Scrutinizing the $\eta$-$\eta'$ mixing, masses and pseudoscalar decay constants in the framework of $U(3)$ chiral effective field theory

Xu-Kun Guo$^a$, Zhi-Hui Guo$^{a,b,*}$, José Antonio Oller$^c$, and Juan José Sanz-Cillero$^d$

$^a$ Department of Physics, Hebei Normal University, Shijiazhuang 050024, People’s Republic of China
$^b$ State Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, CAS, Beijing 100190, People’s Republic of China
$^c$ Departamento de Física, Universidad de Murcia, E-30071 Murcia, Spain
$^d$ Departamento de Física Teórica and Instituto de Física Teórica, IFT-UAM/CSIC, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain

We study the $\eta$-$\eta'$ mixing up to next-to-next-to-leading-order in $U(3)$ chiral perturbation theory in the light of recent lattice simulations and phenomenological inputs. A general treatment for the $\eta$-$\eta'$ mixing at higher orders, with the higher-derivative, kinematic and mass mixing terms, is addressed. The connections between the four mixing parameters in the two-mixing-angle scheme and the low energy constants in the $U(3)$ chiral effective theory are provided both for the singlet-octet and the quark-flavor bases. The axial-vector decay constants of pion and kaon are studied in the same order and confronted with the lattice simulation data as well. The quark-mass dependences of $m_\eta$, $m_{\eta'}$ and $m_K$ are found to be well described at next-to-leading order. Nonetheless, in order to simultaneously describe the lattice data and phenomenological determinations for the properties of light pseudoscalars $\pi$, $K$, $\eta$ and $\eta'$, the next-to-next-to-leading order study is essential. Furthermore, the lattice and phenomenological inputs are well reproduced for reasonable values of low the energy constants, compatible with previous bibliography.

PACS numbers: 12.39.Fe
Keywords: Chiral Lagrangian, $\eta$-$\eta'$ mixing

I. INTRODUCTION

The phenomenology of light flavor pseudoscalar mesons $\eta$ and $\eta'$ provides a valuable window on many important nonperturbative features of Quantum Chromodynamics (QCD). It includes such important aspects as:

- The spontaneous breaking of chiral symmetry, which gives rise to the appearance of the multiplet of light pseudoscalar mesons.
- The $U_A(1)$ anomaly of strong interactions, which gives mass to the singlet $\eta_0$ in $N_C = 3$ QCD, even in the chiral limit.
- The explicit $SU(3)$-flavor symmetry breaking, due to the splitting $m_s \neq \hat{m}$ between the strange and up/down quark masses (the isospin limit, where $m_u = m_d = \hat{m}$ and the electromagnetic corrections are neglected, will be assumed all through the article).
- The $1/N_C$ expansion of QCD in the limit of large $N_C$, with $N_C$ the number of colors in QCD.

The interaction between the pseudo-Nambu-Goldstone bosons (pNGBs) ($\pi, K, \eta_0$) from the spontaneous chiral symmetry breaking can be systematically described through a low-energy effective field theory (EFT) based on $SU(3)_L \times SU(3)_R$ chiral symmetry, namely Chiral Perturbation Theory ($\chi$PT) [1]. Following large–$N_C$ arguments [2], this approach was later extended, incorporating the singlet $\eta_0$ into a $U(3)$ $\chi$PT Lagrangian [3–8]. This combination of $\chi$PT and the $1/N_C$ expansion provides a consistent framework which addresses all the previous issues.

More precisely, in this article we show that this large–$N_C$ $\chi$PT framework yields an excellent description of the $\eta$ and $\eta'$ masses from lattice simulations at different light-quark masses [9–13]. Constraints from phenomenological studies of $\rho$, $\omega$, $\phi$, $J/\psi$ decays [14, 15] and kaon mass lattice simulations [16, 17] are compatible and easily accommodated in a joint fit. The problems arise when one tries to also describe lattice simulations for $F_\pi$, $F_K$ and $F_K/F_\pi$ [16–18].

