{\text{MS}}$ scheme at a scale μ perturbatively at the five-loop level [77] whenever this is needed.

The octet AWI between lattice matrix elements reads

$$\partial_\mu \left\langle \Omega \left| A^8_\mu \right| \mathcal{M} \right\rangle = \frac{2}{3}\,(\widetilde{m}_\ell + 2\widetilde{m}_s) \left\langle \Omega \left| P^8 \right| \mathcal{M} \right\rangle - \frac{2\sqrt{2}}{3}\delta\widetilde{m}\, r_P \left\langle \Omega \left| P^0 \right| \mathcal{M} \right\rangle, \tag{6.9}$$

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 is equivalent to g_P, see eq. (3.3)), 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

[8] The same results can be obtained in the first case for $a = 2$ and the π^- or for $a = 3$ and the π^0 (where the disconnected quark contractions cancel due to isospin symmetry), while the combination $\widetilde{m}_\ell + \widetilde{m}_s$ can also be extracted using $a = 5, 6, 7$ with the appropriate kaon states.

JHEP08(2021)137

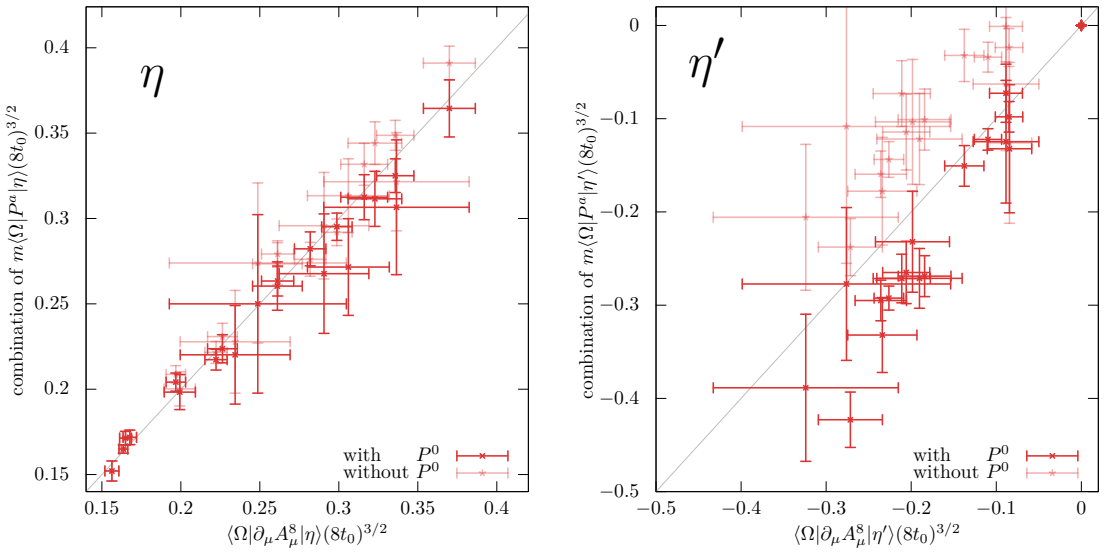


Figure 6. Check of the octet AWI (6.9) for the η and η' states. Light points correspond to the case when the singlet contribution to the octet AWI is neglected.

is trivial if applied to the $\eta' = \eta_0$ state. Note that the left hand side of eq. (6.9) can be replaced with

$$\partial_\mu \left\langle \Omega \left| A^8_\mu \right| \mathcal{M} \right\rangle = Z_A^{-1} \partial_\mu \left\langle \Omega \left| \widehat{A}^8_\mu \right| \mathcal{M} \right\rangle = Z_A^{-1} M^2_{\mathcal{M}} F^8_{\mathcal{M}}. \tag{6.10}$$

Hence, the combination on the right hand side of eq. (6.9) does not depend on the momentum of the meson $\mathcal{M}$.

In figure 6 we check the octet AWI (6.9) at zero momentum directly against our data, utilizing the quark masses computed according to eqs. (6.6) and (6.7). 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 section 5.5 (fit 7). Throughout, we find reasonable agreement between the left and right hand sides of the Ward identity, as shown in figure 6. 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.

In the singlet case, due to the anomaly contribution to eq. (6.4), we would expect that

$$\partial_\mu \left\langle \Omega \left| A^0_\mu \right| \mathcal{M} \right\rangle \neq 2\overline{\widetilde{m}}\, r_P \left\langle \Omega \left| P^0 \right| \mathcal{M} \right\rangle - \frac{2\sqrt{2}}{3} \delta\widetilde{m} \left\langle \Omega \left| P^8 \right| \mathcal{M} \right\rangle, \tag{6.11}$$

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 figure 7. 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

JHEP08(2021)137

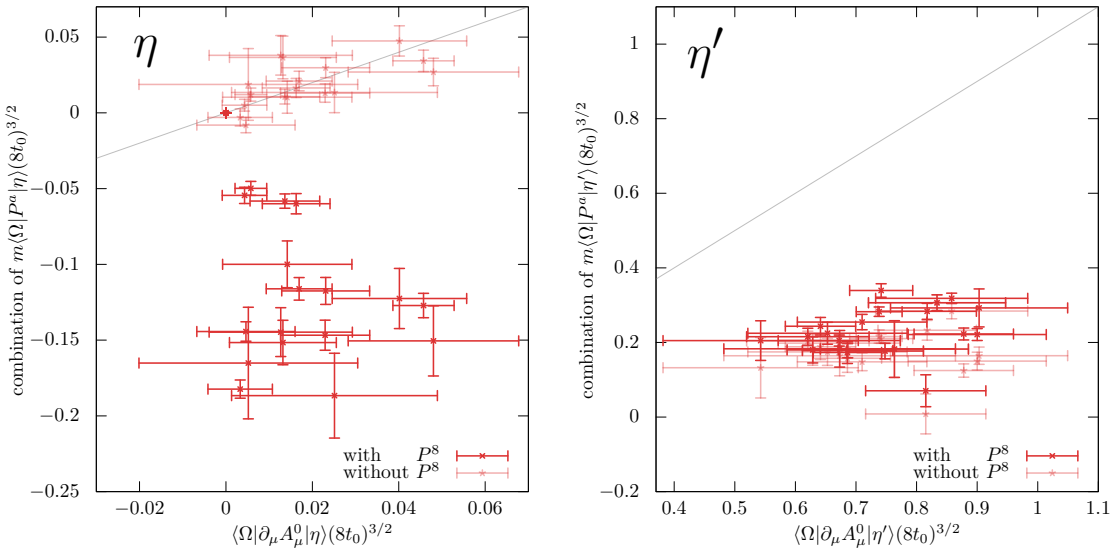


Figure 7. Fermionic contributions to the singlet AWI for the η and η' states. Light points correspond to only the singlet contribution whereas the darker points correspond to all pseudoscalar terms in eq. (6.4), but the anomalous contribution is neglected.

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 subsection.

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

$$a_{\mathcal{M}} = 2\langle\Omega|\widehat{\omega}|\mathcal{M}\rangle \tag{6.12}$$

directly using gluonic correlators, we first compute them via the renormalized singlet AWI (6.4):

$$\begin{aligned} a_{\mathcal{M}}(\mu) &= \sqrt{\frac{2}{3}} Z_A^s(\mu)\partial_\mu \left\langle \Omega \left| A_\mu^0 \right| \mathcal{M} \right\rangle + \frac{2\sqrt{2}}{\sqrt{3}} Z_A \left[\frac{\sqrt{2}}{3}\delta\widetilde{m} \left\langle \Omega \left| P^8 \right| \mathcal{M} \right\rangle - r_P \overline{\widetilde{m}} \left\langle \Omega \left| P^0 \right| \mathcal{M} \right\rangle \right] \\ &= \sqrt{\frac{2}{3}} M_{\mathcal{M}}^2 F_{\mathcal{M}}^0(\mu) + \frac{4}{3\sqrt{3}} \frac{Z_A}{Z_P} \delta\widetilde{m} H_{\mathcal{M}}^8 - \frac{2\sqrt{2}}{\sqrt{3}} r_P \frac{Z_A}{Z_P} \overline{\widetilde{m}} H_{\mathcal{M}}^0, \end{aligned} \tag{6.13}$$

where $H^a_{\mathcal{M}}$ are the renormalized and $\mathcal{O}(a)$ improved pseudoscalar matrix elements, in analogy to eqs. (5.1) and (5.5). This fermionic definition has the advantage that no knowledge of the renormalization factors $Z_{\omega A}$ and Z_ω is needed (see section 6.3 below). Note that $a_{\mathcal{M}}$ depends on the renormalization scale μ and only in the modified scheme, discussed

JHEP08(2021)137

in section 5.2 (which corresponds to $\mu \to \infty$), do the gluonic matrix elements become scale independent.

The improvement coefficients that enter the computation of the singlet decay constants are taken from section 5.5 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_P^s. The latter was defined in eq. (3.3) and persists in the chiral limit. We remark that by replacing $a g_P \mathrm{tr} F_{\mu\nu}\widetilde{F}_{\mu\nu} \mapsto a c_P^s \partial_\mu A_\mu^0$, the definition of the coefficients d_P and $\bar{d}_P$ (and therefore of $\tilde{d}_P$) with respect to [46] is somewhat altered. Some of the above coefficients will be free parameters within a combined fit, incorporating the NLO large-N_c ChPT continuum prediction eqs. (B.9) and (B.11), that we derive in appendix B.

We start by defining partially improved matrix elements at $\mu = \infty$:

$$\check{a}_{\eta^{(\prime)}} = \sqrt{\frac{2}{3}} M^2_{\eta^{(\prime)}} F^0_{\eta^{(\prime)}}(\mu) + \frac{4}{3\sqrt{3}} Z_A \delta\widetilde{m} \check{H}^8_{\eta^{(\prime)}} - \frac{2\sqrt{2}}{\sqrt{3}} r_P Z_A \overline{\widetilde{m}} \check{H}^0_{\eta^{(\prime)}}, \tag{6.14}$$

where $\check{H}^a_{\eta^{(\prime)}} = \langle \Omega | P^a | \eta^{(\prime)} \rangle$ are unimproved pseudoscalar lattice matrix elements and we assume $r_P = 1$. We then carry out a fit according to

$$\begin{aligned} \check{a}_{\eta^{(\prime)}}(a, \overline{M}^2, \delta M^2) = a_{\eta^{(\prime)}}(\overline{M}^2, \delta M^2 | \cdots) \\ - \frac{2\sqrt{2} Z_A}{\sqrt{3}} \overline{\widetilde{m}} \left[3a\tilde{d}_P \overline{m} \check{H}^0_{\eta^{(\prime)}} + a d_P \frac{1}{\sqrt{3}} \left(\sqrt{2} m_\ell \check{H}^\ell_{\eta^{(\prime)}} + m_s \check{H}^s_{\eta^{(\prime)}} \right) + a c_P^s M^2_{\eta^{(\prime)}} F^0_{\eta^{(\prime)}} \right] \\ + \frac{4 Z_A}{3\sqrt{3}} \delta\widetilde{m} \left[3a\tilde{b}_P \overline{m} \check{H}^8_{\eta^{(\prime)}} + a b_P \frac{1}{\sqrt{3}} \left(m_\ell \check{H}^\ell_{\eta^{(\prime)}} - \sqrt{2} m_s \check{H}^s_{\eta^{(\prime)}} \right) + \sqrt{2} a f_P \delta m \check{H}^0_{\eta^{(\prime)}} \right], \end{aligned} \tag{6.15}$$

where the continuum parametrizations $a_{\eta^{(\prime)}}(\overline{M}^2, \delta M^2 | \cdots)$ correspond to eqs. (B.9)–(B.11) and the ellipses represent the six NLO LECs. 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 parameterize the coefficients d_P, $\tilde{d}_P$, f_P and c_P^s (that are functions of g^2) similarly to eq. (5.52) with one parameter each, while b_P and $\tilde{b}_P$ are known non-perturbatively [70]:

$$\begin{aligned} b_P(\beta = 3.4) &= 1.622(74), & b_P(\beta = 3.46) &= 1.592(213), \\ b_P(\beta = 3.55) &= 1.560(165), & b_P(\beta = 3.7) &= 1.696(78), \end{aligned} \tag{6.16}$$

$$\begin{aligned} \tilde{b}_P(\beta = 3.4) &= 0.39(27), & \tilde{b}_P(\beta = 3.46) &= 0.32(20), \\ \tilde{b}_P(\beta = 3.55) &= 0.40(23), & \tilde{b}_P(\beta = 3.7) &= 0.16(13). \end{aligned} \tag{6.17}$$

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 section 5.7. Therefore, in analogy to eq. (5.58), we add these results, given in eq. (5.59), as priors to the χ^2 function. The widths σ are set to the statistical and systematic errors, added in quadrature. 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 figure 8, gives a valid description of the data, with a fully correlated $\chi^2/N_{\mathrm{df}} \approx 34/31 \approx 1.09$. We obtain

$$d_P(g^2) = b_P(g^2) + 6.6(6) g^4, \quad c_P^s(g^2) = -2.4(3) g^4 \quad \text{and} \quad f_P(g^2) = -26(6) g^6 \tag{6.18}$$

JHEP08(2021)137

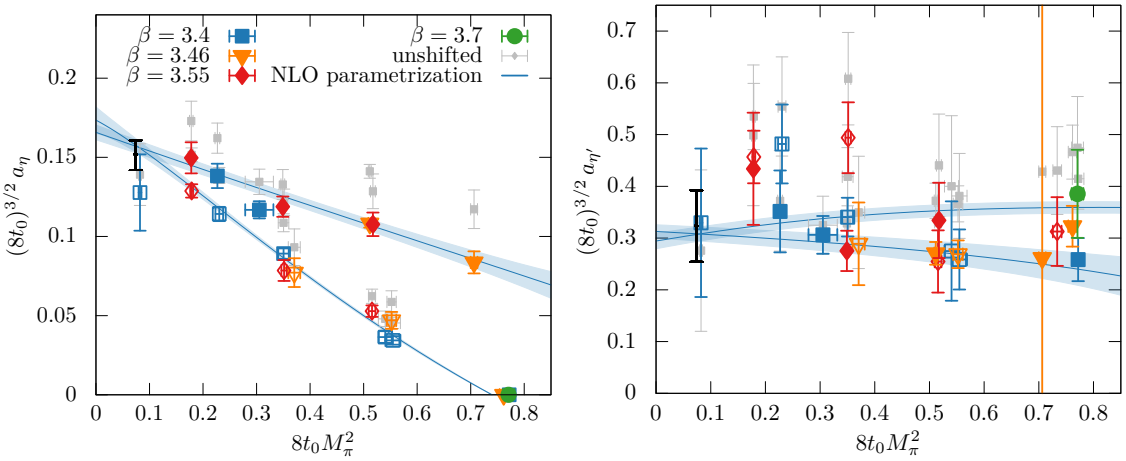


Figure 8. The anomalous gluonic matrix element a_η (left) and $a_{\eta'}$ (right) determined via the singlet AWI from fermionic matrix elements, eq. (6.14). The coloured points have been adjusted for lattice spacing effects, while the grey points indicate the unshifted data. The two curves correspond to the NLO large-N_c ChPT parametrization eqs. (B.9) and (B.11) 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.

for the additional improvement coefficients, setting $\tilde{d}_P = \tilde{b}_P$. The corresponding LECs read

$$\begin{aligned} L_5 &= 1.95\,(^{7}_{3})_{\text{stat}} \cdot 10^{-3}, & L_8 &= 0.97\,(^{4}_{6})_{\text{stat}} \cdot 10^{-3}, & M_0 &= 1.59\,(^{1}_{6})_{\text{stat}}\,(8t_0^{\chi})^{-1/2}, \\ F &= 0.1881\,(^{9}_{24})_{\text{stat}}\,(8t_0^{\chi})^{-1/2}, & \Lambda_1 &= -0.10\,(^{1}_{1})_{\text{stat}}, & \tilde{\Lambda} &= -0.21\,(^{3}_{4})_{\text{stat}}, \end{aligned} \tag{6.19}$$

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 (5.59), obtained from the fit to the masses 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

$$(8t_0^{\text{ph}})^{3/2} a_\eta = 0.1564\,(^{37}_{63}) \quad \text{and} \quad (8t_0^{\text{ph}})^{3/2} a_{\eta'} = 0.308\,(^{16}_{17})\,. \tag{6.20}$$

The NLO prediction eqs. (B.9) and (B.11), using the LECs of eq. (5.59), reads

$$(8t_0^{\text{ph}})^{3/2} a_\eta = 0.1609\,(^{17}_{27}) \quad \text{and} \quad (8t_0^{\text{ph}})^{3/2} a_{\eta'} = 0.383\,(^{11}_{17})\,. \tag{6.21}$$

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. (6.20) and (6.21) as the systematic error associated with taking the physical limit. We discuss the results and quote values in physical units in section 7.4.

JHEP08(2021)137

6.3 Renormalization of the anomaly term and the topological susceptibility

The singlet AWI has received a lot of attention, also in different settings, e.g., regarding the spin structure of the nucleon [85]. It would be desirable to validate the singlet AWI in our lattice study, also in view of confirming a consistent continuum limit extrapolation of the lattice data. Therefore, we will attempt to compute $a_{\eta^{(\prime)}}$ directly, destroying $\eta^{(\prime)}$ states by the topological charge density operator. This however requires an analysis of the renormalization of the anomaly term and its mixing with the derivative of the axialvector current. We start from the singlet AWI in the massless case

$$\partial_\mu \widehat{A}^0_\mu = \sqrt{2N_f}\,\widehat{\omega}. \tag{6.22}$$

Since $\omega(x)$ can be written as the divergence of a topological current, the associated Pontryagin index

$$Q = \int \mathrm{d}^4x\, \omega(x) \tag{6.23}$$

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.[9] $\partial_\mu A^0_\mu$ can and will mix into ω:[10]

$$\widehat{\omega} = Z_\omega \omega + Z_{\omega A} \partial_\mu A^0_\mu, \tag{6.24}$$

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 [86], without any factor Z_ω and the term containing $Z_{\omega A}$ cancels when computing the topological susceptibility τ, defined in eq. (6.25), 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 [87], 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. (6.22).[11] An alternative scheme of renormalizing the singlet axialvector

[9] In our conventions the kinetic term of the Lagrangian reads $\frac{1}{4g_s^2} F^a_{\mu\nu} F^a_{\mu\nu}$. In perturbative QCD the coupling is usually not absorbed into the field and then this term amounts to $-\frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu}$ instead. This would have introduced a factor of g_s^2 in front of $\omega(x)$ within eq. (6.22) and an additional scale dependence of $\widehat{\omega}$, governed by the QCD β-function.

[10] On the lattice with Wilson fermions, in principle ω can also mix with $a^{-1}P^0$. However, such power divergent terms are removed if we define ω after a gradient flow time t that we keep fixed in physical units as the continuum limit is approached [48]. Likewise, $\langle Q^2 \rangle$ will contain a contact term $\langle \omega(0)\omega(0) \rangle$, which diverges in the continuum limit. This divergence is in fact required to reproduce the correct topological susceptibility. However, this term could in principle also mix with lower dimensional operators. The latter possibility is eliminated too, by virtue of the gradient flow.

Also note that $\omega = \partial_\mu K_\mu$ will not mix into lower dimensional operators and neither does the gauge non-invariant Chern-Simons current K_μ interfere with the renormalization of A^0_μ.

[11] In [73] 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 $\gamma_{\omega A} = \gamma^s_A$.