*Corresponding author: zhguo@mail.hebtu.edu.cn
Nevertheless, the issue of these observables in $\chi$PT is known and has been widely discussed in previous bibliography [19–23]. It constitutes a problem in its own and it is not the central goal of this article. It is discussed for sake of completeness and to show its impact in a global fit.

The $\eta$ and $\eta'$ mesons not only attract much attention from the chiral community but they have been also intensively scrutinized in lattice QCD simulations, where enormous progresses have been recently made by different groups [9–13]. Varying the light-quark masses $m$ and $m_s$, both their masses and mixing angles have been extracted in the range 200 MeV $< m_\pi <$ 700 MeV. We will focus on the simulation points with $m_\pi <$ 500 MeV in the present work. By observing the dependence of these observables with the light-quark masses we will determine the $\chi$PT low energy constants (LECs) and further constrain the theoretical models. At the practical level we have recast all $\hat{m}$ dependencies in terms of $m_\pi$ and study the observables as functions of $m_\pi$. The $\eta$ and $\eta'$ lattice simulations have not been thoroughly analyzed in the chiral framework yet and it is the central goal of the present work. However, the numerical uncertainties resulting from our analyses in this work must be taken with a grain of salt as correlations between the different lattice data points and other systematic errors are not considered here.

In addition to lattice QCD, there are also phenomenological studies of the $\eta$ and $\eta'$ mixing, which has been extensively investigated in radiative decays of light-flavor vector resonances $\rho, \omega, \phi$ and $J/\psi \rightarrow VP, P\gamma$ processes [14, 15, 24–28]. In these works, the modern two-mixing-angle scheme for the $\eta$ and $\eta'$ mesons, which was first advocated in Refs. [4, 5], was employed to fit various experimental data. The common methodology in these works is that the two-mixing-angle pattern for the $\eta$ and $\eta'$ is simply adopted to perform the phenomenological discussion and the mixing parameters are then directly determined from data. This is a bottom-up approach to address the $\eta$-$\eta'$ mixing problem and it is quite useful for the phenomenological analysis. Contrary to the bottom-up method, it is also very interesting to study the $\eta$-$\eta'$ mixing from a top-down approach in which one first constructs the relevant $\chi$PT Lagrangian and then calculates the $\eta$-$\eta'$ mixing pattern and parameters in terms of the LECs. In this case, one can predict the $\eta$-$\eta'$ mixing parameters once the values of the unknown LECs are given. The present work belongs to the latter category of top-down approaches.

Though the singlet $\eta_0$ meson, which is the main component of the physical $\eta'$ state, is not a pNGB due to the strong $U_A(1)$ anomaly, it can be formally introduced into $\chi$PT from the large-$N_C$ point of view. The argument is that the quark loop induced $U_A(1)$ anomaly, which is responsible for the large mass of the singlet $\eta_0$, is $1/N_C$ suppressed and hence the $\eta_0$ becomes the ninth pNGB in the large $N_C$ limit [29]. Based on this argument, the leading-order (LO) effective Lagrangian for $U(3)$ $\chi$PT, which simultaneously includes the pNGB octet $\pi, K, \eta_8$ and the singlet $\eta_0$ as dynamical fields, was formulated in Ref. [3]. Later on, a full $O(p^4)$ $U(3)$ chiral Lagrangian was constructed in Ref. [7] and the discussion on the $O(p^6)$ unitary group chiral Lagrangian has been very recently completed in Ref. [30]. Subtle problems about the choice of suitable variables for the higher order $U(3)$ $\chi$PT Lagrangian in the large $N_C$ framework were analyzed in Ref. [6].