JHEP08(2021)137

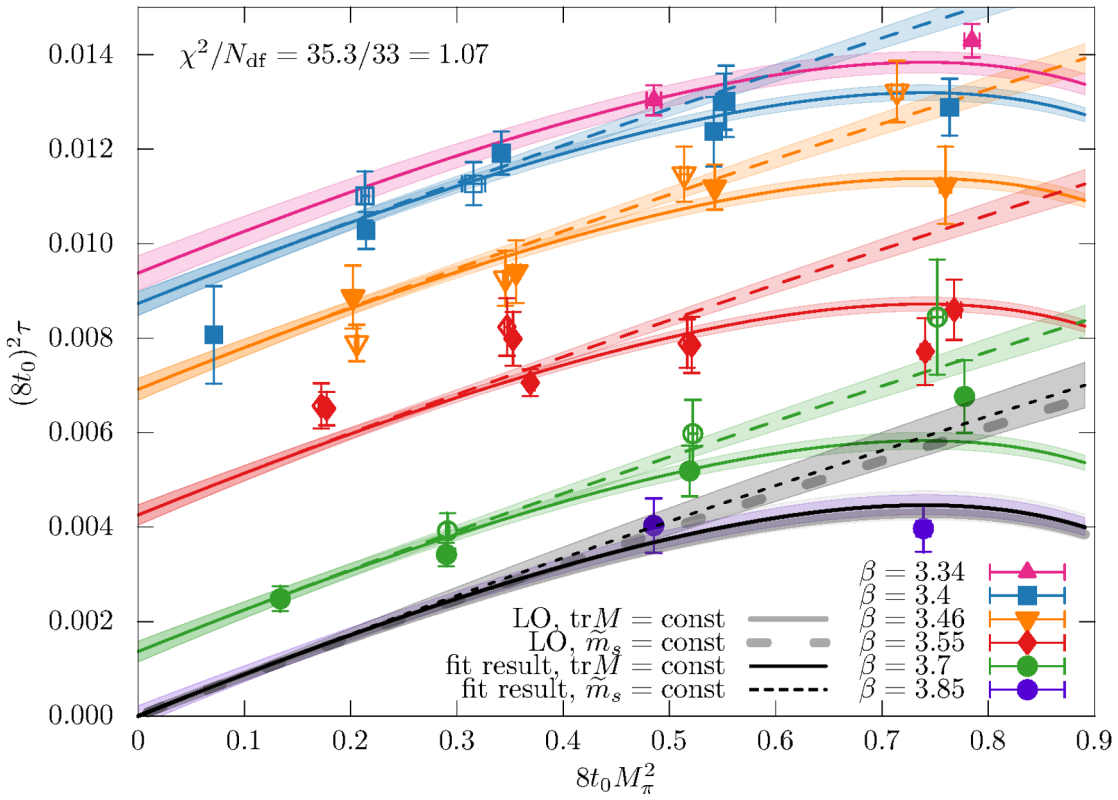


Figure 9. Topological susceptibility for many of the CLS ensembles described in [47]. Filled symbols mark ensembles that are simulated with a constant sum of quark masses (solid lines), open symbols correspond to ensembles with the strange quark mass fixed to approximately the physical value (dashed lines). Lines and shaded regions are the result of a fit to eq. (6.27). 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. (5.59), and setting $Z_\omega = 1$.

current is discussed in section 5.2. In that case, both Z_A^s and $Z_{\omega A}$ have no anomalous dimension. We remark that the $Z_{\omega A}\partial_\mu A_\mu^0$ term will not affect the topological susceptibility

$$\hat{\tau} = \sum_x \langle \widehat{\omega}(0)\widehat{\omega}(x)\rangle = \frac{1}{V}\sum_{x,y}\langle \widehat{\omega}(x)\widehat{\omega}(y)\rangle = \frac{\langle \widehat{Q}^2\rangle}{V} \tag{6.25}$$

since this term does not contribute to the volume sum, due to translational invariance.[12]

Within the numerical computations, we use the field-theoretical definition of ω extracted after evolving the gauge fields to the gradient flow time $\sqrt{8t} = \sqrt{8t_0^*} \approx 0.413\,\text{fm}$. For ensembles with open boundary conditions we keep the same distance $b \gtrsim 1.9\,\text{fm}$ to the boundaries when computing eq. (6.25) as we did in the computation of the fermion loops, cf. section 3.3.

[12]This also holds approximately for open boundary conditions in time, provided $L_t - 2b$ is much larger than the relevant correlation lengths: $\sum_{x_0=-L_t/2+b}^{L_t/2-b}\sum_{\vec{x}}\partial_\mu A_\mu^0(x) = \frac{1}{2a}\sum_{\vec{x}}\left[A_0^0(L_t/2-b) - A_0^0(-L_t/2+b)\right] \to 0$ for $L_t \to \infty$.

JHEP08(2021)137

As a first sanity check, we plot the topological susceptibility in figure 9, where we include most of the CLS ensembles analysed in [47], which adds additional points at finer and coarser lattice spacings. We find large cut-off effects with our definition of the susceptibility, shifting points considerably away from the $N_f = 3$ continuum expectation [3, 88],

$$\hat{\tau} = \frac{F^2}{2}\left(\frac{1}{2M_K^2 - M_\pi^2} + \frac{2}{M_\pi^2}\right)^{-1}. \tag{6.26}$$

Indeed, large cut-off effects have been reported in unquenched simulations previously [89–95]. To confirm $Z_\omega = 1$ numerically, we attempt a simple fit to

$$(8t_0)^2\tau = \frac{(8t_0)^2F^2}{2Z_\omega^2}\left(\frac{1}{2M_K^2 - M_\pi^2} + \frac{2}{M_\pi^2}\right)^{-1} + l_\tau^{(2)}\frac{a^2}{t_0^*} + l_\tau^{(3)}\frac{a^3}{(t_0^*)^{3/2}} + l_\tau^{(4)}\frac{a^4}{(t_0^*)^2}. \tag{6.27}$$

From this four parameter fit with $\chi^2/N_{\mathrm{df}} \approx 35.3/33 \approx 1.07$, we obtain in the continuum limit

$$\frac{\sqrt{8t_0^\chi}F}{Z_\omega} = 0.190(13). \tag{6.28}$$

When assuming $Z_\omega = 1$, this value agrees with our previous result $\sqrt{8t_0^\chi}F = 0.1866(48)$ (see eq. (5.59) of section 5.7). The coefficients of the terms parameterizing the lattice spacing dependence are

$$l_\tau^{(2)} = -0.072(10), \quad l_\tau^{(3)} = 0.355(34) \quad \text{and} \quad l_\tau^{(4)} = -0.324(30), \tag{6.29}$$

resulting in the non-monotonous behaviour observed in figure 9. The alternating sign also explains how the susceptibilities at our finest lattice spacing $a \approx 0.039\,\mathrm{fm}$ ($\beta = 3.85$) 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 (6.27) 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.

6.4 Direct determination of the gluonic matrix elements

We wish to check if the fermionic results that were obtained in section 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

$$a_{\mathcal{M}}(\mu) = 2\,Z_\omega\langle\Omega|\omega|\mathcal{M}\rangle + 2\frac{Z_{\omega A}}{Z_A^s}M_{\mathcal{M}}^2F_{\mathcal{M}}^0(\mu), \tag{6.30}$$

see eq. (6.24). 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 figure 10. Qualitatively these are in agreement with $Z_\omega = 1$ and we suspect that the two outliers are due to lattice artefacts.

JHEP08(2021)137

Based on the evidence presented above, we assume $Z_\omega = 1$, however, $Z_{\omega A}$ is not known and therefore comparing the direct determination (6.30) of the anomaly terms with the corresponding predictions from the singlet AWI eq. (6.13) 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. (6.30), we can isolate the renormalization scale independent ratio

$$\frac{Z_{\omega A}}{Z_A^s} = \frac{a_{\mathcal{M}} - 2\, Z_\omega \langle \Omega | \omega | \mathcal{M} \rangle}{2 M_{\mathcal{M}}^2 F_{\mathcal{M}}^0}. \tag{6.31}$$

Since F_η^0 in the denominator is close to zero and has large relative errors, we only use the η' matrix elements. We plot this ratio for $Z_\omega = 1$ in the right panel of figure 10. 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

$$\begin{aligned} (Z_{\omega_A}/Z_A^s)\,(\beta = 3.4) &= -0.036(13), & (Z_{\omega_A}/Z_A^s)\,(\beta = 3.46) &= -0.065(14), \\ (Z_{\omega_A}/Z_A^s)\,(\beta = 3.55) &= -0.043(16), & (Z_{\omega_A}/Z_A^s)\,(\beta = 3.7) &= -0.10(18). \end{aligned} \tag{6.32}$$

Using these values (and $Z_\omega = 1$), we evaluate eq. (6.30) 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 section 6.2 in the scatter plot figure 11. 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 section 6.3 and the singlet pseudoscalar matrix elements in section 6.2. The qualitative agreement suggests that some of the discretization effects may be similar for both definitions.

7 Summary and comparison to other results

In this section we summarize our results and compare them to other determinations from lattice or phenomenological studies. The meson masses, decay constants, large-N_c U(3) ChPT LECs (and their relation to their SU(3) equivalents) and pseudoscalar gluonic and fermionic matrix elements are presented. 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 $(8t_0^{\text{ph}})^{-1/2} = 475(6)\,\text{MeV}$ [49] and $(8t_0^{\chi})^{-1/2} = 470(7)$, see eqs. (3.5) and (3.6). 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.

JHEP08(2021)137

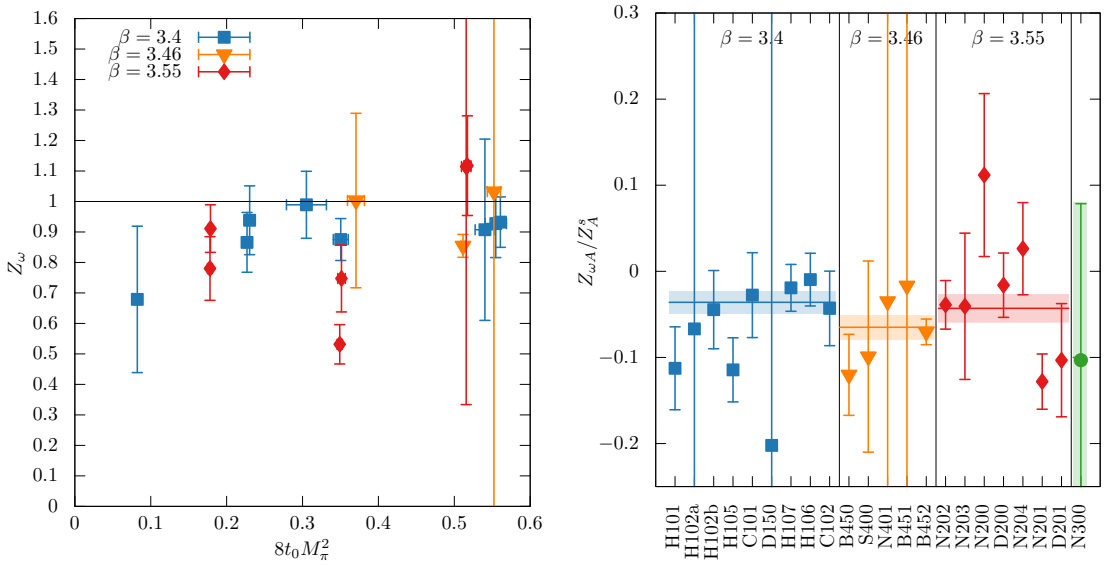


Figure 10. (Left) Z_ω from solving eq. (6.30). The $m_s = m_\ell$ points are not shown since in these cases $F_\eta^0 = 0$ and the equation system is singular. (Right) Values of $Z_{\omega A}/Z_A^s$ from eq. (6.31), assuming $Z_\omega = 1$.

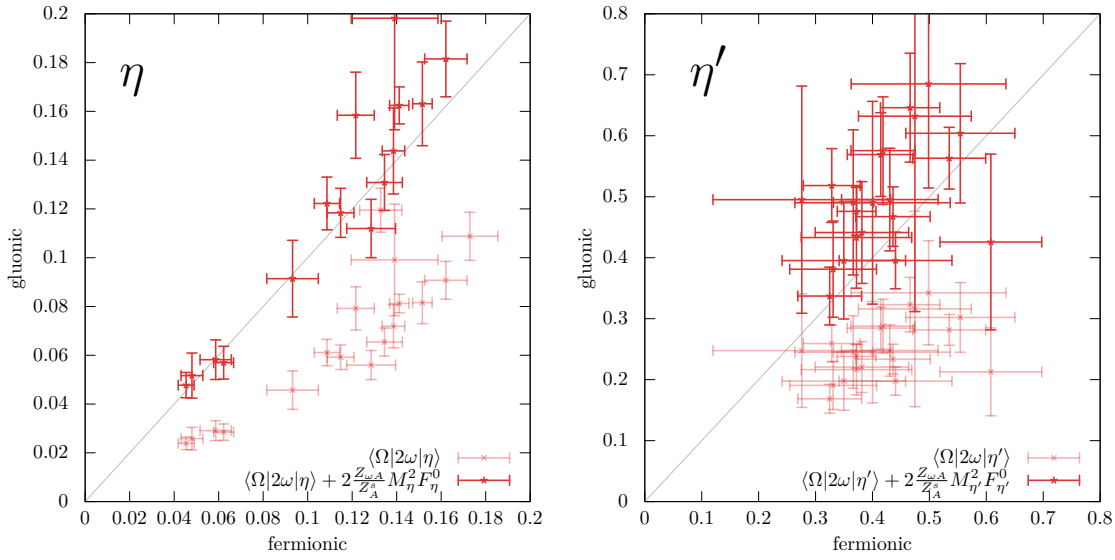


Figure 11. Scatter plot of the fermionic (eq. (6.14), horizontally) and gluonic (eq. (6.30), vertically) determinations of the gluonic matrix elements $a_{\eta^{(\prime)}}$. The mixing with the derivative of the axialvector current is non-negligible: the unrenormalized lattice matrix elements (pale red points) do not agree with the fermionic definition.

JHEP08(2021)137

7.1 The η and η' meson masses

Our final results for the masses of the η and η' mesons are (see section 5.6 and eq. (5.54))

$$M_\eta = 554.7\,(^{4.0}_{6.6})_{\text{stat}}\,(^{2.4}_{2.7})_{\text{syst}}\,(7.0)_{t_0}\text{, MeV} \quad \text{and} \tag{7.1}$$

$$M_{\eta'} = 929.9\,(^{12.9}_{6.0})_{\text{stat}}\,(^{22.9}_{3.3})_{\text{syst}}\,(11.7)_{t_0}\,\text{MeV}, \tag{7.2}$$

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,

$$\text{PDG [63]}: \qquad M_\eta^{\text{ph}} = 547.862(17)\,\text{MeV} \quad \text{and} \quad M_{\eta'}^{\text{ph}} = 957.78(6)\,\text{MeV}. \tag{7.3}$$

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 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 section 5.5. 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 figure 12, 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 [18] is affected by mistuning. The ETMC [19] and JLQCD [34] 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 [19, 35, 36] and JLQCD [34]. The latter $N_f = 2+1$ work utilizes gluonic correlation functions to determine the η' mass. A simple linear extrapolation is

JHEP08(2021)137

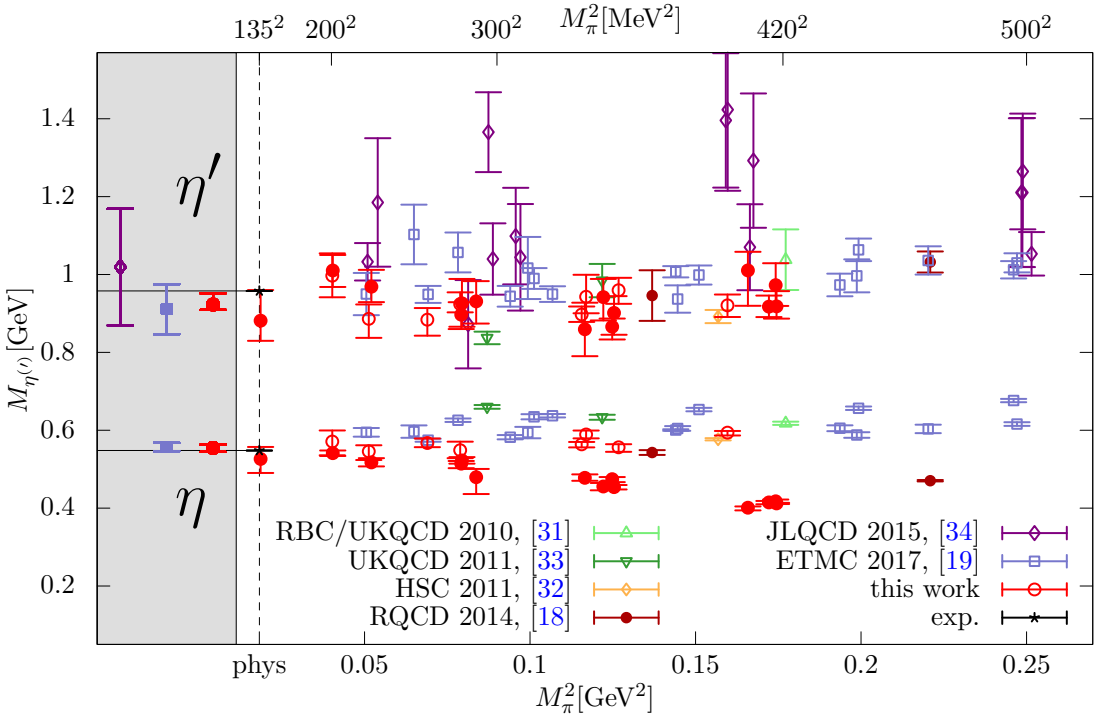


Figure 12. Recent $N_f = 2{+}1(+1)$ lattice results for the masses of the η and η' mesons. Most points have been simulated at approximately physical strange quark masses (open symbols), whereas in this work we also include an additional trajectory along which the average of the quark masses is kept constant (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 [34] (who do not give an estimate of M_η), ETMC [19] and this work.

performed which is justified in view of the large statistical errors. ETMC [19] 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.

JHEP08(2021)137

7.2 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 section 5.7. The two sets of results, detailed in tables 24 and 25 in appendix 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 section 5.2. 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 listed in tables 24 and 25.

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

$$F^8 = \quad 115.0\,(^{1.1}_{1.2})_{\text{stat}}\,(^{1.6}_{2.4})_{\text{syst}}\,(1.5)_{t_0}\,\text{MeV}, \tag{7.4}$$

$$\theta_8 = -25.8\,(^{1.2}_{2.1})_{\text{stat}}\,(^{2.2}_{0.3})^{\circ}_{\text{syst}}\,, \tag{7.5}$$

$$F^0(\mu = 2\,\text{GeV}) = \quad 100.1\,(^{7}_{1.9})_{\text{stat}}\,(^{2.0}_{2.7})_{\text{syst}}\,(1.3)_{t_0}\,\text{MeV}, \tag{7.6}$$

$$\theta_0 = -8.1\,(^{1.0}_{1.1})_{\text{stat}}\,(^{1.5}_{1.5})^{\circ}_{\text{syst}}\,, \tag{7.7}$$

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

$$F^{\ell}(\mu = 2\,\text{GeV}) = 88.28\,(^{1.20}_{2.02})_{\text{stat}}\,(^{3.00}_{1.74})_{\text{syst}}\,(1.12)_{t_0}\,\text{MeV}, \tag{7.8}$$

$$\phi_{\ell}(\mu = 2\,\text{GeV}) = 36.2\,(^{1.1}_{2.0})_{\text{stat}}\,(^{1.3}_{0.4})^{\circ}_{\text{syst}}\,, \tag{7.9}$$

$$F^{s}(\mu = 2\,\text{GeV}) = 124.3\,(^{1.7}_{1.6})_{\text{stat}}\,(^{2.7}_{4.3})_{\text{syst}}\,(1.6)_{t_0}\,\text{MeV}, \tag{7.10}$$

$$\phi_{s}(\mu = 2\,\text{GeV}) = 37.9\,(^{1.0}_{1.3})_{\text{stat}}\,(^{1.4}_{0.8})^{\circ}_{\text{syst}}\,, \tag{7.11}$$

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 [41, 96, 97] and to NLO in large-N_c ChPT it is equivalent to neglecting OZI-rule violating terms, specifically those involving Λ_1. At NLO the latter parameter is related to the angles and decay constants via [6, 41]

$$\frac{\sqrt{2}}{3} F_{\pi}^2 \Lambda_1 = F^{\ell} F^{s} \sin(\phi_{\ell} - \phi_s). \tag{7.12}$$

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

JHEP08(2021)137

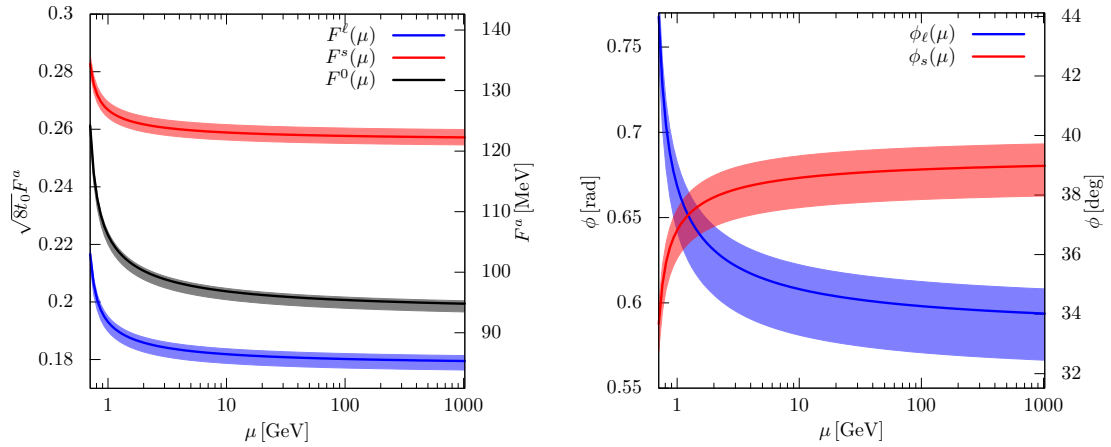


Figure 13. Scale dependence of the decay constants and their mixing angles in the flavour basis. The approximation $\phi_\ell \approx \phi_s \approx \phi$ is only valid for μ inbetween about 1 GeV and 2 GeV. In the same region the decay constants vary considerably with the scale. The (asymmetric) errors indicated by the coloured bands are statistical only.

Dependence on the QCD scale. In effect, the assumption $\Lambda_1 = 0$ renders the singlet decay constant independent of the scale since [6, 98]

$$\mu \frac{\mathrm{d}}{\mathrm{d}\mu} \frac{F_0(\mu)}{\sqrt{1+\Lambda_1(\mu)}} = 0. \tag{7.13}$$

The results in eqs. (7.8)–(7.11) show that at $\mu = 2\,\mathrm{GeV}$ the angles almost agree within errors. However, our estimate for $\Lambda_1 = -0.25(5)$ (see section 5.7) 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 figure 13. 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 figure 14, where we display both OZI violating LECs [99],

$$\Lambda_1(\mu) = \left(\frac{Z_A^s(\mu)}{Z_A^{s\prime}}\right)^2 (1+\Lambda_1(\mu=\infty)) - 1, \quad \Lambda_2(\mu) = \frac{Z_A^s(\mu)}{Z_A^{s\prime}}(1+\Lambda_2(\mu=\infty)) - 1. \tag{7.14}$$

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 [81, 100].

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 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

JHEP08(2021)137

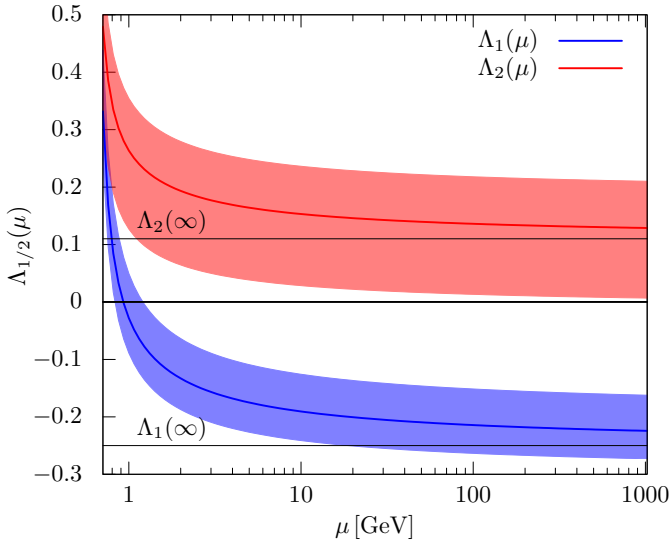


Figure 14. Scale dependence of the large-N_c ChPT LECs Λ_1 and Λ_2.

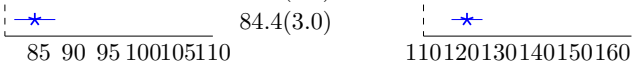
processes dominated by different physical scales, the predictions for, e.g., F^0 will be affected. Clearly, results obtained using the FKS scheme should be compared at the scale where Λ_1 vanishes. One possibility to mitigate this problem is to compare results in the octet/singlet basis (where only F^0 depends on μ) and form the scale independent combination $F^0/\sqrt{1+\Lambda_1}$ [6]. Our result for the latter reads

$$F^0/\sqrt{1+\Lambda_1} = 107.3\,(^{1.5}_{7})_{\text{stat}}\,(^{1.3}_{1.3})_{\text{syst}}\,(1.4)_{t_0}\,\text{MeV}. \tag{7.15}$$

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 tables 16 and 17, 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 [6], 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 [97] 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 [19, 36] 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. [81] 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. [107] extended the analysis by also utilizing the decay constant results from ETMC [19]. However, additional constraints on the parameters still seem to be necessary in order to obtain stable NNLO results. Bickert et al. [42] also perform an NNLO analysis, in this case combining LECs obtained from the literature and derived from

JHEP08(2021)137

ref	F^ℓ/MeV	F^s/MeV
Benayoun et al. [101]	—	—
Escribano and Frere [102]	100.4(2.8)	152.8(5.5)
Escribano et al. [103]	97.6(1.0)	144(20)
Chen et al. [104]	105.9(3.7)	143.6(5.5)
Escribano et al. [105]	98.7(9)	128(13)
Escribano et al. [98]	94.9(3.7)	125.3(3.7)
Leutwyler [6]	—	—
Feldmann [97]	98.5(1.8)	123.4(5.5)
Guo et al. [81] NLO-A	94.1(2.5)	122.3(5.2)
Guo et al. [81] NNLO-B	92.7(1.4)	139(15)
Bickert et al. [42] NLO-I	—	—
[42] NNLO w/o Ci ($\mu_{\rm EFT} = 1$GeV)	—	—
[42] NNLO w/ Ci ($\mu_{\rm EFT} = 1$GeV)	—	—
Ding et al. [106]	101	138
ETMC [19]	88.4(5.4)	125.6(2.6)
Gu et al. [107] 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. [101]	—	—
Escribano and Frere [102]	39.9(1.3)°	41.4(1.4)°
Escribano et al. [103]	40.3(1.8)°	40.3(1.8)°
Chen et al. [104]	34.5(1.8)°	36.0(1.4)°
Escribano et al. [105]	39.3(1.2)°	39.2(1.2)°
Escribano et al. [98]	39.6(2.3)°	40.8(1.8)°
Leutwyler [6]	—	—
Feldmann [97]	39.3(1.0)°	39.3(1.0)°
Guo et al. [81] NLO-A	40.4(4.8)°	39.9(2.8)°
Guo et al. [81] NNLO-B	35.8(1.2)°	37.1(1.4)°
Bickert et al. [42] NLO-I	—	—
[42] NNLO w/o Ci ($\mu_{\rm EFT} = 1$GeV)	—	—
[42] NNLO w/ Ci ($\mu_{\rm EFT} = 1$GeV)	—	—
Ding et al. [106]	42.8°	42.8°
ETMC [19]	38.8(3.3)°	38.8(3.3)°
Gu et al. [107] 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)°

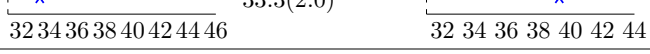


Table 16. Comparison of recent phenomenological and lattice results for the (scale dependent) decay constants in the angle representation for the light/strange flavour basis. The results from this work are presented at three different scales.

JHEP08(2021)137

ref	F^8/MeV	F^0/MeV
Benayoun et al. [101]	125.2(9)	121.5(2.8)
Escribano and Frere [102]	139.0(4.6)	118.8(3.7)
Escribano et al. [103]	—	—
Chen et al. [104]	133.5(3.7)	117.8(5.5)
Escribano et al. [105]	112.4(9.2)	105.9(5.5)
Escribano et al. [98]	117.0(1.8)	105.0(4.6)
Leutwyler [6]	118	—
Feldmann [97]	116.0(3.7)	107.8(2.8)
Guo et al. [81] NLO-A	113.2(4.4)	104.9(2.9)
Guo et al. [81] NNLO-B	126(12)	109.1(6.0)
Bickert et al. [42] NLO-I	116.0(9)	—
[42] NNLO w/o Ci ($\mu_{\rm EFT} = 1$GeV)	117.9(1.8)	—
[42] NNLO w/ Ci ($\mu_{\rm EFT} = 1$GeV)	109(7)	—
Ding et al. [106]	123.4	116.0
ETMC [19]	—	—
Gu et al. [107] NNLO-A9p(F_π)	113.1(2.1)	106.0(4.4)
eq. (7.16)	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. [101]	$-20.4(1.0)^\circ$	$-0.1(1.0)^\circ$
Escribano and Frere [102]	$-23.8(1.4)^\circ$	$-2.4(1.9)^\circ$
Escribano et al. [103]	—	—
Chen et al. [104]	$-26.7(1.8)^\circ$	$-11.0(1.0)^\circ$
Escribano et al. [105]	$-21.3(3.5)^\circ$	$-11.3(3.9)^\circ$
Escribano et al. [98]	$-21.2(1.9)^\circ$	$-6.9(2.4)^\circ$
Leutwyler [6]	-20°	-4°
Feldmann [97]	$-21.2(1.9)^\circ$	$-9.2(1.7)^\circ$
Guo et al. [81] NLO-A	$-21.5(4.5)^\circ$	$-7.2(2.5)^\circ$
Guo et al. [81] NNLO-B	$-27.9(1.7)^\circ$	$-6.8(3.8)^\circ$
Bickert et al. [42] NLO-I	$-21.7(7)^\circ$	$-0.5(7)^\circ$
[42] NNLO w/o Ci ($\mu_{\rm EFT} = 1$GeV)	$-12.6(6.1)^\circ$	$-6.3(6.5)^\circ$
[42] NNLO w/ Ci ($\mu_{\rm EFT} = 1$GeV)	$-34(22)^\circ$	$-33(24)^\circ$
Ding et al. [106]	-21°	-2.8°
ETMC [19]	—	—
Gu et al. [107] NNLO-A9p(F_π)	$-26.1(2.5)^\circ$	$-7.0(2.1)^\circ$
eq. (7.16)	—	—
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 17. 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 $F_{\pi^+} = 92.1\,\text{MeV}$ [63] to convert decay constants given as a multiple of F_π and eq. (7.16) refers to the NLO result from literature pion and kaon decay constants. The results from this work are presented at three different scales.

JHEP08(2021)137

ref	Λ_1	$F^0/\sqrt{1+\Lambda_1}$/MeV
Benayoun et al. [101]	0.20(4)	*110.9(4.8)*
Escribano and Frere [102]	0.34(10)	*102.6(8.2)*
Escribano et al. [103]	0	—
Chen et al. [104]	—	—
Escribano et al. [105]	0	—
Escribano et al. [98]	0.01(13)	*105(11)*
Leutwyler [6]	—	101
Feldmann [97]	0.0(3)	107.8(2.8)
Guo et al. [81] NLO-A	0.02(8)	*102.8(7.0)*
Guo et al. [81] NNLO-B	−0.04(14)	*111(14)*
Bickert et al. [42] NLO-I	—	104.1(0)
[42] NNLO w/o Ci ($\mu_{\rm EFT}$ = 1GeV)	0	79.2(9)
[42] NNLO w/ Ci ($\mu_{\rm EFT}$ = 1GeV)	0	76.4(9)
Ding et al. [106]	—	—
ETMC [19]	0	—
Gu et al. [107] NNLO-A9p(F_π)	0.24(21)	*95(20)*
eq. (7.16)	—	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)

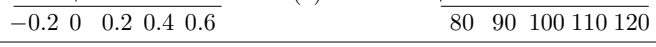


Table 18. Comparison of determinations of the (scale dependent) large-N_c ChPT LEC Λ_1 and the scale independent combination of Λ_1 and F^0, where eq. (7.16) refers to the NLO result from literature pion and kaon decay constants. The values indicated in Italics have been computed by us from F_0 and Λ_1 with naive error propagation. The results from this work are presented at three different scales.

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 [101, 108] employs vector meson dominance, while others involve more phenomenologically driven extractions, for example, Escribano et al. [98, 102, 103, 105], 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 ChPT is made and allows to predict some of the LECs. Chen et al. [104] 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. [106] employ coupled gap and Bethe-Salpeter equations.

We also include the values $F^8 = 115.2(1.3)\,\text{MeV}$ and $F^0/\sqrt{1+\Lambda_1} = 104.3(1.1)\,\text{MeV}$ in tables 17 and 18 (labelled as "eq. (7.16)"). These are obtained from the identities

$$(F^8)^2 = \frac{4F_K^2 - F_\pi^2}{3} \quad \text{and} \quad (F^0)^2 = \frac{2F_K^2 + F_\pi^2}{3}\,(1+\Lambda_1)\,, \tag{7.16}$$

JHEP08(2021)137

which hold at NLO in large-N_c ChPT [6], using the values $F_{K^+}/F_{\pi^+} = 1.193(2)$ [109] and $F_{\pi^+} = 92.1(8)\,\text{MeV}$ [63, 109] 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 phenomenological NLO large-N_c ChPT determinations. 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 [42, 81]. 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 [18]. 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 [36]. This study was updated in [19] 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 section 7.1. 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 [97]. 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. (7.16) into the flavour basis results in

$$(F^\ell)^2 = F_\pi^2 + \frac{2}{3}\Lambda_1(2F_K^2 + F_\pi^2) \quad \text{and} \quad (F^s)^2 = 2F_K^2 - F_\pi^2 - \frac{1}{3}\Lambda_1(2F_K^2 + F_\pi^2). \tag{7.17}$$

Setting $\Lambda_1 = 0$ and plugging the experimental ratio $F_{K^+}/F_{\pi^+} = 1.193(2)$ [109] into eq. (7.17) gives

$$F^\ell/F_\pi = 1 \quad \text{and} \quad F^s/F_K = \sqrt{2 - \frac{F_{\pi^+}^2}{F_{K^+}^2}} = 1.139(2). \tag{7.18}$$

In the FKS model the flavour basis AWIs (6.5) factorize into anomalous and non-anomalous contributions and one obtains the relations [96, 97]

$$2m_\ell H^\ell = M_\pi^2 F^\ell \qquad \text{and} \qquad 2m_s H^s = (2M_K^2 - M_\pi^2)F^s, \tag{7.19}$$

JHEP08(2021)137

where $H^q = \sqrt{(H^q_\eta)^2 + (H^q_{\eta'})^2}$ for $q = \ell, s$ and the $H^q_{\eta^{(\prime)}}$ are defined in eq. (4.19). Therefore, in this approximation, in the flavour basis the mass mixing angle $\phi_{\mathrm{SU}(3)}$, the mixing angles ϕ^ℓ and ϕ^s and the mixing angles for pseudoscalar matrix elements are all equal: $\phi_{\mathrm{SU}(3)} = \phi = \phi_s = \phi_\ell = \phi_{\mathrm{PS}}$, where [110]

$$\phi_{\mathrm{SU}(3)} = \arcsin \sqrt{\frac{\left(M^2_{\eta'} - (2M^2_K - M^2_\pi)\right)(M^2_\eta - M_\pi)}{(M^2_{\eta'} - M^2_\eta)(2M^2_K - 2M^2_\pi)}} = 42.4^\circ. \tag{7.20}$$

ETMC [19] compute the mixing angle from the averaged ratios of their pseudoscalar matrix elements:

$$\tan\phi_{\mathrm{PS}} = \sqrt{\tan\phi^\ell_{\mathrm{PS}} \tan\phi^s_{\mathrm{PS}}} \quad \text{with} \quad \tan\phi^\ell_{\mathrm{PS}} = \frac{H^\ell_{\eta'}}{H^\ell_\eta} \quad \text{and} \quad \tan\phi^s_{\mathrm{PS}} = -\frac{H^s_\eta}{H^s_{\eta'}}. \tag{7.21}$$

The decay constants are obtained, using eq. (7.19). 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

$$\text{ETMC [19]:} \qquad F^\ell/F_\pi = 0.960(59), \quad F^s/F_K = 1.143(24) \quad \text{and} \quad \phi = 38.8(3.3)^\circ, \tag{7.22}$$

which within errors agrees with the values computed within the FKS model in eqs. (7.18) and (7.20) 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 $\mu = 1\,\mathrm{GeV}$ (see table 16), 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 tables 16, 17 and 18 (see also eqs. (7.18) and (7.20)) show that our ab-initio values determined at $\mu = 1\,\mathrm{GeV}$ agree well with those derived by employing the FKS scheme. We can go further and directly check the relations (7.19), (7.21) and (7.17), also away from the physical point. Figure 15 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 $\mu = 1\,\mathrm{GeV}$. 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 $\sqrt{2F^2_K - F^2_\pi}$. 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.

JHEP08(2021)137

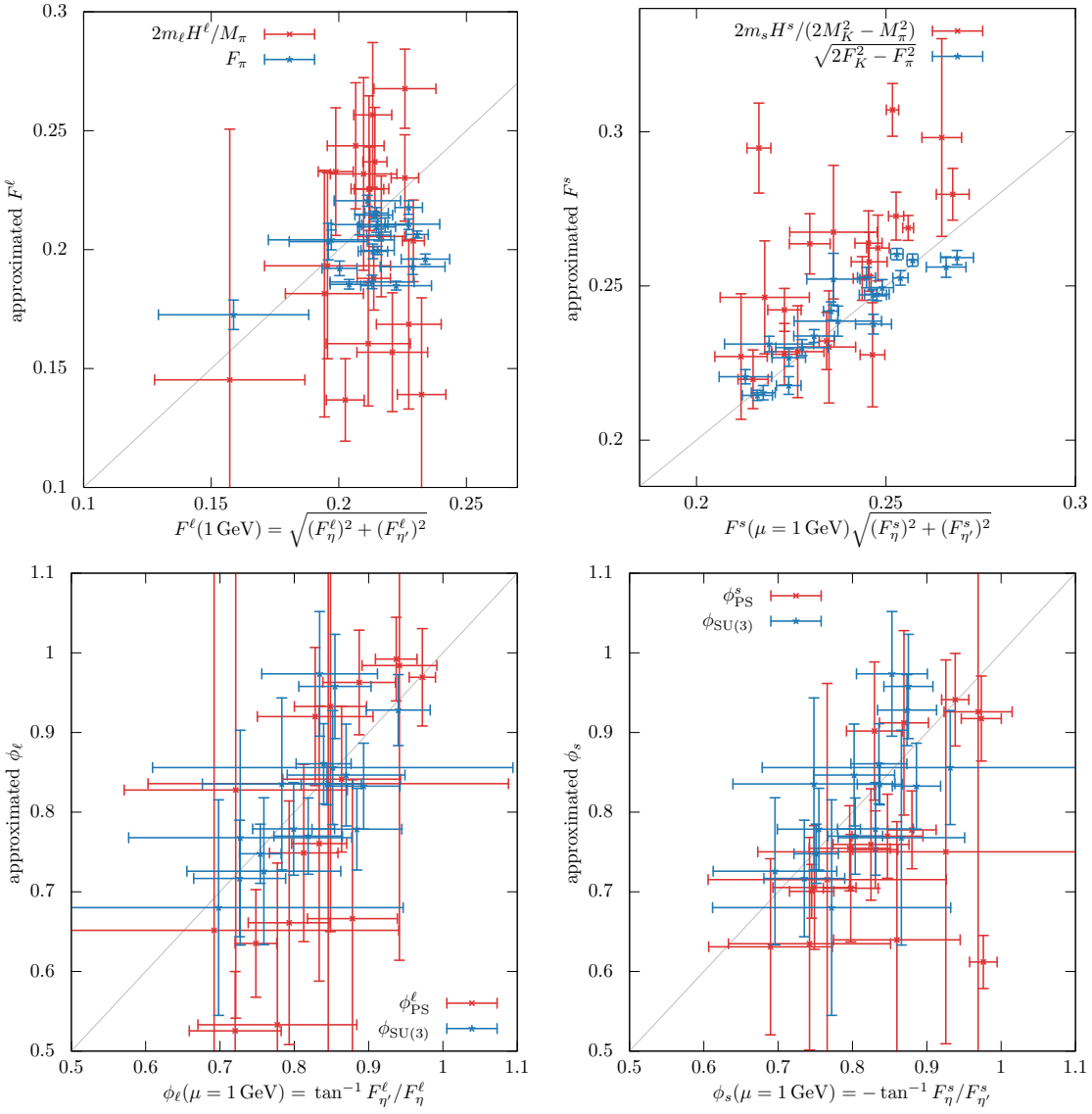


Figure 15. Light and strange decay constants (top) and angles (bottom) determined on each ensemble at the QCD scale $\mu = 1\,\mathrm{GeV}$. The central values and errors in the x-direction indicate the results obtained directly from the axialvector matrix elements, while the position and error in the y-direction indicates the values constructed from the pseudoscalar matrix elements (red, eqs. (7.19) and (7.21)). For the decay constants, the expectations derived from combinations of the pion and kaon masses and decay constants (blue, eq. (7.17)) are also displayed, while for the angles $\phi_{\mathrm{SU(3)}}$ (eq. (7.20)) is also shown. The blue points have been shifted slightly to the right for better visibility.

JHEP08(2021)137

Our direct QCD results can be used as input to theory calculations and we consider one important example in section 7.5, namely the light-cone sum rule computation of the $\gamma\gamma^* \to \eta^{(\prime)}$ transition form factors.