The standard power counting employed in $SU(2)$ and $SU(3)$ $\chi$PT in powers of the external momenta and quark masses [1], is not valid any more in $U(3)$ $\chi$PT, due to the appearance of the large $\eta_0$ mass. However, since the singlet $\eta_0$ mass squared behaves like $1/N_C$ in large $N_C$ limit, the $\eta_0$ mass can be harmonized with the other two expansion parameters if one assigns the same counting to $1/N_C$, the squared momenta $p^2$ and the light quark masses $m_q$. As a result of this, in order to have a systematic power counting, the combined expansions on momentum, light quark masses and $1/N_C$ are mandatory in $U(3)$ $\chi$PT [6, 7]. We will work in this combined expansion in our study and denote it as $\delta$ expansion throughout the paper, where $O(\delta) \sim O(p^2) \sim O(m_q) \sim O(1/N_C)$. This counting rule is different from the one proposed in Ref. [31], where the $\eta_0$ mass is counted as $O(1)$ and the infrared regularization method is employed to handle the chiral loops.

Some recent works in Refs. [8, 32–35] have addressed the $\eta$-$\eta'$ mixing in the chiral framework up to next-leading order (NLO). As an improvement, we will perform the systematic study of the $\eta$-$\eta'$ mixing in the $\delta$-expansion scheme up to next-to-next-to-leading order (NNLO) and take into account the very recent lattice simulation data, which are not considered in the previous works [8, 32–35]. In addition, we also simultaneously analyze the $m_s$ dependences of other physical quantities from lattice simulations, such as the axial $\pi, K$ decay constants and the mass ratio of the strange and up/down quarks, in order to further constrain the $\chi$PT LECs.

This article is organized as follows. In Sect. II, we introduce the theoretical framework and calculate the relevant physical quantities. In Sect. III, the phenomenological discussions will be presented. Conclusions will be given in Sect. IV. Further details about the calculations up to NNLO are relegated to App. A.
II. THEORETICAL FRAMEWORK

A. Relevant chiral Lagrangian

At leading order in the $\delta$ expansion, i.e. $\mathcal{O}(\delta^0)$, the $U(3)\chi$PT Lagrangian consists of three operators

$$\mathcal{L}^{(\delta^0)} = \frac{F^2}{4} (u_\mu u^\mu) + \frac{F^2}{4} (\chi^+) + \frac{F^2}{12} M_0^2 X^2,$$

where the chiral building blocks are defined as [1, 6–8]

$$\begin{align*}
U &= U^2 = e^{i\frac{2K}{3}}, \\
u_\mu &= i u^\dagger D_\mu u^\mu, \\
D_\mu U &= \delta_\mu U - i(v_\mu + a_\mu)U + iU(v_\mu - a_\mu),
\end{align*}$$

with the pNGB octet-singlet matrix

$$\Phi = \begin{pmatrix}
\frac{1}{\sqrt{2}} \pi^0 + \frac{1}{\sqrt{6}} \eta_8 + \frac{1}{\sqrt{3}} \eta_0 \\
\pi^- \\
K^- \\
\frac{1}{\sqrt{2}} \pi^0 + \frac{1}{\sqrt{6}} \eta_8 + \frac{1}{\sqrt{3}} \eta_0 \\
K^0 \\
\frac{-2}{\sqrt{6}} \eta_8 + \frac{1}{\sqrt{3}} \eta_0
\end{pmatrix},$$

and $s, p, v, m$, being the external scalar, pseudoscalar, vector and axial-vector sources, respectively. The coupling $F$ appearing in Eqs. (1) and (2) corresponds to the pNGB axial decay constant in the large $N_C$ and chiral limits. The light quark masses are introduced by setting $(s + ip) = \text{diag}\{m, \hat{m}, m_s\}$, being $\hat{m}$ the averaged up and down quark masses and $m_s$ of the strange quark.

Notice the structure of the LO Lagrangian in Eq. (1): the first operator is of $\mathcal{O}(N_C, p^2)$ type, the second one corresponds to the type of $\mathcal{O}(N_C, m_s)$ and the last one stems from the QCD $U_A(1)$ anomaly and is of $\mathcal{O}(N_{C}^0, p^3)$ type, where $U$ is counted as $\mathcal{O}(1), F^2 \sim \mathcal{O}(N_C)$ and $M_0^2 \sim \mathcal{O}(N_{C}^{-1})$ in the classification $\mathcal{O}(N_{C}^1, p^2), M_0$ of the EFT Lagrangian operators in Eq. (1). In the following, we will denote the chiral expansions in powers of squared momenta $p^2$ and quark masses $m_q$ simply as a generic expansion in $p^2$.