7.3 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:

$$M_0 = 761\,(^{13}_{21})_{\text{stat}}\,(^{18}_{11})_{\text{syst}}\,(11)_{t_0}\,\text{MeV}, \tag{7.23}$$

$$\Lambda_1 = -0.25\,(^{1}_{4})_{\text{stat}}\,(^{6}_{2})_{\text{syst}}\,, \tag{7.24}$$

$$\Lambda_2 = 0.11\,(^{5}_{5})_{\text{stat}}\,(^{7}_{10})_{\text{syst}}\,, \tag{7.25}$$

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 section 7.2. To aid comparison with literature values we consider the scale independent combinations [6]:

$$M_0/\sqrt{1+\Lambda_1} = 877\,(^{12}_{10})_{\text{stat}}\,(^{21}_{8})_{\text{syst}}\,(13)_{t_0}\,\text{MeV} \quad \text{and} \quad \tilde{\Lambda} = \Lambda_1 - 2\Lambda_2 = -0.46(19). \tag{7.26}$$

Previous determinations of these quantities include:

Leutwyler [6]:	$M_0/\sqrt{1+\Lambda_1} \approx 899\,\text{MeV}$	and	$\tilde{\Lambda} = -0.31$,	(7.27)
Benayoun et al. [101]:			$\tilde{\Lambda} = -0.42(6)$,	(7.28)
Guo et al. [81]:	$M_0/\sqrt{1+\Lambda_1} = 804(80)\,\text{MeV}$	and	$\tilde{\Lambda} = -0.37(17)$,	(7.29)
Bickert et al. [42]:	$M_0/\sqrt{1+\Lambda_1} = 950(7)\,\text{MeV}$	and	$\tilde{\Lambda} = -0.34(5)$,	(7.30)

where except for [42], 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 [81], which utilizes lattice data, however, considering the large uncertainty quoted in this reference, there is no significant disagreement. The determination from [42] 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 [42] when comparing predictions for the decay constants and angles, cf. table 17. 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

$$F = 87.71\,(^{1.44}_{1.57})_{\text{stat}}\,(^{2.69}_{81})_{\text{syst}}\,(1.31)_{t_0}\,\text{MeV}. \tag{7.31}$$

JHEP08(2021)137

This result agrees within errors with the NLO values presented in [81] and [42]:

$$\text{Guo et al. [81]:}\quad F = 92.1(6)\,\text{MeV}, \qquad \text{Bickert et al. [42]:}\quad F = 90.73(11)\,\text{MeV}. \tag{7.32}$$

However, the corresponding NNLO analyses give somewhat lower values of F,

$$\text{Guo et al. [81]:}\quad F = 80.8(6.3)\,\text{MeV}, \qquad \text{Bickert et al. [42]:}\quad F = 79.46(6.59)\,\text{MeV}, \tag{7.33}$$

which within errors still agree with our result (7.31). In the simulations of [111] 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

$$\text{Hernandez et al. [111]:}\quad F(N_f = 3, N_c = 3) = 68(7)\,\text{MeV} \tag{7.34}$$

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 [112], 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 [113]. Thus, neglecting NNLO contributions including L_4-terms may lead to larger values of F. However, also our fits in appendix A, including only the NNLO loop contributions (see the discussion of section 5.5 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

$$L_5 = 1.66\,(11)_{\text{stat}}\,(20)_{\text{syst}} \cdot 10^{-3} \quad \text{and} \quad L_8 = 1.08\,(9)_{\text{stat}}\,(9)_{\text{syst}} \cdot 10^{-3}. \tag{7.35}$$

These values agree reasonably well with those obtained from other NLO large-N_c ChPT studies, e.g.,

$$\text{Leutwyler [6]:}\quad L_5 = 2.2 \cdot 10^{-3} \quad \text{and} \quad L_8 = 1.0 \cdot 10^{-3}, \tag{7.36}$$

$$\text{Guo et al. [81]:}\quad L_5 = 1.47(29) \cdot 10^{-3} \quad \text{and} \quad L_8 = 1.08(6) \cdot 10^{-3}, \tag{7.37}$$

$$\text{Bickert et al. [42]:}\quad L_5 = 1.86(6) \cdot 10^{-3} \quad \text{and} \quad L_8 = 0.78(5) \cdot 10^{-3}. \tag{7.38}$$

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 [42, 83, 114]

$$L_5(\mu_{\text{EFT}}) = L_5^{\text{SU(3)}}(\mu_{\text{SU(3)}}) + \frac{3}{8}\frac{1}{16\pi^2}\ln\left(\frac{\mu_{\text{SU(3)}}}{\mu_{\text{EFT}}}\right) \tag{7.39}$$

$$L_8(\mu_{\text{EFT}}) = L_8^{\text{SU(3)}}(\mu_{\text{SU(3)}}) + \frac{5}{48}\frac{1}{16\pi^2}\ln\left(\frac{\mu_{\text{SU(3)}}}{\mu_{\text{EFT}}}\right) + \frac{1}{12}\frac{1}{16\pi^2}\ln\left(\frac{\mu_{\text{match}}}{\mu_{\text{EFT}}}\right), \tag{7.40}$$

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 [115] from a $N_f = 2 + 1$

JHEP08(2021)137

	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	$a^2_{\eta'}/a^2_\eta$
$\mu = 1\,\text{GeV}$	$-22.1\,(^{3}_{5})_{\text{stat}}\,(2.8)^\circ_{\text{syst}}$	$6.09\,(^{27}_{53})_{\text{stat}}\,(2.05)_{\text{syst}}$
$\mu = 2\,\text{GeV}$	$-24.0\,(^{4}_{1.0})_{\text{stat}}\,(3.2)^\circ_{\text{syst}}$	$5.03\,(^{19}_{45})_{\text{stat}}\,(1.94)_{\text{syst}}$
$\mu = 10\,\text{GeV}$	$-25.3\,(^{4}_{1.1})_{\text{stat}}\,(3.6)^\circ_{\text{syst}}$	$4.46\,(^{16}_{41})_{\text{stat}}\,(1.86)_{\text{syst}}$
$\mu = \infty$	$-26.9\,(^{4}_{1.2})_{\text{stat}}\,(4.1)^\circ_{\text{syst}}$	$3.88\,(^{14}_{38})_{\text{stat}}\,(1.78)_{\text{syst}}$

Table 19. Gluonic matrix elements of the η and η' and combinations thereof at various scales.

lattice study of the pion and kaon masses and decay constants. Here, we quote the values presented in the FLAG review [109] for $\mu_{\text{SU(3)}} = 0.770\,\text{GeV}$

$$\text{MILC [115]:}\quad L_5^{\text{SU(3)}} = 0.98(38)\cdot 10^{-3}\quad\text{and}\quad L_8^{\text{SU(3)}} = 0.42(27)\cdot 10^{-3}. \tag{7.41}$$

The agreement with our (scale independent) results improves for $\mu_{\text{EFT}} < 0.770\,\text{GeV}$ and $\mu_{\text{match}} > 0.770\,\text{GeV}$.

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 appendix A. 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. (A.9)) that are slightly smaller and larger than our NLO values quoted in eq. (7.35), 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 [42, 81].

7.4 Pseudoscalar gluonic and fermionic matrix elements

We determined the anomaly matrix elements a_η and $a_{\eta'}$ in section 6.2 from a fit to combinations of axialvector and pseudoscalar matrix elements, eq. (6.13). 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 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

JHEP08(2021)137

axialvector current. With $N_f = 3$ active quark flavours, at 2 GeV we obtain:

$$a_\eta(\mu = 2\,\text{GeV}) = 0.01700\,(^{40}_{69})_{\text{stat}}\,(48)_{\text{syst}}\,(66)_{t_0}\ \text{GeV}^3, \tag{7.42}$$

$$a_{\eta'}(\mu = 2\,\text{GeV}) = 0.0381\,(^{18}_{17})_{\text{stat}}\,(80)_{\text{syst}}\,(17)_{t_0}\ \text{GeV}^3. \tag{7.43}$$

The systematic error is computed as the difference between our results from a direct NLO fit to the $a_{\eta^{(\prime)}}$ data (see eq. (6.20)), that included lattice correction terms, and the continuum NLO large-N_c ChPT prediction (see eq. (6.21)), 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 table 19 and compare to literature values in table 20, where the scale is not specified. These analyses are based on, e.g., QCD sum rule calculations [116, 117], large-N_c ChPT [9, 97] and related state mixing models that include a pseudoscalar glueball [118, 119]. We find agreement with the references that give error estimates, with the exception of [9].

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

$$\begin{aligned} m_\ell H^\ell_\eta &= 0.0021\,(^{3}_{2})_{\text{stat}}\,(13)_{\text{syst}}(0)_{t_0}\ \text{GeV}^3, & m_s H^s_\eta &= -0.0173\,(^{3}_{2})_{\text{stat}}\,(17)_{\text{syst}}(7)_{t_0}\ \text{GeV}^3, \\ m_\ell H^\ell_{\eta'} &= 0.0045\,(^{10}_{8})_{\text{stat}}\,(40)_{\text{syst}}(0)_{t_0}\ \text{GeV}^3, & m_s H^s_{\eta'} &= 0.0309\,(^{15}_{5})_{\text{stat}}\,(50)_{\text{syst}}(10)_{t_0}\ \text{GeV}^3 \end{aligned} \tag{7.44}$$

for the pseudoscalar fermionic matrix elements, where again the systematic error is the difference with respect to the NLO ChPT predictions eqs. (B.5)–(B.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,

$$\begin{aligned} \text{Feldmann [96]:}\quad m_\ell H^\ell_\eta &= 0.0010\,\text{GeV}^3, & m_s H^s_\eta &= -0.026\,\text{GeV}^3, \\ m_\ell H^\ell_{\eta'} &= 0.0008\,\text{GeV}^3, & m_s H^s_{\eta'} &= 0.032\,\text{GeV}^3, \end{aligned} \tag{7.45}$$

and the very similar numbers of a QCD sum rule calculation,

$$\begin{aligned} \text{Singh [117]:}\quad m_\ell H^\ell_\eta &= 0.00105(14)\,\text{GeV}^3, & m_s H^s_\eta &= -0.0284(55)\,\text{GeV}^3, \\ m_\ell H^\ell_{\eta'} &= 0.000782(250)\,\text{GeV}^3, & m_s H^s_{\eta'} &= 0.0379(71)\,\text{GeV}^3. \end{aligned} \tag{7.46}$$

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 [97],

$$\theta_y(\mu = 2\,\text{GeV}) = -\arctan\left(\frac{a_\eta(2\,\text{GeV})}{a_{\eta'}(2\,\text{GeV})}\right) = -24.0\,(^{4}_{1.0})_{\text{stat}}\,(3.2)^\circ_{\text{syst}}\,. \tag{7.47}$$

JHEP08(2021)137

ref	a_η/GeV^3	$a_{\eta'}/\text{GeV}^3$
Novikov et al. [116]	0.021	0.035
Feldmann [97]	0.023	0.058
Beneke and Neubert [9]	0.022(2)	0.057(2)
Cheng et al. [118]	0.026(28)	0.054(57)
Singh [117]	0.0220(50)	0.037(10)
Qin et al. [119]	0.016	0.051
Ding et al. [106]	0.024	0.051
this work at $\mu = 1\text{GeV}$	0.0172(10)	0.0424(84)
this work at $\mu = 2\text{GeV}$	0.0170(10)	0.0381(84)
this work at $\mu = \infty$	0.0168(10)	0.0330(83)

Table 20. Literature values for the anomaly matrix elements in comparison with our results at various scales. Note that the error bars of [118] 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 [116, 120],

$$R(J/\psi) = \frac{\Gamma[J/\psi \to \eta'\gamma]}{\Gamma[J/\psi \to \eta\gamma]} = \frac{a_{\eta'}^2}{a_\eta^2}\left(\frac{k_{\eta'}}{k_\eta}\right)^3, \tag{7.48}$$

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 [63], we obtain $(k_{\eta'}/k_\eta)^3 \approx 0.8137$. Our result for the ratio $a_{\eta'}^2/a_\eta^2$, listed in table 19, gives at $\mu = 2\,\text{GeV}$

$$R(J/\psi, \mu = 2\,\text{GeV}) = 5.03\,(^{19}_{45})_{\text{stat}}\,(1.94)_{\text{syst}}\,. \tag{7.49}$$

Note that $a_{\eta'}$ depends strongly on the scale, see table 19, 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 $\Gamma(J/\psi \to \eta\gamma)/\Gamma_{\text{total}} = 1.108(27)\cdot 10^{-3}$ and $\Gamma(J/\psi \to \eta'\gamma)/\Gamma_{\text{total}} = 5.25(7)\cdot 10^{-3}$ result in

$$\text{PDG [63]:} \quad R(J/\psi) = 4.74(13), \tag{7.50}$$

which is very close to our value with $N_f = 3$ at $\mu = 2\,\text{GeV}$. Clearly, a more precise comparison should take $\mathcal{O}(\alpha_s)$ corrections into account.

7.5 Transition form factors $F_{\gamma\gamma^*\to\eta}(Q^2)$ and $F_{\gamma\gamma^*\to\eta'}(Q^2)$

Photoproduction is the simplest hard process involving $\eta^{(\prime)}$ mesons. The corresponding transition form factors $F_{\gamma\gamma^*\to\eta}(Q^2)$ and $F_{\gamma\gamma^*\to\eta'}(Q^2)$ at large photon virtualities $Q^2 = |q^2|$ have been studied, e.g., in [10, 12, 106] and assessed phenomenologically for instance in [98, 105]. 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

JHEP08(2021)137

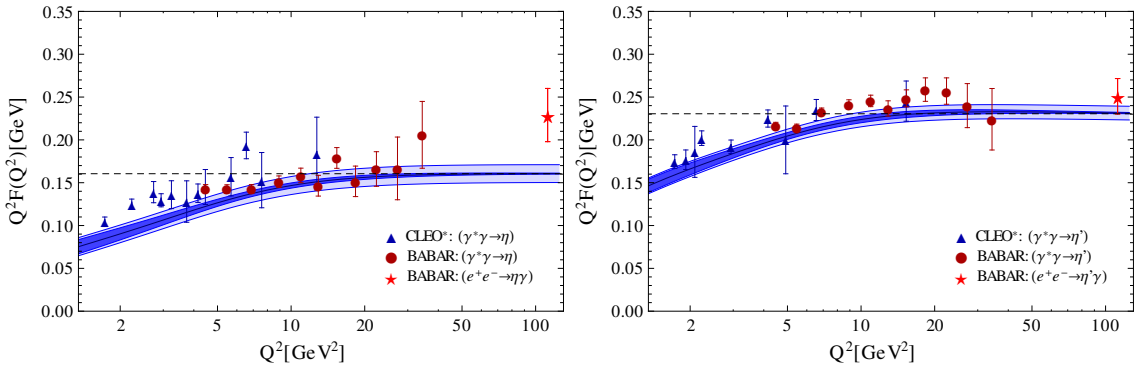


Figure 16. Predictions for the transition form factor $F_{\gamma\gamma^*\to\eta}(Q^2)$ (left) and $F_{\gamma\gamma^*\to\eta'}(Q^2)$ (right) following the approach in [12], using the decay constants determined in our work as input parameters.

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 low scale, and, second, whether η' contains a significant two-gluon component, see, e.g., [10, 12] for a discussion.

In what follows we briefly discuss the impact of our results on predictions of these transition form factors. One has to keep in mind that such predictions are affected by higher-twist and end-point ("Feynman") contributions that are formally suppressed by a power of the photon virtuality Q^2. However, at the virtualities covered by present day experiments these corrections are still significant. In figure 16 we show the QCD prediction for the form factors, complemented by taking into account power-suppressed contributions, using dispersion techniques and quark-hadron duality (light-cone sum rules, LCSRs), see [12] for details. The results are compared to the experimental data for the space-like form factors by the CLEO [121] and BaBar [122] collaborations, and we also include BaBar's time-like data points at $q^2 = 112\,\text{GeV}^2$ [123] as stars.

The difference with respect to the original calculation is that in figure 16 the lattice values of the decay constants determined in this work are used as an input, whereas in [12] the decay constants were taken from [96] under the (ad hoc) assumption that they correspond to the scale 1 GeV. Using lattice results removes this scale uncertainty. 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 table 25) down to $\mu_0 = 1.51\,\text{GeV}$ [124] (see section 5.2), 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 figure 16 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.

JHEP08(2021)137

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 [12], the corresponding parameters are chosen from the fit to the pion transition form factor in the same approach [125]. 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 figure 16 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 works 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:

$$\lim_{Q^2\to\infty} Q^2 F_{\gamma\gamma^*\to\eta^{(\prime)}} = \frac{2}{\sqrt{3}}\left(F^8_{\eta^{(\prime)}} + 2\sqrt{2}F^0_{\eta^{(\prime)}}(\mu=\infty, N_f=4)\right). \tag{7.51}$$

We obtain for $N_f = 4$:

$$\begin{aligned}\lim_{Q^2\to\infty} Q^2 F_{\gamma\gamma^*\to\eta}(Q^2) &= 160.5(10.0)\,\text{MeV},\\ \lim_{Q^2\to\infty} Q^2 F_{\gamma\gamma^*\to\eta'}(Q^2) &= 230.5(10.1)\,\text{MeV}.\end{aligned} \tag{7.52}$$

These asymptotic values are shown as dashed lines in figure 16. 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 [19]. Also the matching to the $N_f = 4$ theory somewhat reduces the value. As already emphasized in [12], the effect due to the scale dependence is enhanced for the η' form factor because in this case the two terms in eq. (7.51) 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 [8]. 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.

8 Conclusions

In this study we determined the η and η' masses, their singlet and octet decay constants and gluonic matrix elements without model assumptions in $N_f = 2+1$ QCD as well as the LECs of large-N_c ChPT at NLO. This was achieved by analysing several gauge ensembles, employing non-perturbatively improved Sheikholeslami-Wilson fermions. The twenty-one large volume CLS ensembles that were used are distributed across four lattice spacings $0.050\,\text{fm} \lesssim a \lesssim 0.086\,\text{fm}$, along two distinct quark mass trajectories that both lead down

JHEP08(2021)137

to and include the physical point, which enables a controlled continuum extrapolation of the quark mass dependence of all the observables.

The main results on the masses and decay constants are shown in figure 5 and summarized in sections 7.1–7.2. For the masses we agree with experiment within about one standard deviation and achieve a precision that has considerably improved with respect to previous lattice studies, while fully controlling all systematic errors. Adding all errors in quadrature, we obtain in the continuum limit

$$M_\eta = 554.7(9.2)\,\text{MeV} \quad \text{and} \quad M_{\eta'} = 930(21)\,\text{MeV}. \tag{8.1}$$

Our results for the decay constants are the first to be directly determined from the axialvector matrix elements and they are at a similar level of accuracy in terms of the quoted errors as existing results from the literature that rely on model assumptions and experimental data. The reasonable agreement found with many of these values confirms some of the approximations made and sheds light on their range of validity. In the octet/singlet mixing scheme, defined in eq. (2.12), the η and η' decay constants can be parameterized as follows:

$$\begin{aligned} F^8 &= 115.0(2.8)\,\text{MeV}, & \theta_8 &= -25.8(2.3)^\circ, \\ F^0(\mu = 2\,\text{GeV}) &= 100.1(3.0)\,\text{MeV}, & \theta_0 &= -8.1(1.8)^\circ, \end{aligned} \tag{8.2}$$

where $F_\pi \approx 92\,\text{MeV}$ and the value of the scale dependent singlet decay constant is given in the $\overline{\text{MS}}$ scheme for $N_f = 3$. The corresponding results in different parametrizations and at various renormalization scales are given in table 25 of appendix E. Computing for the first time a value for F^0 at a definite QCD scale, enabled us to improve on the prediction [12] of the transition form factors $\gamma\gamma^* \to \eta^{(\prime)}$, which is presented in section 7.5.

The continuum extrapolation and physical mass point interpolation benefit from the large parameter space explored. We carry out simultaneous NLO large-N_c ChPT fits to the two masses and four decay constants, while including all possible correlations among these. A number of different fits are performed in order to quantify the systematic errors and the parametrization employed yields a consistent set of low energy constants (LECs), see section 7.3. By taking the renormalization group running of the singlet axialvector current into account, we can for the first time determine the OZI-rule violating LECs Λ_1 and Λ_2 at well-defined scales and find that Λ_1 is small only around $\mu = 1\,\text{GeV}$, while Λ_2 cannot be neglected at any scale, see figure 14. The NLO large-N_c ChPT LECs read:

$$\begin{aligned} M_0(\mu = 2\,\text{GeV}) &= 818(27)\,\text{MeV}, & F &= 87.7(2.8)\,\text{MeV}, \\ \Lambda_1(\mu = 2\,\text{GeV}) &= -0.13(5), & L_5 &= 1.66(23)\cdot 10^{-3}, \\ \Lambda_2(\mu = 2\,\text{GeV}) &= 0.19(10), & L_8 &= 1.08(13)\cdot 10^{-3}. \end{aligned} \tag{8.3}$$

Note that M_0, Λ_1 and Λ_2 depend on the QCD scale. Meson loops do not contribute at NLO and, therefore, the above LECs are independent of the ChPT renormalization scale.

Using the axialvector and pseudoscalar matrix elements of the η and η' mesons, we were able to test the octet and singlet AWIs and to determine the gluonic matrix elements

JHEP08(2021)137

$a_{\eta^{(\prime)}} = \langle \Omega | 2\omega | \eta^{(\prime)} \rangle$. The results are discussed in section 7.4. We found consistency with the prediction from large-N_c ChPT (derived in appendix B), see figure 8, and we successfully checked our computation against the same matrix elements determined directly from the gluonic correlation functions, after carrying out the appropriate renormalization. As a by-product we also confirmed for the first time in $N_f = 2 + 1$ QCD that the topological susceptibility, while significantly affected by lattice corrections, is well described by the leading order ChPT expectation with only one LEC, F, see figure 9. Our first ab-initio calculation of the anomaly matrix elements gives at the physical point in the continuum limit, in the $\overline{\text{MS}}$ scheme for $N_f = 3$,

$$a_\eta(\mu = 2\,\text{GeV}) = 0.0170(10)\,\text{GeV}^3 \quad \text{and} \quad a_{\eta'}(\mu = 2\,\text{GeV}) = 0.0381(84)\,\text{GeV}^3, \tag{8.4}$$

where the mixing angle

$$\theta_y(\mu = 2\,\text{GeV}) = -\arctan\left(\frac{a_\eta}{a_{\eta'}}\right) = -24.0(3.3)^\circ \tag{8.5}$$

at this scale is close to θ_8 as expected in the FKS state mixing model [96, 110]. While θ_8 is scale independent, the value of $-\theta_y$ increases with the renormalization scale, see table 19. Using the above result, we find excellent agreement with the ratio of the experimental decay rates for $J/\psi \to \eta'\gamma$ and $J/\psi \to \eta\gamma$.

In general, we find NLO large-N_c U(3) ChPT to describe our data reasonably well, however, there is some tension regarding the LECs between the gluonic and fermionic matrix elements, in particular regarding Λ_1. In view of this, a NNLO description may be desirable, also with respect to a matching to SU(3) ChPT, where meson loops already contribute at NLO. Simply adding the meson loop contributions that enter at NNLO in large-N_c ChPT to the NLO parametrization gives a less satisfactory description of the data. Therefore, ideally, one would carry out a full NNLO analysis. Constraining the additional LECs in this case will require data on additional ensembles, in particular along the $m_s = m_\ell$ line in the quark mass plane, and a simultaneous analysis of the masses and decay constants of the whole nonet of light mesons.

Many phenomenological descriptions of experimental data give numbers for the matrix elements that agree with or are close to those of our QCD calculation. For the first time, we presented results at an unambiguous QCD renormalization scale with a reliable quantification of the systematic uncertainties. The values of the decay constants and anomaly matrix elements calculated here, therefore, constitute very valuable input to theory predictions that are related to upcoming experiments, e.g., at Belle II [8].

Acknowledgments

The authors thank T. Feldmann, K. Ottnad and S. Scherer for discussions and clarifications regarding their articles. We thank F. Joswig and I. Kanamori for alerting us of misprints in an earlier draft. J. S. thanks W. Söldner for discussions. 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

JHEP08(2021)137

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 thank our colleagues in CLS [http://wiki-zeuthen.desy.de/CLS/CLS] for the joint effort in the generation of the gauge field ensembles. The authors gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time through the John von Neumann Institute for Computing (NIC) on the super-computers JUQUEEN [126], JUWELS [127] and in particular JURECA-Booster [128] 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 [129] and QPACE 3 machines. We use the multigrid solver of [59, 130–132] for the inversion of the Dirac operator.

A 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 section 5.3.2 do not contain chiral logarithms or a dependence on the EFT renormalization scale μ_{EFT}. We define the loop functions

$$A_0(M^2) = -M^2 \log\left(\frac{M^2}{\mu^2_{\text{EFT}}}\right). \tag{A.1}$$

The octet and singlet decay constants of eqs. (5.43)–(5.46) receive the additional contributions [42, 81]

$$F_\eta^{8\,\text{NLO+loops}} = F_\eta^{8\,\text{NLO}} + \frac{3}{32\pi^2 F}\cos(\theta) A_0(M_K^2), \tag{A.2}$$

$$F_{\eta'}^{8\,\text{NLO+loops}} = F_{\eta'}^{8\,\text{NLO}} + \frac{3}{32\pi^2 F}\sin(\theta) A_0(M_K^2), \tag{A.3}$$

$$F_\eta^{0\,\text{NLO+loops}} = F_\eta^{0\,\text{NLO}} - \frac{1}{32\pi^2 F}\sin(\theta) A_0(M_\pi^2), \tag{A.4}$$

$$F_{\eta'}^{0\,\text{NLO+loops}} = F_{\eta'}^{0\,\text{NLO}} + \frac{1}{32\pi^2 F}\cos(\theta) A_0(M_\pi^2). \tag{A.5}$$

Moreover, the quark mass dependence of the mass mixing angle θ, defined in eq. (5.25), changes as the entries of the square mass matrix eq. (5.19) also receive additional contri-

JHEP08(2021)137

butions. Specifically, we have to add to eqs. (5.40)–(5.42) [42]:[13]

$$
\begin{aligned}
(\mu_8^{\text{NLO+loops}})^2 = (\mu_8^{\text{NLO}})^2 + \frac{1}{48\pi^2 F^2}\Bigg[& \left(\frac{3}{2}\overline{M}^2 - \frac{1}{2}\delta M^2\right) A_0(M_\pi^2) \\
& - \left(4\overline{M}^2 + \frac{2}{3}\delta M^2\right) A_0(M_K^2) \\
& + \left(\frac{5}{4}\overline{M}^2 + \frac{7}{12}\delta M^2\right)\left(A_0(\tilde{M}_\eta^2) + A_0(\tilde{M}_{\eta'}^2)\right) \\
& + \frac{2\sqrt{2}\sin(2\theta^{\text{LO}}) + \cos(2\theta^{\text{LO}})}{4}\left(\overline{M}^2 + \delta M^2\right)\left(A_0(\tilde{M}_\eta^2) - A_0(\tilde{M}_{\eta'}^2)\right)\Bigg], \qquad \text{(A.6)}
\end{aligned}
$$

$$
\begin{aligned}
(\mu_0^{\text{NLO+loops}})^2 = (\mu_0^{\text{NLO}})^2 + \frac{1}{48\pi^2 F^2}\Bigg[& \left(3\overline{M}^2 - \delta M^2\right) A_0(M_\pi^2) \\
& + \left(4\overline{M}^2 + \frac{2}{3}\delta M^2\right) A_0(M_K^2) \\
& + \left(\overline{M}^2 + \frac{1}{6}\delta M^2\right)\left(A_0(\tilde{M}_\eta^2) + A_0(\tilde{M}_{\eta'}^2)\right) \\
& + \frac{2\sqrt{2}\sin(2\theta^{\text{LO}}) + \cos(2\theta^{\text{LO}})}{6}\delta M^2\left(A_0(\tilde{M}_\eta^2) - A_0(\tilde{M}_{\eta'}^2)\right)\Bigg], \qquad \text{(A.7)}
\end{aligned}
$$

$$
\begin{aligned}
(\mu_{80}^{\text{NLO+loops}})^2 = (\mu_{80}^{\text{NLO}})^2 + \frac{\sqrt{2}}{48\pi^2 F^2}\Bigg[& \left(\frac{3}{2}\overline{M}^2 - \frac{1}{2}\delta M^2\right) A_0(M_\pi^2) \\
& - \left(\overline{M}^2 + \frac{1}{6}\delta M^2\right) A_0(M_K^2) \\
& - \left(\frac{1}{4}\overline{M}^2 + \frac{5}{12}\delta M^2\right)\left(A_0(\tilde{M}_\eta^2) + A_0(\tilde{M}_{\eta'}^2)\right) \\
& - \frac{2\sqrt{2}\sin(2\theta^{\text{LO}}) + \cos(2\theta^{\text{LO}})}{4}\left(\overline{M}^2 + \frac{1}{3}\delta M^2\right)\left(A_0(\tilde{M}_\eta^2) - A_0(\tilde{M}_{\eta'}^2)\right)\Bigg]. \qquad \text{(A.8)}
\end{aligned}
$$

θ^{LO} corresponds to the mass mixing angle eq. (5.25), evaluated at LO, eqs. (5.36)–(5.38). $\tilde{M}_\eta$ and $\tilde{M}_{\eta'}$ denote the η and η' masses, computed at LO via eqs. (5.26)–(5.28) and (5.36)–(5.38).

Carrying out the analysis of our masses and decay constants, including the NNLO loops, we obtain for the LECs at $\mu_{\text{EFT}} = 0.770\,\text{GeV}$ in the $N_f = 3$ $\overline{\text{MS}}$ scheme at $\mu = \infty$:

$$
\begin{aligned}
L_5 &= 1.97\,\binom{16}{11}_{\text{stat}}\binom{0}{19}_a\binom{0}{23}_\chi \times 10^{-3}, \\
L_8 &= 0.848\,\binom{126}{109}_{\text{stat}}\binom{0}{124}_a\binom{0}{113}_\chi \times 10^{-3}, \\
M_0 &= 1.78\,\binom{3}{5}_{\text{stat}}\binom{2}{2}_a\binom{0}{0}_\chi\binom{1}{1}_{\text{renorm}}(8t_0^\chi)^{-1/2} \\
&= 837\,\binom{13}{23}_{\text{stat}}\binom{11}{11}_{\text{syst}}(12)_{t_0}\,\text{MeV}, \\
F &= 0.1680\,\binom{38}{66}_{\text{stat}}\binom{69}{0}_a\binom{39}{0}_\chi\binom{4}{1}_{\text{renorm}}(8t_0^\chi)^{-1/2} \\
&= 78.97\,\binom{1.78}{3.10}_{\text{stat}}\binom{3.71}{17}_{\text{syst}}(1.18)_{t_0}\,\text{MeV},
\end{aligned}
$$

[13] Note that there are misprints within the normalizations of eqs. (C11)–(C13) in [42].

JHEP08(2021)137

$$\begin{aligned}
\Lambda_1 &= -0.10\,\binom{2}{4}_{\text{stat}}\,\binom{5}{4}_a\,\binom{0}{2}_\chi\,\binom{4}{2}_{\text{renorm}}\,,\\
\tilde{\Lambda} &= -1.0\,\binom{2}{3}_{\text{stat}}\,\binom{4}{0}_a\,\binom{2}{0}_\chi\,\binom{0}{1}_{\text{renorm}}\,,\\
\Lambda_2 &= 0.45\,\binom{16}{12}_{\text{stat}}\,\binom{1}{19}_a\,\binom{0}{12}_\chi\,\binom{6}{3}_{\text{renorm}}
\end{aligned} \tag{A.9}$$

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.

B NLO expressions for pseudoscalar and gluonic matrix elements

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. (6.9) 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

$$F_n^8 M_n^2 = \frac{2}{\sqrt{3}} m_\ell H_n^\ell - 2\sqrt{\frac{2}{3}} m_s H_n^s, \tag{B.1}$$

$$F_n^0 M_n^2 = 2\sqrt{\frac{2}{3}} m_\ell H_n^\ell + \frac{2}{\sqrt{3}} m_s H_n^s + \sqrt{\frac{2}{3}} a_n, \tag{B.2}$$

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. (B.1) and (B.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. (5.43)–(5.46) while M_η^2 and $M_{\eta'}^2$ can be parameterized in terms of the squared mass matrix (with elements eqs. (5.40)–(5.42)) via the rotation (5.20)–(5.21) with the angle defined in eq. (5.25) 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

$$2m_\ell B_0 = M_\pi^2 \left(1 - 8\frac{2L_8 - L_5}{F^2} M_\pi^2\right), \tag{B.3}$$

$$(m_\ell + m_s)\, B_0 = M_K^2 \left(1 - 8\frac{2L_8 - L_5}{F^2} M_K^2\right). \tag{B.4}$$

JHEP08(2021)137

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 (B.1)–(B.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

$$
\begin{aligned}
m_\ell H_\eta^\ell = \cos\theta F \Bigg\{ & \left(\overline{M}^2 - \frac{\delta M^2}{3}\right)\frac{\sqrt{3}}{12}\left[(1-2\Lambda_1+4\Lambda_2)\left(1-\frac{1}{\cos(2\theta)}\right)+2\right] \\
& + M_0^2\left(1-\frac{3\overline{M}^2}{\delta M^2}\right)\frac{\sqrt{3}}{12}(1-\Lambda_1)\left(1-\frac{1}{\cos(2\theta)}\right) \\
& + \frac{\overline{M}^2\delta M^2}{3}\left[\frac{5\sqrt{3}}{3}\left(2\frac{2L_8-L_5}{F^2}+\frac{L_5}{F^2}\right)\left(1+\frac{1}{5\cos(2\theta)}\right)\right. \\
& \qquad \left. + \sqrt{6}\frac{L_5}{F^2}\tan\theta\left(1-\frac{1}{3\cos(2\theta)}\right)\right] \\
& + \overline{M}^2 M_0^2\left[\frac{2L_8-L_5}{F^2}\left(\frac{5\sqrt{3}}{3}+2\sqrt{3}\frac{\overline{M}^2}{\delta M^2}\right)\right. \\
& \qquad \left. - \left(\frac{\sqrt{6}}{3}\tan\theta+\sqrt{3}\right)\frac{L_5}{F^2}\right]\left(1-\frac{1}{\cos(2\theta)}\right) \\
& - \frac{\delta M^2 M_0^2}{3}\left[\frac{7\sqrt{3}}{3}\frac{2L_8-L_5}{F^2} - \left(\sqrt{3}+\frac{\sqrt{6}}{3}\tan\theta\right)\frac{L_5}{F^2}\right]\left(1-\frac{1}{\cos(2\theta)}\right) \\
& - 2\sqrt{3}\overline{M}^4\frac{2L_8-L_5}{F^2}\left(1-\frac{1}{3\cos(2\theta)}\right) \\
& - \delta M^4\left[\frac{4\sqrt{3}}{27}\frac{2L_8-L_5}{F^2}\left(1+\frac{1}{\cos(2\theta)}\right)+\frac{\sqrt{3}}{27}\frac{L_5}{F^2}\left(5+\frac{1}{\cos(2\theta)}\right)\right. \\
& \qquad \left. + \frac{\sqrt{6}}{9}\tan\theta\frac{L_5}{F^2}\left(1-\frac{1}{3\cos(2\theta)}\right)\right]\Bigg\},
\end{aligned}
\tag{B.5}
$$

JHEP08(2021)137

$$
\begin{aligned}
m_s H_\eta^s = \cos\theta F \Bigg\{ & -\frac{\sqrt{6}}{8}\left(\overline{M}^2+\frac{2}{3}\delta M^2\right)\left[\left(1-\frac{\Lambda_1}{3}+\frac{2\Lambda_2}{3}\right)\left(1-\frac{1}{\cos(2\theta)}\right)+\frac{4}{3\cos(2\theta)}\right] \\
& -\frac{\sqrt{6}}{12}M_0^2(1-\Lambda_1)\left(1+\frac{3\overline{M}^2}{2\delta M^2}\right)\left(1-\frac{1}{\cos(2\theta)}\right) \\
& +\frac{\overline{M}^2\delta M^2}{3}\left[\frac{5\sqrt{6}}{3}\frac{2L_8-L_5}{F^2}\left(1+\frac{1}{5\cos(\theta)}\right)\right. \\
& \qquad -\frac{L_5}{F^2}\left(\frac{2\sqrt{6}}{3}\left(1+\frac{1}{2\cos(2\theta)}\right)\right. \\
& \qquad\qquad \left.\left.+\sqrt{3}\tan\theta\left(1+\frac{1}{3\cos(2\theta)}\right)\right)\right]+ \\
& +\overline{M}^2M_0^2\left[\left(\frac{5\sqrt{6}}{6}+\sqrt{6}\frac{\overline{M}^2}{\delta M^2}\right)\frac{2L_8-L_5}{F^2}-\frac{\sqrt{3}}{3}\frac{L_5}{F^2}\tan\theta\right]\left(1-\frac{1}{\cos(2\theta)}\right) \\
& -\delta M^2M_0^2\left(\frac{7\sqrt{6}}{18}\frac{2L_8-L_5}{F^2}+\frac{2\sqrt{3}}{9}\frac{L_5}{F^2}\tan\theta\right)\left(1-\frac{1}{\cos(2\theta)}\right) \\
& +\sqrt{6}\overline{M}^4\frac{2L_8-L_5}{F^2}\left(1+\frac{1}{3\cos(2\theta)}\right) \\
& -\frac{\delta M^4}{9}\left[\frac{\sqrt{6}}{6}\frac{2L_8-L_5}{F^2}\left(31+\frac{13}{\cos(2\theta)}\right)+\frac{2\sqrt{6}}{3}\frac{L_5}{F^2}\left(2+\frac{1}{\cos(2\theta)}\right)\right. \\
& \qquad \left.+2\sqrt{3}\tan\theta\frac{L_5}{F^2}\left(1+\frac{1}{3\cos(2\theta)}\right)\right]\Bigg\},
\end{aligned}
\tag{B.6}
$$

$$
\begin{aligned}
m_\ell H_{\eta'}^\ell = \sin\theta F \Bigg\{ & \left(\overline{M}^2-\frac{\delta M^2}{3}\right)\frac{\sqrt{3}}{12}\left[(1-2\Lambda_1+4\Lambda_2)\left(1+\frac{1}{\cos(2\theta)}\right)+2\right] \\
& +M_0^2\left(1-\frac{3\overline{M}^2}{\delta M^2}\right)\frac{\sqrt{3}}{12}(1-\Lambda_1)\left(1+\frac{1}{\cos(2\theta)}\right) \\
& +\frac{\overline{M}^2\delta M^2}{3}\left[\frac{5\sqrt{3}}{3}\left(2\frac{2L_8-L_5}{F^2}+\frac{L_5}{F^2}\right)\left(1-\frac{1}{5\cos(2\theta)}\right)\right. \\
& \qquad \left.-\sqrt{6}\frac{L_5}{F^2}\cot\theta\left(1+\frac{1}{3\cos(2\theta)}\right)\right] \\
& +\overline{M}^2M_0^2\left[\frac{2L_8-L_5}{F^2}\left(\frac{5\sqrt{3}}{3}+2\sqrt{3}\frac{\overline{M}^2}{\delta M^2}\right)\right. \\
& \qquad \left.+\left(\frac{\sqrt{6}}{3}\cot\theta-\sqrt{3}\right)\frac{L_5}{F^2}\right]\left(1+\frac{1}{\cos(2\theta)}\right) \\
& -\frac{\delta M^2M_0^2}{3}\left[\frac{7\sqrt{3}}{3}\frac{2L_8-L_5}{F^2}+\left(-\sqrt{3}+\frac{\sqrt{6}}{3}\cot\theta\right)\frac{L_5}{F^2}\right]\left(1+\frac{1}{\cos(2\theta)}\right) \\
& -2\sqrt{3}\overline{M}^4\frac{2L_8-L_5}{F^2}\left(1+\frac{1}{3\cos(2\theta)}\right) \\
& -\frac{\sqrt{3}}{27}\delta M^4\left[4\frac{2L_8-L_5}{F^2}\left(1-\frac{1}{\cos(2\theta)}\right)+\left(5-\frac{1}{\cos(2\theta)}\right)\frac{L_5}{F^2}\right. \\
& \qquad \left.-3\sqrt{2}\cot\theta\frac{L_5}{F^2}\left(1+\frac{1}{3\cos(2\theta)}\right)\right]\Bigg\},
\end{aligned}
\tag{B.7}
$$

JHEP08(2021)137

$$
\begin{aligned}
m_s H^s_{\eta'} = \sin\theta F \Bigg\{ & -\frac{\sqrt{6}}{8}\left(\overline{M}^2+\frac{2}{3}\delta M^2\right)\left[\left(1-\frac{\Lambda_1}{3}+\frac{2\Lambda_2}{3}\right)\left(1+\frac{1}{\cos(2\theta)}\right)-\frac{4}{3\cos(2\theta)}\right] \\
& -\frac{\sqrt{6}}{12}M_0^2(1-\Lambda_1)\left(1+\frac{3\overline{M}^2}{2\delta M^2}\right)\left(1+\frac{1}{\cos(2\theta)}\right) \\
& +\frac{\overline{M}^2\delta M^2}{3}\left[\frac{5\sqrt{6}}{3}\frac{2L_8-L_5}{F^2}\left(1-\frac{1}{5\cos(\theta)}\right)\right. \\
& \qquad \left. -\frac{L_5}{F^2}\left(\frac{2\sqrt{6}}{3}\left(1-\frac{1}{2\cos(2\theta)}\right)-\sqrt{3}\cot\theta\left(1-\frac{1}{3\cos(2\theta)}\right)\right)\right] \\
& +\overline{M}^2M_0^2\left[\left(\frac{5\sqrt{6}}{6}+\sqrt{6}\frac{\overline{M}^2}{\delta M^2}\right)\frac{2L_8-L_5}{F^2}+\frac{\sqrt{3}}{3}\frac{L_5}{F^2}\cot\theta\right]\left(1+\frac{1}{\cos(2\theta)}\right) \\
& -\delta M^2M_0^2\left(\frac{7\sqrt{6}}{18}\frac{2L_8-L_5}{F^2}-\frac{2\sqrt{3}}{9}\frac{L_5}{F^2}\cot\theta\right)\left(1+\frac{1}{\cos(2\theta)}\right) \\
& +\sqrt{6}\,\overline{M}^4\frac{2L_8-L_5}{F^2}\left(1-\frac{1}{3\cos(2\theta)}\right) \\
& -\frac{\delta M^4}{9}\left[\frac{\sqrt{6}}{6}\frac{2L_8-L_5}{F^2}\left(31-\frac{13}{\cos(2\theta)}\right)+\frac{4\sqrt{6}}{6}\frac{L_5}{F^2}\left(2-\frac{1}{\cos(2\theta)}\right)\right. \\
& \qquad \left. -2\sqrt{3}\frac{L_5}{F^2}\cot\theta\left(1-\frac{1}{3\cos(2\theta)}\right)\right]\Bigg\}. \qquad \text{(B.8)}
\end{aligned}
$$

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:

$$
\begin{aligned}
a_\eta = \cos\theta F \Bigg\{ & \frac{\sqrt{3}}{6}\overline{M}^2\left(-1+\Lambda_1-2\Lambda_2+3\frac{M_0^2}{\delta M^2}(1-\Lambda_1)\right)\left(1-\frac{1}{\cos(2\theta)}\right) \qquad \text{(B.9)} \\
& +\frac{\sqrt{3}}{9}\delta M^2\left[(-\Lambda_1+2\Lambda_2)\left(1-\frac{1}{\cos(2\theta)}\right)+2\right] \\
& -\frac{4\sqrt{3}}{9}\overline{M}^2\delta M^2\left[\frac{2L_8-L_5}{F^2}\left(5+\frac{1}{\cos(2\theta)}\right)+3\frac{L_5}{F^2}\right] \\
& -\frac{2\sqrt{3}}{3}\overline{M}^2M_0^2\left(5\frac{2L_8-L_5}{F^2}-2\frac{L_5}{F^2}\right)\left(1-\frac{1}{\cos(2\theta)}\right) \\
& +\frac{2\sqrt{3}}{9}\delta M^2M_0^2\left(7\frac{2L_8-L_5}{F^2}-4\frac{L_5}{F^2}\right)\left(1-\frac{1}{\cos(2\theta)}\right) \\
& +\frac{4\sqrt{3}}{3}\overline{M}^4\left(1-3\frac{M_0^2}{\delta M^2}\right)\frac{2L_8-L_5}{F^2}\left(1-\frac{1}{\cos(2\theta)}\right) \\
& +\frac{\sqrt{3}}{27}\delta M^4\left[\frac{2L_8-L_5}{F^2}\left(26+\frac{14}{\cos(2\theta)}\right)+8\frac{L_5}{F^2}\right]\Bigg\}+
\end{aligned}
$$

JHEP08(2021)137

$$
\begin{aligned}
&+\sin\theta F\Bigg\{-\frac{\sqrt{6}}{3}\overline{M}^2\left[1+\frac{\Lambda_1}{2\cos(2\theta)}+\Lambda_2\left(1-\frac{1}{\cos(2\theta)}\right)\right]\\
&\quad-\frac{\sqrt{6}}{18}\delta M^2\left(1+\frac{\Lambda_1}{2}\right)\left(1+\frac{1}{\cos(2\theta)}\right)\\
&\quad-\frac{\sqrt{6}}{6}M_0^2\left(1-\frac{\Lambda_1}{2}\right)\left(1-\frac{1}{\cos(2\theta)}\right)\\
&\quad-\frac{4\sqrt{6}}{9}\frac{L_5}{F^2}\overline{M}^2\delta M^2+\frac{4\sqrt{6}}{3}\overline{M}^2M_0^2\frac{L_5}{F^2}\left(1-\frac{1}{\cos(2\theta)}\right)-\frac{4\sqrt{6}}{3}\frac{L_5}{F^2}\overline{M}^4\\
&\quad-\frac{2\sqrt{6}}{3}\delta M^4\left[\left(1+\frac{1}{3\cos(2\theta)}\right)\frac{2L_8-L_5}{F^2}-\frac{4}{9}\frac{L_5}{F^2}\right]\Bigg\},
\end{aligned}
$$

$$
\begin{aligned}
a_{\eta'}=\cos\theta F\Bigg\{&\frac{\sqrt{6}}{3}\overline{M}^2\left[1-\frac{\Lambda_1}{2\cos(2\theta)}+\Lambda_2\left(1+\frac{1}{\cos(2\theta)}\right)\right]\\
&+\frac{\sqrt{6}}{18}\delta M^2\left(1+\frac{\Lambda_1}{2}\right)\left(1-\frac{1}{\cos(2\theta)}\right)\\
&+\frac{\sqrt{6}}{6}M_0^2\left(1-\frac{\Lambda_1}{2}\right)\left(1+\frac{1}{\cos(2\theta)}\right)\\
&+\frac{4\sqrt{6}}{9}\overline{M}^2\delta M^2\frac{L_5}{F^2}-\frac{4\sqrt{6}}{3}\overline{M}^2M_0^2\frac{L_5}{F^2}\left(1+\frac{1}{\cos(2\theta)}\right)\\
&+\frac{4\sqrt{6}}{3}\overline{M}^4\frac{L_5}{F^2}\\
&+\frac{2\sqrt{6}}{3}\delta M^4\left[\frac{2L_8-L_5}{F^2}\left(1-\frac{1}{3\cos(2\theta)}\right)-\frac{4}{9}\frac{L_5}{F^2}\right]\Bigg\}
\end{aligned}
\tag{B.10}
$$

$$
\begin{aligned}
&+\sin\theta F\Bigg\{-\frac{\sqrt{3}}{6}\overline{M}^2\left(1-\Lambda_1+2\Lambda_2-3\frac{M_{0^2}}{\delta M^2}(1-\Lambda_1)\right)\left(1+\frac{1}{\cos(2\theta)}\right)\\
&\quad+\frac{2\sqrt{3}}{9}\delta M^2\left[1+\left(-\frac{\Lambda_1}{2}+\Lambda_2\right)\left(1+\frac{1}{\cos(2\theta)}\right)\right]\\
&\quad-\frac{4\sqrt{3}}{9}\overline{M}^2\delta M^2\left[\frac{2L_8-L_5}{F^2}\left(5-\frac{1}{\cos(2\theta)}\right)+3\frac{L_5}{F^2}\right]\\
&\quad-\frac{2\sqrt{3}}{3}\overline{M}^2M_0^2\left(5\frac{2L_8-L_5}{F^2}-2\frac{L_5}{F^2}\right)\left(1+\frac{1}{\cos(2\theta)}\right)\\
&\quad+\frac{2\sqrt{3}}{9}\delta M^2M_0^2\left(7\frac{2L_8-L_5}{F^2}-4\frac{L_5}{F^2}\right)\left(1+\frac{1}{\cos(2\theta)}\right)\\
&\quad+\frac{4\sqrt{3}}{3}\overline{M}^4\frac{2L_8-L_5}{F^2}\left(1-3\frac{M_0^2}{\delta M^2}\right)\left(1+\frac{1}{\cos(2\theta)}\right)\\
&\quad+\frac{2\sqrt{3}}{27}\delta M^4\left[\frac{2L_8-L_5}{F^2}\left(13-\frac{7}{\cos(2\theta)}\right)+4\frac{L_5}{F^2}\right]\Bigg\}.
\end{aligned}
\tag{B.11}
$$

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. (5.25)

JHEP08(2021)137

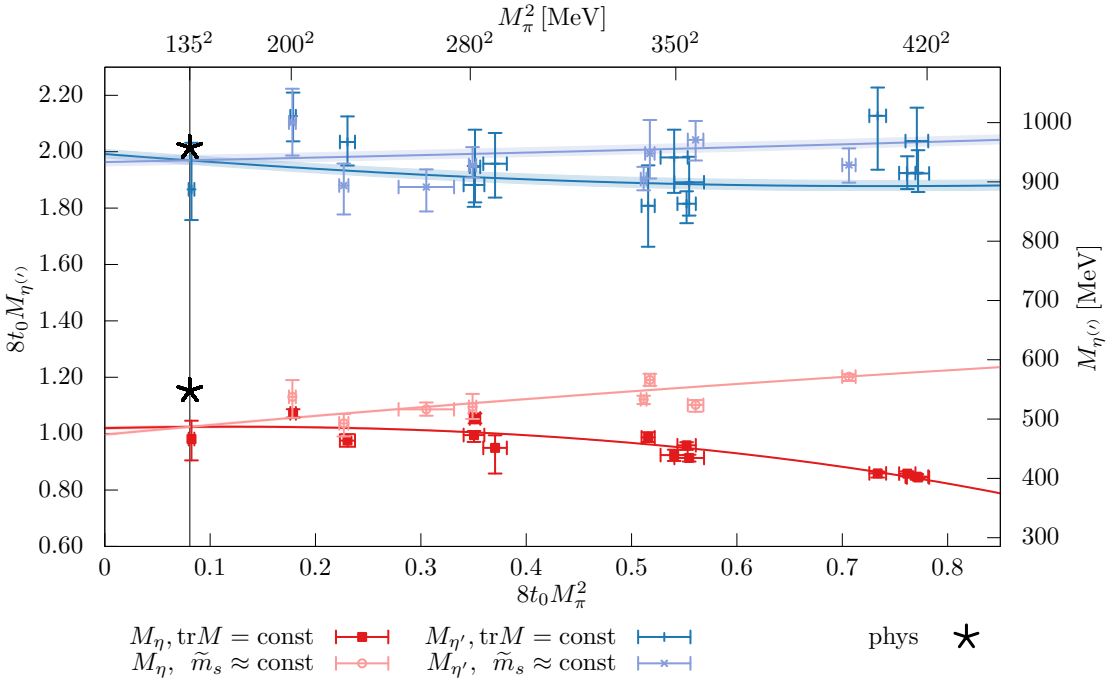


Figure 17. LO parametrization of the masses of the η and the η' mesons for our two trajectories in the quark mass plane. The data are corrected for lattice spacing effects according to the fit.

and (5.36)–(5.38) as a function of the δM^2 and M_0^2. Above, we use the NLO expression for θ, eqs. (5.25) and (5.40)–(5.42).

C The leading order fit

We show in figure 17 the analogue of figure 5 for our fit to the LO parametrization eqs. (5.26)–(5.28) and (5.36)–(5.38). No simultaneous fit of the masses and the decay constants can be carried out since $F_\eta^8 \neq F_\eta^0$ and $F_\eta^0 \neq -F_{\eta'}^8$, which is why only the masses are included. The continuum parametrization shown depend on a single parameter, the LEC $M_0 \approx 785\,\text{MeV}$, as detailed in section 5.6. Although the raw lattice data fall onto continuous curves (see the upper panel of figure 5), lattice correction terms $\propto a^2\delta M^2$ had to be added for each particle to obtain $\chi^2/N_\text{df} \approx 91/41$. In figure 17 the shifted data are shown, along with the continuum limit curves that depend only on the parameter M_0. Since in our NLO fits 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.

D Continuum limit fit parameters

The parametrizations of lattice artefacts within our simultaneous fits to the masses and decay constants are defined in section 5.4. In section 5.5 we explain how 17 different

JHEP08(2021)137

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 21. Fit results for the unknown $\mathcal{O}(a)$ improvement coefficients, see eqs. (5.50), (5.51) and (5.52). The fit ids are defined in table 7.

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 22. Fit parameters, accompanying quadratic lattice effects for the octet decay constants, see eq. (5.53). The fit ids are defined in table 7.

JHEP08(2021)137

id	$\chi^2/N_{\rm df}$	$l_{F^0_\eta}$	$m_{F^0_\eta}$	$n_{F^0_\eta}$	$l_{F^0_{\eta'}}$	$m_{F^0_{\eta'}}$	$n_{F^0_{\eta'}}$
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 23. Fit parameters, accompanying quadratic lattice effects for the singlet decay constants, see eq. (5.53). The fit ids are defined in table 7.

parametrizations were selected. These are enumerated and defined in table 7. The six continuum limit fit parameters (LECs) for each of these fits are given in table 8. Here, in table 21 we list the results for the unknown $\mathcal{O}(a)$ improvement coefficients within eqs. (5.50) and (5.51). Their parametrizations are given in eq. (5.52). In tables 22 and 23 we list the $\mathcal{O}(a^2)$ coefficients, defined in eq. (5.53) for both octet and both singlet decay constants, respectively.

E Decay constants in various representations

We list the four decay constants, in units of $(8t_0^{\rm ph})^{-1/2}$ and in MeV. In table 24 we collect the results of our simultaneous fits to the masses and decay constants, including the statistical and systematic errors, while in table 25 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. (2.11). In addition, we give the parameters of the corresponding two-angle representations eqs. (2.12) and (2.13). 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 $\overline{\rm MS}$ scheme for $N_f = 3$ active flavours. Only the octet decay constants F^8_η, $F^8_{\eta'}$ 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 $\tan(\theta_0) = -F^0_\eta/F^0_{\eta'}$.

JHEP08(2021)137

Table 24. Decay constants in various representations and at several renormalization scales.

octet/singlet basis, state representation			
F_η^8		$0.2219\,\binom{18}{37}_{\text{stat}}\,\binom{17}{24}_{a}\,\binom{10}{2}_{\chi}\,(8t_0^{\text{ph}})^{-1/2}$	$105.4\,\binom{9}{1.8}_{\text{stat}}\,\binom{9}{1.1}_{\text{syst}}\,(1.3)_{t_0}$ MeV
$F_{\eta'}^8$		$-0.0939\,\binom{28}{100}_{\text{stat}}\,\binom{84}{0}_{a}\,\binom{58}{82}_{\chi}\,(8t_0^{\text{ph}})^{-1/2}$	$-44.6\,\binom{1.3}{4.8}_{\text{stat}}\,\binom{4.9}{3.9}_{\text{syst}}\,(6)_{t_0}$ MeV
F_η^0	$\mu =$ ∞	$0.0224\,\binom{53}{30}_{\text{stat}}\,\binom{28}{0}_{a}\,\binom{5}{21}_{\chi}\,\binom{20}{8}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$10.6\,\binom{2.5}{1.4}_{\text{stat}}\,\binom{1.4}{1.4}_{\text{syst}}\,(1)_{t_0}$ MeV
	1 GeV	$0.0255\,\binom{60}{35}_{\text{stat}}\,\binom{32}{0}_{a}\,\binom{5}{24}_{\chi}\,\binom{23}{9}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$12.1\,\binom{2.9}{1.6}_{\text{stat}}\,\binom{1.6}{1.6}_{\text{syst}}\,(2)_{t_0}$ MeV
	2 GeV	$0.0241\,\binom{57}{33}_{\text{stat}}\,\binom{30}{0}_{a}\,\binom{5}{23}_{\chi}\,\binom{22}{9}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$11.4\,\binom{2.7}{1.5}_{\text{stat}}\,\binom{1.5}{1.5}_{\text{syst}}\,(1)_{t_0}$ MeV
	10 GeV	$0.0233\,\binom{55}{32}_{\text{stat}}\,\binom{29}{0}_{a}\,\binom{5}{22}_{\chi}\,\binom{21}{8}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$11.1\,\binom{2.6}{1.5}_{\text{stat}}\,\binom{1.5}{1.5}_{\text{syst}}\,(1)_{t_0}$ MeV
$F_{\eta'}^0$	$\mu =$ ∞	$0.1974\,\binom{14}{48}_{\text{stat}}\,\binom{0}{31}_{a}\,\binom{4}{27}_{\chi}\,\binom{52}{26}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$93.77\,\binom{67}{2.29}_{\text{stat}}\,\binom{1.24}{3.18}_{\text{syst}}\,(1.18)_{t_0}$ MeV
	1 GeV	$0.2247\,\binom{16}{55}_{\text{stat}}\,\binom{0}{36}_{a}\,\binom{5}{31}_{\chi}\,\binom{60}{29}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$106.7\,\binom{8}{2.6}_{\text{stat}}\,\binom{1.4}{3.6}_{\text{syst}}\,(1.3)_{t_0}$ MeV
	2 GeV	$0.2122\,\binom{15}{52}_{\text{stat}}\,\binom{0}{34}_{a}\,\binom{5}{29}_{\chi}\,\binom{56}{28}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$100.8\,\binom{7}{2.5}_{\text{stat}}\,\binom{1.3}{3.4}_{\text{syst}}\,(1.3)_{t_0}$ MeV
	10 GeV	$0.2051\,\binom{15}{50}_{\text{stat}}\,\binom{0}{33}_{a}\,\binom{5}{28}_{\chi}\,\binom{54}{27}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$97.41\,\binom{69}{2.38}_{\text{stat}}\,\binom{1.29}{3.30}_{\text{syst}}\,(1.23)_{t_0}$ MeV
octet/singlet basis, angle representation			
F^8		$0.2410\,\binom{23}{16}_{\text{stat}}\,\binom{11}{50}_{a}\,\binom{38}{12}_{\chi}\,(8t_0^{\text{ph}})^{-1/2}$	$114.5\,\binom{1.1}{8}_{\text{stat}}\,\binom{1.9}{2.5}_{\text{syst}}\,(1.4)_{t_0}$ 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	$\mu =$ ∞	$0.1987\,\binom{12}{42}_{\text{stat}}\,\binom{0}{31}_{a}\,\binom{2}{29}_{\chi}\,\binom{54}{27}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$94.37\,\binom{57}{2.01}_{\text{stat}}\,\binom{1.26}{3.26}_{\text{syst}}\,(1.19)_{t_0}$ MeV
	1 GeV	$0.2262\,\binom{14}{48}_{\text{stat}}\,\binom{0}{35}_{a}\,\binom{2}{33}_{\chi}\,\binom{62}{30}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$107.4\,\binom{6}{2.3}_{\text{stat}}\,\binom{1.4}{3.7}_{\text{syst}}\,(1.4)_{t_0}$ MeV
	2 GeV	$0.2136\,\binom{13}{46}_{\text{stat}}\,\binom{0}{33}_{a}\,\binom{2}{31}_{\chi}\,\binom{58}{29}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$101.5\,\binom{6}{2.2}_{\text{stat}}\,\binom{1.4}{3.5}_{\text{syst}}\,(1.3)_{t_0}$ MeV
	10 GeV	$0.2064\,\binom{12}{44}_{\text{stat}}\,\binom{0}{32}_{a}\,\binom{2}{30}_{\chi}\,\binom{56}{28}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$98.04\,\binom{59}{2.09}_{\text{stat}}\,\binom{1.31}{3.39}_{\text{syst}}\,(1.24)_{t_0}$ MeV
θ_0		$-0.113\,\binom{15}{29}_{\text{stat}}\,\binom{0}{15}_{a}\,\binom{11}{3}_{\chi}$	$-6.5\,\binom{9}{1.7}_{\text{stat}}\,\binom{6}{9}^{\circ}_{\text{syst}}$
light/strange basis, state representation			
F_η^ℓ	$\mu =$ ∞	$0.1464\,\binom{27}{21}_{\text{stat}}\,\binom{23}{2}_{a}\,\binom{3}{14}_{\chi}\,\binom{17}{7}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$69.56\,\binom{1.28}{1.01}_{\text{stat}}\,\binom{1.13}{1.05}_{\text{syst}}\,(88)_{t_0}$ 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
$F_{\eta'}^\ell$	$\mu =$ ∞	$0.1070\,\binom{24}{88}_{\text{stat}}\,\binom{26}{3}_{a}\,\binom{11}{44}_{\chi}\,\binom{37}{20}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$50.82\,\binom{1.16}{4.18}_{\text{stat}}\,\binom{1.65}{2.75}_{\text{syst}}\,(64)_{t_0}$ 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_η^s	$\mu =$ ∞	$(-0.1683\,\binom{59}{30}_{\text{stat}}\,\binom{36}{5}_{a}\,\binom{4}{19}_{\chi}\,\binom{12}{4}_{\text{renorm}}\,8t_0^{\text{ph}})^{-1/2}$	$-79.93\,\binom{2.78}{1.43}_{\text{stat}}\,\binom{1.72}{1.09}_{\text{syst}}\,(1.01)_{t_0}$ 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
$F_{\eta'}^s$	$\mu =$ ∞	$0.1906\,\binom{71}{31}_{\text{stat}}\,\binom{0}{75}_{a}\,\binom{70}{63}_{\chi}\,\binom{38}{17}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$90.55\,\binom{3.37}{1.48}_{\text{stat}}\,\binom{3.41}{5.00}_{\text{syst}}\,(1.14)_{t_0}$ 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	$0.656\,\binom{12}{44}_{\text{stat}}\,\binom{7}{6}_{a}\,\binom{7}{14}_{\chi}\,\binom{11}{6}_{\text{renorm}}$	$37.6\,\binom{7}{2.5}_{\text{stat}}\,\binom{8}{9}^{\circ}_{\text{syst}}$

Continued on next page

JHEP08(2021)137

Table 24 (continued): Decay constants at various scales.

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 25. Decay constants, using the experimental η and η' masses as additional input (priors, see section 5.7) in various representations and at several renormalization scales.

octet/singlet basis, state representation				
F_η^8			$0.2180\,\binom{25}{32}_{\text{stat}}\,\binom{15}{22}_{a}\,\binom{33}{0}_{\chi}\,(8t_0^{\text{ph}})^{-1/2}$	$103.5\,\binom{1.2}{1.5}_{\text{stat}}\,\binom{1.7}{1.1}_{\text{syst}}\,(1.3)_{t_0}$ MeV
$F_{\eta'}^8$			$-0.105\,\binom{5}{9}_{\text{stat}}\,\binom{7}{0}_{a}\,\binom{7}{0}_{\chi}\,(8t_0^{\text{ph}})^{-1/2}$	$-50.0\,\binom{2.2}{4.2}_{\text{stat}}\,\binom{4.8}{3}_{\text{syst}}\,(6)_{t_0}$ MeV
F_η^0	$\mu =$	∞	$0.0276\,\binom{34}{36}_{\text{stat}}\,\binom{52}{0}_{a}\,\binom{0}{50}_{\chi}\,\binom{26}{11}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$13.1\,\binom{1.6}{1.7}_{\text{stat}}\,\binom{2.5}{2.7}_{\text{syst}}\,(2)_{t_0}$ MeV
		1 GeV	$0.0314\,\binom{39}{41}_{\text{stat}}\,\binom{60}{0}_{a}\,\binom{0}{56}_{\chi}\,\binom{30}{12}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$14.9\,\binom{1.9}{2.0}_{\text{stat}}\,\binom{2.9}{3.0}_{\text{syst}}\,(2)_{t_0}$ MeV
		2 GeV	$0.0297\,\binom{37}{39}_{\text{stat}}\,\binom{56}{0}_{a}\,\binom{0}{53}_{\chi}\,\binom{28}{11}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$14.1\,\binom{1.7}{1.8}_{\text{stat}}\,\binom{2.7}{2.9}_{\text{syst}}\,(2)_{t_0}$ MeV
		10 GeV	$0.0287\,\binom{36}{38}_{\text{stat}}\,\binom{54}{0}_{a}\,\binom{0}{51}_{\chi}\,\binom{27}{11}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$13.6\,\binom{1.7}{1.8}_{\text{stat}}\,\binom{2.6}{2.8}_{\text{syst}}\,(2)_{t_0}$ MeV
$F_{\eta'}^0$	$\mu =$	∞	$0.1941\,\binom{15}{39}_{\text{stat}}\,\binom{4}{10}_{a}\,\binom{34}{4}_{\chi}\,\binom{44}{23}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$92.21\,\binom{69}{1.85}_{\text{stat}}\,\binom{1.96}{2.14}_{\text{syst}}\,(1.16)_{t_0}$ MeV
		1 GeV	$0.2210\,\binom{17}{44}_{\text{stat}}\,\binom{5}{34}_{a}\,\binom{39}{4}_{\chi}\,\binom{50}{26}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$105.0\,\binom{8}{2.1}_{\text{stat}}\,\binom{2.2}{2.9}_{\text{syst}}\,(1.3)_{t_0}$ MeV
		2 GeV	$0.2087\,\binom{16}{42}_{\text{stat}}\,\binom{4}{32}_{a}\,\binom{37}{4}_{\chi}\,\binom{47}{24}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$99.14\,\binom{74}{1.99}_{\text{stat}}\,\binom{2.11}{2.70}_{\text{syst}}\,(1.25)_{t_0}$ MeV
		10 GeV	$0.2017\,\binom{15}{40}_{\text{stat}}\,\binom{4}{31}_{a}\,\binom{36}{4}_{\chi}\,\binom{45}{24}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$95.79\,\binom{72}{1.92}_{\text{stat}}\,\binom{2.04}{2.61}_{\text{syst}}\,(1.21)_{t_0}$ MeV
octet/singlet basis, angle representation				
F^8			$0.2421\,\binom{22}{26}_{\text{stat}}\,\binom{8}{50}_{a}\,\binom{32}{12}_{\chi}\,(8t_0^{\text{ph}})^{-1/2}$	$115.0\,\binom{1.1}{1.2}_{\text{stat}}\,\binom{1.6}{2.4}_{\text{syst}}\,(1.5)_{t_0}$ MeV
θ_8			$-0.450\,\binom{21}{36}_{\text{stat}}\,\binom{24}{0}_{a}\,\binom{29}{0}_{\chi}\,\binom{1}{5}_{\text{renorm}}$	$-25.8\,\binom{1.2}{2.1}_{\text{stat}}\,\binom{2.2}{0.3}^{\circ}_{\text{syst}}$
F^0	$\mu =$	∞	$0.1961\,\binom{13}{37}_{\text{stat}}\,\binom{12}{6}_{a}\,\binom{28}{7}_{\chi}\,\binom{47}{24}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$93.14\,\binom{62}{1.75}_{\text{stat}}\,\binom{1.83}{2.27}_{\text{syst}}\,(1.18)_{t_0}$ MeV
		1 GeV	$0.2232\,\binom{15}{42}_{\text{stat}}\,\binom{13}{26}_{a}\,\binom{32}{8}_{\chi}\,\binom{53}{27}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$106.0\,\binom{7}{2.0}_{\text{stat}}\,\binom{2.1}{2.9}_{\text{syst}}\,(1.3)_{t_0}$ MeV
		2 GeV	$0.2108\,\binom{14}{40}_{\text{stat}}\,\binom{12}{25}_{a}\,\binom{30}{8}_{\chi}\,\binom{50}{26}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$100.1\,\binom{7}{1.9}_{\text{stat}}\,\binom{2.0}{2.7}_{\text{syst}}\,(1.3)_{t_0}$ MeV
		10 GeV	$0.2037\,\binom{14}{38}_{\text{stat}}\,\binom{12}{24}_{a}\,\binom{29}{7}_{\chi}\,\binom{49}{25}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$96.76\,\binom{65}{1.82}_{\text{stat}}\,\binom{1.90}{2.60}_{\text{syst}}\,(1.22)_{t_0}$ MeV
θ_0			$-0.141\,\binom{18}{20}_{\text{stat}}\,\binom{0}{27}_{a}\,\binom{27}{0}_{\chi}$	$-8.1\,\binom{1.0}{1.1}_{\text{stat}}\,\binom{1.5}{1.5}^{\circ}_{\text{syst}}$
light/strange basis, state representation				
F_η^ℓ	$\mu =$	∞	$0.1484\,\binom{20}{21}_{\text{stat}}\,\binom{31}{0}_{a}\,\binom{0}{21}_{\chi}\,\binom{21}{9}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$70.48\,\binom{93}{99}_{\text{stat}}\,\binom{1.53}{1.42}_{\text{syst}}\,(89)_{t_0}$ MeV
		1 GeV	$0.1515\,\binom{23}{26}_{\text{stat}}\,\binom{37}{0}_{a}\,\binom{0}{27}_{\chi}\,\binom{24}{10}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$71.96\,\binom{1.11}{1.22}_{\text{stat}}\,\binom{1.81}{1.71}_{\text{syst}}\,(91)_{t_0}$ MeV
		2 GeV	$0.1501\,\binom{22}{23}_{\text{stat}}\,\binom{34}{0}_{a}\,\binom{0}{24}_{\chi}\,\binom{22}{9}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$71.29\,\binom{1.03}{1.10}_{\text{stat}}\,\binom{1.68}{1.58}_{\text{syst}}\,(90)_{t_0}$ MeV
		10 GeV	$0.1493\,\binom{21}{22}_{\text{stat}}\,\binom{33}{0}_{a}\,\binom{0}{23}_{\chi}\,\binom{22}{9}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$70.90\,\binom{98}{1.04}_{\text{stat}}\,\binom{1.61}{1.50}_{\text{syst}}\,(90)_{t_0}$ MeV
$F_{\eta'}^\ell$	$\mu =$	∞	$0.09773\,\binom{362}{686}_{\text{stat}}\,\binom{441}{12}_{a}\,\binom{364}{0}_{\chi}\,\binom{285}{164}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$46.42\,\binom{1.72}{3.26}_{\text{stat}}\,\binom{2.83}{1.35}_{\text{syst}}\,(59)_{t_0}$ MeV
		1 GeV	$0.1197\,\binom{40}{71}_{\text{stat}}\,\binom{44}{2}_{a}\,\binom{36}{0}_{\chi}\,\binom{33}{19}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$56.84\,\binom{1.88}{3.38}_{\text{stat}}\,\binom{2.86}{1.59}_{\text{syst}}\,(72)_{t_0}$ MeV
		2 GeV	$0.1096\,\binom{40}{70}_{\text{stat}}\,\binom{44}{1}_{a}\,\binom{36}{0}_{\chi}\,\binom{31}{18}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$52.08\,\binom{1.90}{3.32}_{\text{stat}}\,\binom{2.84}{1.48}_{\text{syst}}\,(66)_{t_0}$ MeV
		10 GeV	$0.1039\,\binom{38}{69}_{\text{stat}}\,\binom{44}{1}_{a}\,\binom{36}{0}_{\chi}\,\binom{30}{17}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$49.35\,\binom{1.82}{3.29}_{\text{stat}}\,\binom{2.83}{1.42}_{\text{syst}}\,(62)_{t_0}$ MeV
F_η^s	$\mu =$	∞	$-0.1620\,\binom{44}{40}_{\text{stat}}\,\binom{48}{7}_{a}\,\binom{0}{56}_{\chi}\,\binom{16}{6}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$-76.97\,\binom{2.07}{1.91}_{\text{stat}}\,\binom{2.29}{2.76}_{\text{syst}}\,(97)_{t_0}$ MeV
		1 GeV	$-0.1598\,\binom{46}{42}_{\text{stat}}\,\binom{52}{6}_{a}\,\binom{0}{60}_{\chi}\,\binom{18}{7}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$-75.93\,\binom{2.20}{2.00}_{\text{stat}}\,\binom{2.49}{2.97}_{\text{syst}}\,(96)_{t_0}$ MeV
		2 GeV	$-0.1609\,\binom{45}{41}_{\text{stat}}\,\binom{50}{6}_{a}\,\binom{0}{58}_{\chi}\,\binom{17}{7}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$-76.41\,\binom{2.14}{1.96}_{\text{stat}}\,\binom{2.40}{2.87}_{\text{syst}}\,(97)_{t_0}$ MeV
		10 GeV	$-0.1614\,\binom{44}{41}_{\text{stat}}\,\binom{49}{7}_{a}\,\binom{0}{57}_{\chi}\,\binom{16}{7}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$-76.68\,\binom{2.11}{1.93}_{\text{stat}}\,\binom{2.35}{2.82}_{\text{syst}}\,(97)_{t_0}$ MeV

Continued on next page

JHEP08(2021)137

Table 25 (continued): Decay constants at various scales.

$F^s_{\eta'}$	$\mu =$	∞	$0.1980\,\binom{59}{44}_{\text{stat}}\,\binom{0}{74}_{a}\,\binom{24}{58}_{\chi}\,\binom{35}{16}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$94.06\,\binom{2.80}{2.08}_{\text{stat}}\,\binom{1.37}{4.77}_{\text{syst}}\,(1.19)_{t_0}$ MeV
		1 GeV	$0.2135\,\binom{58}{47}_{\text{stat}}\,\binom{0}{76}_{a}\,\binom{27}{58}_{\chi}\,\binom{39}{18}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$101.4\,\binom{2.8}{2.2}_{\text{stat}}\,\binom{1.5}{4.9}_{\text{syst}}\,(1.3)_{t_0}$ MeV
		2 GeV	$0.2064\,\binom{58}{45}_{\text{stat}}\,\binom{0}{75}_{a}\,\binom{25}{58}_{\chi}\,\binom{37}{17}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$98.06\,\binom{2.78}{2.15}_{\text{stat}}\,\binom{1.45}{4.85}_{\text{syst}}\,(1.24)_{t_0}$ MeV
		10 GeV	$0.2024\,\binom{59}{45}_{\text{stat}}\,\binom{0}{74}_{a}\,\binom{25}{58}_{\chi}\,\binom{36}{17}_{\text{renorm}}\,(8t_0^{\text{ph}})^{-1/2}$	$96.13\,\binom{2.79}{2.12}_{\text{stat}}\,\binom{1.41}{4.81}_{\text{syst}}\,(1.21)_{t_0}$ MeV
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	$\mu =$	∞	$0.582\,\binom{19}{36}_{\text{stat}}\,\binom{11}{4}_{a}\,\binom{20}{0}_{\chi}\,\binom{7}{5}_{\text{renorm}}$	$33.3\,\binom{1.1}{2.1}_{\text{stat}}\,\binom{1.3}{0.4}^{\circ}_{\text{syst}}$
		1 GeV	$0.669\,\binom{19}{34}_{\text{stat}}\,\binom{6}{5}_{a}\,\binom{21}{0}_{\chi}\,\binom{6}{5}_{\text{renorm}}$	$38.3\,\binom{1.1}{1.9}_{\text{stat}}\,\binom{1.3}{0.4}^{\circ}_{\text{syst}}$
		2 GeV	$0.631\,\binom{19}{35}_{\text{stat}}\,\binom{8}{4}_{a}\,\binom{20}{0}_{\chi}\,\binom{7}{5}_{\text{renorm}}$	$36.2\,\binom{1.1}{2.0}_{\text{stat}}\,\binom{1.3}{0.4}^{\circ}_{\text{syst}}$
		10 GeV	$0.608\,\binom{19}{35}_{\text{stat}}\,\binom{9}{4}_{a}\,\binom{19}{0}_{\chi}\,\binom{7}{5}_{\text{renorm}}$	$34.8\,\binom{1.1}{2.0}_{\text{stat}}\,\binom{1.3}{0.4}^{\circ}_{\text{syst}}$
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	$\mu =$	∞	$0.686\,\binom{17}{23}_{\text{stat}}\,\binom{6}{4}_{a}\,\binom{24}{0}_{\chi}\,\binom{6}{14}_{\text{renorm}}$	$39.3\,\binom{1.0}{1.3}_{\text{stat}}\,\binom{1.5}{0.8}^{\circ}_{\text{syst}}$
		1 GeV	$0.643\,\binom{16}{22}_{\text{stat}}\,\binom{6}{4}_{a}\,\binom{23}{0}_{\chi}\,\binom{6}{14}_{\text{renorm}}$	$36.8\,\binom{0.9}{1.3}_{\text{stat}}\,\binom{1.4}{0.8}^{\circ}_{\text{syst}}$
		2 GeV	$0.662\,\binom{17}{23}_{\text{stat}}\,\binom{6}{4}_{a}\,\binom{23}{0}_{\chi}\,\binom{6}{14}_{\text{renorm}}$	$37.9\,\binom{1.0}{1.3}_{\text{stat}}\,\binom{1.4}{0.8}^{\circ}_{\text{syst}}$
		10 GeV	$0.673\,\binom{17}{23}_{\text{stat}}\,\binom{6}{4}_{a}\,\binom{23}{0}_{\chi}\,\binom{6}{14}_{\text{renorm}}$	$38.6\,\binom{1.0}{1.3}_{\text{stat}}\,\binom{1.4}{0.8}^{\circ}_{\text{syst}}$

References

[1] E. Witten, *Instantons, the quark model, and the* $1/N$ *expansion*, *Nucl. Phys. B* **149** (1979) 285 [INSPIRE].

[2] G. Veneziano, U(1) *without instantons*, *Nucl. Phys. B* **159** (1979) 213 [INSPIRE].

[3] P. Di Vecchia and G. Veneziano, *Chiral dynamics in the large N limit*, *Nucl. Phys. B* **171** (1980) 253 [INSPIRE].

[4] K. Kawarabayashi and N. Ohta, *The problem of* η *in the large N limit: Effective Lagrangian approach*, *Nucl. Phys. B* **175** (1980) 477 [INSPIRE].

[5] P. Di Vecchia, F. Nicodemi, R. Pettorino and G. Veneziano, *Large n, chiral approach to pseudoscalar masses, mixings and decays*, *Nucl. Phys. B* **181** (1981) 318 [INSPIRE].

[6] H. Leutwyler, *On the* $1/N$ *expansion in chiral perturbation theory*, *Nucl. Phys. B Proc. Suppl.* **64** (1998) 223 [hep-ph/9709408] [INSPIRE].

[7] L. Gan, B. Kubis, E. Passemar and S. Tulin, *Precision tests of fundamental physics with* η *and* η' *mesons*, arXiv:2007.00664 [INSPIRE].

[8] Belle-II collaboration, *The Belle II Physics Book*, *PTEP* **2019** (2019) 123C01 [*Erratum ibid.* **2020** (2020) 029201] [arXiv:1808.10567] [INSPIRE].

JHEP08(2021)137

[9] M. Beneke and M. Neubert, *Flavor singlet B decay amplitudes in QCD factorization*, *Nucl. Phys. B* **651** (2003) 225 [hep-ph/0210085] [INSPIRE].

[10] P. Kroll and K. Passek-Kumericki, *The two gluon components of the η and η' mesons to leading twist accuracy*, *Phys. Rev. D* **67** (2003) 054017 [hep-ph/0210045] [INSPIRE].

[11] P. Ball and G.W. Jones, *$B \to \eta^{(\prime)}$ form factors in QCD*, *JHEP* **08** (2007) 025 [arXiv:0706.3628] [INSPIRE].

[12] S.S. Agaev, V.M. Braun, N. Offen, F.A. Porkert and A. Schäfer, *Transition form factors $\gamma^*\gamma \to \eta$ and $\gamma^*\gamma \to \eta'$ in QCD*, *Phys. Rev. D* **90** (2014) 074019 [arXiv:1409.4311] [INSPIRE].

[13] A.E. Blechman, S. Mantry and I.W. Stewart, *Heavy quark symmetry in isosinglet nonleptonic B-decays*, *Phys. Lett. B* **608** (2005) 77 [hep-ph/0410312] [INSPIRE].

[14] L.A. Harland-Lang, V.A. Khoze, M.G. Ryskin and W.J. Stirling, *Central exclusive production as a probe of the gluonic component of the η' and η mesons*, *Eur. Phys. J. C* **73** (2013) 2429 [arXiv:1302.2004] [INSPIRE].

[15] RQCD collaboration, *Light-cone distribution amplitudes of pseudoscalar mesons from lattice QCD*, *JHEP* **08** (2019) 065 [*Addendum ibid.* **11** (2020) 037] [arXiv:1903.08038] [INSPIRE].

[16] A. Gérardin, H.B. Meyer and A. Nyffeler, *Lattice calculation of the pion transition form factor with $N_f = 2 + 1$ Wilson quarks*, *Phys. Rev. D* **100** (2019) 034520 [arXiv:1903.09471] [INSPIRE].

[17] G.S. Bali et al., *Pion distribution amplitude from Euclidean correlation functions: Exploring universality and higher-twist effects*, *Phys. Rev. D* **98** (2018) 094507 [arXiv:1807.06671] [INSPIRE].

[18] G.S. Bali, S. Collins, S. Dürr and I. Kanamori, *$D_s \to \eta, \eta'$ semileptonic decay form factors with disconnected quark loop contributions*, *Phys. Rev. D* **91** (2015) 014503 [arXiv:1406.5449] [INSPIRE].

[19] ETM collaboration, *Flavor-singlet meson decay constants from $N_f = 2 + 1 + 1$ twisted mass lattice QCD*, *Phys. Rev. D* **97** (2018) 054508 [arXiv:1710.07986] [INSPIRE].

[20] Y. Kuramashi, M. Fukugita, H. Mino, M. Okawa and A. Ukawa, *η' meson mass in lattice QCD*, *Phys. Rev. Lett.* **72** (1994) 3448 [INSPIRE].

[21] L. Venkataraman and G. Kilcup, *The η' meson with staggered fermions*, hep-lat/9711006 [INSPIRE].

[22] W.A. Bardeen, A. Duncan, E. Eichten and H. Thacker, *Anomalous chiral behavior in quenched lattice QCD*, *Phys. Rev. D* **62** (2000) 114505 [hep-lat/0007010] [INSPIRE].

[23] TXL and T(X)L collaborations, *Flavor singlet pseudoscalar masses in $N_f = 2$ QCD*, *Phys. Rev. D* **63** (2001) 074503 [hep-lat/0010005] [INSPIRE].

[24] UKQCD collaboration, *The η and η' mesons in QCD*, *Phys. Lett. B* **491** (2000) 123 [*Erratum ibid.* **551** (2003) 391] [hep-lat/0006020] [INSPIRE].

[25] SESAM and T(X)L collaborations, *Quark mass effects on the topological susceptibility in QCD*, *Phys. Rev. D* **64** (2001) 054502 [hep-lat/0102002] [INSPIRE].

[26] CP-PACS collaboration, *Flavor singlet meson mass in the continuum limit in two flavor lattice QCD*, *Phys. Rev. D* **67** (2003) 074503 [hep-lat/0211040] [INSPIRE].

JHEP08(2021)137

[27] K. Hashimoto and T. Izubuchi, *η' meson from two flavor dynamical domain wall fermions*, *Prog. Theor. Phys.* **119** (2008) 599 [arXiv:0803.0186] [INSPIRE].

[28] ETM collaboration, *The η' meson from lattice QCD*, *Eur. Phys. J. C* **58** (2008) 261 [arXiv:0804.3871] [INSPIRE].

[29] W. Sun et al., *Glueball spectrum from $N_f = 2$ lattice QCD study on anisotropic lattices*, *Chin. Phys. C* **42** (2018) 093103 [arXiv:1702.08174] [INSPIRE].

[30] P. Dimopoulos et al., *Topological susceptibility and η' meson mass from $N_f = 2$ lattice QCD at the physical point*, *Phys. Rev. D* **99** (2019) 034511 [arXiv:1812.08787] [INSPIRE].

[31] N.H. Christ et al., *The η and η' mesons from Lattice QCD*, *Phys. Rev. Lett.* **105** (2010) 241601 [arXiv:1002.2999] [INSPIRE].

[32] J.J. Dudek, R.G. Edwards, B. Joo, M.J. Peardon, D.G. Richards and C.E. Thomas, *Isoscalar meson spectroscopy from lattice QCD*, *Phys. Rev. D* **83** (2011) 111502 [arXiv:1102.4299] [INSPIRE].

[33] UKQCD collaboration, *A study of the η and η' mesons with improved staggered fermions*, *Phys. Rev. D* **86** (2012) 014504 [arXiv:1112.4384] [INSPIRE].

[34] JLQCD collaboration, *η' meson mass from topological charge density correlator in QCD*, *Phys. Rev. D* **92** (2015) 111501 [arXiv:1509.00944] [INSPIRE].

[35] ETM collaboration, *η and η' mesons from $N_f = 2 + 1 + 1$ twisted mass lattice QCD*, *JHEP* **11** (2012) 048 [arXiv:1206.6719] [INSPIRE].

[36] ETM collaboration, *η and η' mixing from Lattice QCD*, *Phys. Rev. Lett.* **111** (2013) 181602 [arXiv:1310.1207] [INSPIRE].

[37] A.Y. Kotov, M.P. Lombardo and A.M. Trunin, *Fate of the η' in the quark gluon plasma*, *Phys. Lett. B* **794** (2019) 83 [arXiv:1903.05633] [INSPIRE].

[38] M. Bruno et al., *Simulation of QCD with $N_f = 2 + 1$ flavors of non-perturbatively improved Wilson fermions*, *JHEP* **02** (2015) 043 [arXiv:1411.3982] [INSPIRE].

[39] RQCD collaboration, *Lattice simulations with $N_f = 2 + 1$ improved Wilson fermions at a fixed strange quark mass*, *Phys. Rev. D* **94** (2016) 074501 [arXiv:1606.09039] [INSPIRE].

[40] J. Kodaira, *QCD Higher Order Effects in Polarized Electroproduction: Flavor Singlet Coefficient Functions*, *Nucl. Phys. B* **165** (1980) 129 [INSPIRE].

[41] T. Feldmann, P. Kroll and B. Stech, *Mixing and decay constants of pseudoscalar mesons*, *Phys. Rev. D* **58** (1998) 114006 [hep-ph/9802409] [INSPIRE].

[42] P. Bickert, P. Masjuan and S. Scherer, *η-η' Mixing in Large-N_c Chiral Perturbation Theory*, *Phys. Rev. D* **95** (2017) 054023 [arXiv:1612.05473] [INSPIRE].

[43] J. Schechter, A. Subbaraman and H. Weigel, *Effective hadron dynamics: From meson masses to the proton spin puzzle*, *Phys. Rev. D* **48** (1993) 339 [hep-ph/9211239] [INSPIRE].

[44] M. Bruno et al., *The Λ-parameter in 3-flavour QCD and $\alpha_s(m_Z)$ by the ALPHA collaboration*, *PoS* **LATTICE2016** (2016) 197 [arXiv:1701.03075] [INSPIRE].

[45] M. Lüscher and S. Schaefer, *Lattice QCD without topology barriers*, *JHEP* **07** (2011) 036 [arXiv:1105.4749] [INSPIRE].

[46] T. Bhattacharya, R. Gupta, W. Lee, S.R. Sharpe and J.M.S. Wu, *Improved bilinears in lattice QCD with non-degenerate quarks*, *Phys. Rev. D* **73** (2006) 034504 [hep-lat/0511014] [INSPIRE].

JHEP08(2021)137

[47] RQCD collaboration, *Scale setting and the light hadron spectrum in $N_f = 2 + 1$ QCD with Wilson fermions*, in preparation.

[48] M. Lüscher, *Properties and uses of the Wilson flow in lattice QCD*, *JHEP* **08** (2010) 071 [*Erratum ibid.* **03** (2014) 092] [arXiv:1006.4518] [INSPIRE].

[49] M. Bruno, T. Korzec and S. Schaefer, *Setting the scale for the CLS $2 + 1$ flavor ensembles*, *Phys. Rev. D* **95** (2017) 074504 [arXiv:1608.08900] [INSPIRE].

[50] ALPHA collaboration, *QCD coupling from a nonperturbative determination of the three-flavor Λ parameter*, *Phys. Rev. Lett.* **119** (2017) 102001 [arXiv:1706.03821] [INSPIRE].

[51] O. Bär and M. Golterman, *Chiral perturbation theory for gradient flow observables*, *Phys. Rev. D* **89** (2014) 034505 [*Erratum ibid.* **89** (2014) 099905] [arXiv:1312.4999] [INSPIRE].

[52] S. Güsken, U. Löw, K.H. Mütter, R. Sommer, A. Patel and K. Schilling, *Nonsinglet axial vector couplings of the baryon octet in lattice QCD*, *Phys. Lett. B* **227** (1989) 266 [INSPIRE].

[53] M. Falcioni, M.L. Paciello, G. Parisi and B. Taglienti, *Again on* SU(3) *glueball mass*, *Nucl. Phys. B* **251** (1985) 624 [INSPIRE].

[54] S. Bernardson, P. McCarty and C. Thron, *Monte Carlo methods for estimating linear combinations of inverse matrix entries in lattice QCD*, *Comput. Phys. Commun.* **78** (1993) 256 [INSPIRE].

[55] C. Thron, S.J. Dong, K.F. Liu and H.P. Ying, *Padé-Z_2 estimator of determinants*, *Phys. Rev. D* **57** (1998) 1642 [hep-lat/9707001] [INSPIRE].

[56] SESAM collaboration, *Observation of string breaking in QCD*, *Phys. Rev. D* **71** (2005) 114513 [hep-lat/0505012] [INSPIRE].

[57] G.S. Bali, S. Collins and A. Schäfer, *Effective noise reduction techniques for disconnected loops in Lattice QCD*, *Comput. Phys. Commun.* **181** (2010) 1570 [arXiv:0910.3970] [INSPIRE].

[58] A. Frommer, K. Kahl, S. Krieg, B. Leder and M. Rottmann, *Adaptive aggregation based Domain Decomposition Multigrid for the lattice Wilson Dirac operator*, *SIAM J. Sci. Comput.* **36** (2014) A1581 [arXiv:1303.1377] [INSPIRE].

[59] S. 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*, (2014), DOI [arXiv:1412.2629] [INSPIRE].

[60] P. Georg, D. Richtmann and T. Wettig, *DD-αAMG on QPACE 3*, *EPJ Web Conf.* **175** (2018) 02007 [arXiv:1710.07041] [INSPIRE].

[61] C. Michael, *Adjoint sources in lattice gauge theory*, *Nucl. Phys. B* **259** (1985) 58 [INSPIRE].

[62] M. Lüscher and U. Wolff, *How to calculate the elastic scattering matrix in two-dimensional quantum field theories by numerical simulation*, *Nucl. Phys. B* **339** (1990) 222 [INSPIRE].

[63] Particle Data Group collaboration, *Review of Particle Physics*, *PTEP* **2020** (2020) 083C01 [INSPIRE].

[64] G. Bali, S. Collins and J. Simeth, *η and η' masses and decay constants*, *EPJ Web Conf.* **175** (2018) 05028 [arXiv:1710.06733] [INSPIRE].

JHEP08(2021)137

[65] X. Feng, K. Jansen and D.B. Renner, *The $\pi^+\pi^+$ scattering length from maximally twisted mass lattice QCD*, *Phys. Lett. B* **684** (2010) 268 [arXiv:0909.3255] [INSPIRE].

[66] T. Umeda, *A constant contribution in meson correlators at finite temperature*, *Phys. Rev. D* **75** (2007) 094502 [hep-lat/0701005] [INSPIRE].

[67] B. Blossier, M. Della Morte, G. von Hippel, T. Mendes and R. Sommer, *On the generalized eigenvalue method for energies and matrix elements in lattice field theory*, *JHEP* **04** (2009) 094 [arXiv:0902.1265] [INSPIRE].

[68] ALPHA collaboration, *Non-perturbative improvement of the axial current in $N_f = 3$ lattice QCD with Wilson fermions and tree-level improved gauge action*, *Nucl. Phys. B* **896** (2015) 555 [arXiv:1502.04999] [INSPIRE].

[69] P. Korcyl and G.S. Bali, *Non-perturbative determination of improvement coefficients using coordinate space correlators in $N_f = 2 + 1$ lattice QCD*, *Phys. Rev. D* **95** (2017) 014505 [arXiv:1607.07090] [INSPIRE].

[70] G.S. Bali, K.G. Chetyrkin, P. Korcyl and J. Simeth, *Non-perturbative determination of quark-mass independent improvement coefficients in $n_f = 2 + 1$ lattice QCD*, in preparation (2021).

[71] M. Dalla Brida, T. Korzec, S. Sint and P. Vilaseca, *High precision renormalization of the flavour non-singlet Noether currents in lattice QCD with Wilson quarks*, *Eur. Phys. J. C* **79** (2019) 23 [arXiv:1808.09236] [INSPIRE].

[72] A. Gérardin, T. Harris and H.B. Meyer, *Nonperturbative renormalization and $O(a)$-improvement of the nonsinglet vector current with $N_f = 2 + 1$ Wilson fermions and tree-level Symanzik improved gauge action*, *Phys. Rev. D* **99** (2019) 014519 [arXiv:1811.08209] [INSPIRE].

[73] S.A. Larin, *The Renormalization of the axial anomaly in dimensional regularization*, *Phys. Lett. B* **303** (1993) 113 [hep-ph/9302240] [INSPIRE].

[74] T. Ahmed, L. Chen and M. Czakon, *Renormalization of the flavor-singlet axial-vector current and its anomaly in dimensional regularization*, *JHEP* **05** (2021) 087 [arXiv:2101.09479] [INSPIRE].

[75] M.F. Zoller, *OPE of the pseudoscalar gluonium correlator in massless QCD to three-loop order*, *JHEP* **07** (2013) 040 [arXiv:1304.2232] [INSPIRE].

[76] M. Constantinou, M. Hadjiantonis, H. Panagopoulos and G. Spanoudes, *Singlet versus nonsinglet perturbative renormalization of fermion bilinears*, *Phys. Rev. D* **94** (2016) 114513 [arXiv:1610.06744] [INSPIRE].

[77] P.A. Baikov, K.G. Chetyrkin and J.H. Kühn, *Quark mass and field anomalous dimensions to $\mathcal{O}(\alpha_s^5)$*, *JHEP* **10** (2014) 076 [arXiv:1402.6611] [INSPIRE].

[78] F. Herren and M. Steinhauser, *Version 3 of RunDec and CRunDec*, *Comput. Phys. Commun.* **224** (2018) 333 [arXiv:1703.03751] [INSPIRE].

[79] K.G. Chetyrkin, J.H. Kühn and M. Steinhauser, *RunDec: A Mathematica package for running and decoupling of the strong coupling and quark masses*, *Comput. Phys. Commun.* **133** (2000) 43 [hep-ph/0004189] [INSPIRE].

[80] J. Gasser and H. Leutwyler, *Chiral Perturbation Theory: Expansions in the mass of the strange quark*, *Nucl. Phys. B* **250** (1985) 465 [INSPIRE].

JHEP08(2021)137

[81] X.-K. Guo, Z.-H. Guo, J.A. Oller and J.J. Sanz-Cillero, *Scrutinizing the η-η' mixing, masses and pseudoscalar decay constants in the framework of* U(3) *chiral effective field theory*, *JHEP* **06** (2015) 175 [arXiv:1503.02248] [INSPIRE].

[82] J. Orear, *Least squares when both variables have uncertainties*, *Am. J. Phys.* **50** (1982) 912 [INSPIRE].

[83] R. Kaiser and H. Leutwyler, *Large N_c in chiral perturbation theory*, *Eur. Phys. J. C* **17** (2000) 623 [hep-ph/0007101] [INSPIRE].

[84] ALPHA collaboration, *Non-perturbative quark mass renormalisation and running in $N_f = 3$ QCD*, *Eur. Phys. J. C* **78** (2018) 387 [arXiv:1802.05243] [INSPIRE].

[85] J. Liang, Y.-B. Yang, T. Draper, M. Gong and K.-F. Liu, *Quark spins and Anomalous Ward Identity*, *Phys. Rev. D* **98** (2018) 074505 [arXiv:1806.08366] [INSPIRE].

[86] L. Giusti, G.C. Rossi, M. Testa and G. Veneziano, *The $U_A(1)$ problem on the lattice with Ginsparg-Wilson fermions*, *Nucl. Phys. B* **628** (2002) 234 [hep-lat/0108009] [INSPIRE].

[87] L. Del Debbio and C. Pica, *Topological susceptibility from the overlap*, *JHEP* **02** (2004) 003 [hep-lat/0309145] [INSPIRE].

[88] H. Leutwyler and A.V. Smilga, *Spectrum of Dirac operator and role of winding number in QCD*, *Phys. Rev. D* **46** (1992) 5607 [INSPIRE].

[89] MILC collaboration, *Topological susceptibility with the asqtad action*, *Phys. Rev. D* **81** (2010) 114501 [arXiv:1003.5695] [INSPIRE].

[90] A. Chowdhury et al., *Topological susceptibility in Lattice QCD with unimproved Wilson fermions*, *Phys. Lett. B* **707** (2012) 228 [arXiv:1110.6013] [INSPIRE].

[91] ETM collaboration, *Topological susceptibility from the twisted mass Dirac operator spectrum*, *JHEP* **02** (2014) 119 [arXiv:1312.5161] [INSPIRE].

[92] ALPHA collaboration, *Topological susceptibility and the sampling of field space in $N_f = 2$ lattice QCD simulations*, *JHEP* **08** (2014) 150 [arXiv:1406.5363] [INSPIRE].

[93] C. Bonati et al., *Axion phenomenology and θ-dependence from $N_f = 2 + 1$ lattice QCD*, *JHEP* **03** (2016) 155 [arXiv:1512.06746] [INSPIRE].

[94] S. Borsányi et al., *Calculation of the axion mass based on high-temperature lattice quantum chromodynamics*, *Nature* **539** (2016) 69 [arXiv:1606.07494] [INSPIRE].

[95] C. Alexandrou et al., *Topological susceptibility from twisted mass fermions using spectral projectors and the gradient flow*, *Phys. Rev. D* **97** (2018) 074503 [arXiv:1709.06596] [INSPIRE].

[96] T. Feldmann, P. Kroll and B. Stech, *Mixing and decay constants of pseudoscalar mesons: The sequel*, *Phys. Lett. B* **449** (1999) 339 [hep-ph/9812269] [INSPIRE].

[97] T. Feldmann, *Quark structure of pseudoscalar mesons*, *Int. J. Mod. Phys. A* **15** (2000) 159 [hep-ph/9907491] [INSPIRE].

[98] R. Escribano, S. Gonzàlez-Solís, P. Masjuan and P. Sanchez-Puertas, *η' transition form factor from space- and timelike experimental data*, *Phys. Rev. D* **94** (2016) 054033 [arXiv:1512.07520] [INSPIRE].

[99] R. Kaiser and H. Leutwyler, *Pseudoscalar decay constants at large N_c*, in *Workshop on Methods of Nonperturbative Quantum Field Theory*, (1998) [hep-ph/9806336] [INSPIRE].

JHEP08(2021)137

[100] Z.-H. Guo and J.A. Oller, *Resonances from meson-meson scattering in* U(3) *CHPT*, *Phys. Rev. D* **84** (2011) 034005 [arXiv:1104.2849] [INSPIRE].

[101] M. Benayoun, L. DelBuono and H.B. O'Connell, *VMD, the WZW Lagrangian and ChPT: The third mixing angle*, *Eur. Phys. J. C* **17** (2000) 593 [hep-ph/9905350] [INSPIRE].

[102] R. Escribano and J.-M. Frère, *Study of the η-η' system in the two mixing angle scheme*, *JHEP* **06** (2005) 029 [hep-ph/0501072] [INSPIRE].

[103] R. Escribano, P. Masjuan and P. Sanchez-Puertas, *η and η' transition form factors from rational approximants*, *Phys. Rev. D* **89** (2014) 034014 [arXiv:1307.2061] [INSPIRE].

[104] Y.-H. Chen, Z.-H. Guo and B.-S. Zou, *Unified study of $J/\psi \to PV$, $P\gamma^{(*)}$ and light hadron radiative processes*, *Phys. Rev. D* **91** (2015) 014010 [arXiv:1411.1159] [INSPIRE].

[105] R. Escribano, P. Masjuan and P. Sanchez-Puertas, *The η transition form factor from space- and time-like experimental data*, *Eur. Phys. J. C* **75** (2015) 414 [arXiv:1504.07742] [INSPIRE].

[106] M. Ding et al., *$\gamma^*\gamma \to \eta, \eta'$ transition form factors*, *Phys. Rev. D* **99** (2019) 014014 [arXiv:1810.12313] [INSPIRE].

[107] X.-W. Gu, C.-G. Duan and Z.-H. Guo, *Updated study of the η-η' mixing and the thermal properties of light pseudoscalar mesons at low temperatures*, *Phys. Rev. D* **98** (2018) 034007 [arXiv:1803.07284] [INSPIRE].

[108] M. Benayoun, L. DelBuono, S. Eidelman, V.N. Ivanchenko and H.B. O'Connell, *Radiative decays, nonet symmetry and* SU(3) *breaking*, *Phys. Rev. D* **59** (1999) 114027 [hep-ph/9902326] [INSPIRE].

[109] Flavour Lattice Averaging Group collaboration, *FLAG Review 2019*, *Eur. Phys. J. C* **80** (2020) 113 [arXiv:1902.08191] [INSPIRE].

[110] T. Feldmann and P. Kroll, *Mixing of pseudoscalar mesons*, *Phys. Scripta T* **99** (2002) 13 [hep-ph/0201044] [INSPIRE].

[111] P. Hernández, C. Pena and F. Romero-López, *Large N_c scaling of meson masses and decay constants*, *Eur. Phys. J. C* **79** (2019) 865 [arXiv:1907.11511] [INSPIRE].

[112] J. Bijnens and G. Ecker, *Mesonic low-energy constants*, *Ann. Rev. Nucl. Part. Sci.* **64** (2014) 149 [arXiv:1405.6488] [INSPIRE].

[113] J. Bijnens and I. Jemos, *A new global fit of the L_i^r at next-to-next-to-leading order in Chiral Perturbation Theory*, *Nucl. Phys. B* **854** (2012) 631 [arXiv:1103.5945] [INSPIRE].

[114] P. Herrera-Siklódy, J.I. Latorre, P. Pascual and J. Taron, *Chiral effective Lagrangian in the large N_c limit: The nonet case*, *Nucl. Phys. B* **497** (1997) 345 [hep-ph/9610549] [INSPIRE].

[115] MILC collaboration, *Results for light pseudoscalar mesons*, *PoS* **LATTICE2010** (2010) 074 [arXiv:1012.0868] [INSPIRE].

[116] V.A. Novikov, M.A. Shifman, A.I. Vainshtein and V.I. Zakharov, *A theory of the $J/\psi \to \eta(\eta')\gamma$ decays*, *Nucl. Phys. B* **165** (1980) 55 [INSPIRE].

[117] J.P. Singh, *Light quark pseudoscalar densities and anomaly matrix elements for η and η' mesons*, *Phys. Rev. D* **88** (2013) 096005 [arXiv:1307.3311] [INSPIRE].

[118] H.-Y. Cheng, H.-n. Li and K.-F. Liu, *Pseudoscalar glueball mass from $\eta - eta' - G$ mixing*, *Phys. Rev. D* **79** (2009) 014024 [arXiv:0811.2577] [INSPIRE].

JHEP08(2021)137

[119] W. Qin, Q. Zhao and X.-H. Zhong, *Revisiting the pseudoscalar meson and glueball mixing and key issues in the search for a pseudoscalar glueball state*, *Phys. Rev. D* **97** (2018) 096002 [arXiv:1712.02550] [INSPIRE].

[120] H. Goldberg, *The quark-glue structure of the η and η' mesons with application to $\psi \to \eta(\eta')\gamma$ and $\psi' \to \psi\eta$*, *Phys. Rev. Lett.* **44** (1980) 363 [INSPIRE].

[121] CLEO collaboration, *Measurements of the meson-photon transition form-factors of light pseudoscalar mesons at large momentum transfer*, *Phys. Rev. D* **57** (1998) 33 [hep-ex/9707031] [INSPIRE].

[122] BABAR collaboration, *Measurement of the $\gamma\gamma^* \to \eta$ and $\gamma\gamma^* \to \eta'$ transition form factors*, *Phys. Rev. D* **84** (2011) 052001 [arXiv:1101.1142] [INSPIRE].

[123] BABAR collaboration, *Measurement of the η and η' transition form-factors at $q^2 = 112$-GeV^2*, *Phys. Rev. D* **74** (2006) 012002 [hep-ex/0605018] [INSPIRE].

[124] LHC HIGGS CROSS SECTION WORKING GROUP collaboration, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the nature of the Higgs sector*, arXiv:1610.07922 [INSPIRE].

[125] S.S. Agaev, V.M. Braun, N. Offen and F.A. Porkert, *BELLE data on the $\pi^0\gamma * \gamma$ form factor: A game changer?*, *Phys. Rev. D* **86** (2012) 077504 [arXiv:1206.3968] [INSPIRE].

[126] Jülich Supercomputing Centre, *JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre*, *JLSRF* **1** (2015) A1.

[127] Jülich Supercomputing Centre, *JUWELS: Modular Tier-0/1 Supercomputer at the Jülich Supercomputing Centre*, *JLSRF* **5** (2019) A135.

[128] Jülich Supercomputing Centre, *JURECA: Modular supercomputer at Jülich Supercomputing Centre*, *JLSRF* **4** (2018) A132.

[129] P. Arts et al., *QPACE 2 and Domain Decomposition on the Intel Xeon Phi*, *PoS* **LATTICE2014** (2015) 021 [arXiv:1502.04025] [INSPIRE].

[130] S. Heybrock, M. Rottmann, P. Georg and T. Wettig, *Adaptive algebraic multigrid on SIMD architectures*, *PoS* **LATTICE2015** (2016) 036 [arXiv:1512.04506] [INSPIRE].

[131] D. Richtmann, S. Heybrock and T. Wettig, *Multiple right-hand-side setup for the DD-αAMG*, *PoS* **LATTICE2015** (2016) 035 [arXiv:1601.03184] [INSPIRE].

[132] P. Georg, D. Richtmann and T. Wettig, *pMR: A high-performance communication library*, *PoS* **LATTICE2016** (2017) 361 [arXiv:1701.08521] [INSPIRE].

JHEP08(2021)137