The NLO $U(3)$ chiral Lagrangian, i.e., $\mathcal{O}(\delta)$, contains $\mathcal{O}(N_C, p^4)$ and $\mathcal{O}(N_{C}^0, p^2)$ operators. The relevant ones in our work read [6]

$$\mathcal{L}^{(\delta)} = L_5(u^\mu u_\mu \chi^+) + \frac{L_6}{2} (\chi^+ \chi^+ + \chi^- \chi^-) + \frac{F^2}{12} L_1 \partial_\mu N \partial_\mu X - \frac{F^2}{12} L_2 X \partial_\mu (\chi^-),$$

with the dimensionless LECs’ scaling like $L_5, L_6 \sim \mathcal{O}(N_C)$ and $L_1, L_2 \sim \mathcal{O}(N_{C}^{-1})$.

At NNLO, i.e. $\mathcal{O}(\delta^2)$, there are three types of operators: $\mathcal{O}(N_{C}^{-1}, p^2), \mathcal{O}(N_{C}^0, p^4)$ and $\mathcal{O}(N_C, p^6)$. Their explicit forms read [7, 36]

$$\begin{align*}
\mathcal{L}^{(\delta^2)} &= \frac{F^2}{4} v_2^{(2)} X^2 \chi^+ + L_4(u^\mu u_\mu) \chi^+ + L_6(\chi^+ \chi^+) + L_7(\chi^- \chi^-) + L_8(u_\mu u_\mu) \chi^+ + L_{25} X \chi^- \\
&+ C_{12}(h_\mu h_\mu \chi^+) + C_{14}(u^\mu u_\mu \chi^+) + C_{17}(u_\mu u_\mu \chi^+) \\
&+ C_{19}(\chi^+ \chi^- \chi^+) + C_{31}(\chi^+ \chi^- \chi^-),
\end{align*}$$

where the first line corresponds to the $\mathcal{O}(N_{C}^{-1}, p^2)$ type, the second line is of the $\mathcal{O}(N_{C}^0, p^4)$ type and the last two lines are of the $\mathcal{O}(N_C, p^6)$ type. The LECs carry the scalings $v_2^{(2)} \sim \mathcal{O}(N_{C}^{-2}), L_4, L_6, L_7, L_18, L_{25} \sim \mathcal{O}(N_{C}^0)$ and $C_{12}, C_{14}, C_{17}, C_{19}, C_{31} \sim \mathcal{O}(N_C)$. Notice that we have only shown the operators at different $\delta$ orders in Eqs. (1), (4) and (5) that are pertinent to our present study, not aiming to give the complete sets of operators. The conventions to label the LO, NLO and NNLO operators in Eqs. (1), (4) and (5) follow closely the notations in Refs. [6, 7, 36]. Unless it is explicitly stated, the LECs will correspond to the $U(3)$ $\chi$PT and must not be confused with those in $SU(3)$ $\chi$PT. The matching between these two EFTs can be found in Ref. [6]. The terms $L_j$ are denoted as $\beta_j$ in Ref. [7, 8].

Comparing the $U(3)$ and $SU(3)$ theories one can observe that some terms have been reshuffled in the $\delta$ expansion of the $U(3)$ Lagrangian. For example, the $L_{i=1,2,3,4,5,6,7,8}$ terms are NLO in $SU(3)$ $\chi$PT, but they are now split into NLO and NNLO in the $\delta$ expansion (see Eqs. (4) and (5)). We have several additional new operators, namely the last one in Eq. (1), the $\Lambda_{i=1,2}$ in Eq. (4) and the $v_2^{(2)}$, $L_{18}, L_{25}$ terms in Eq. (5), that are absent in the $SU(3)$ $\chi$PT case. Finally, the chiral loops start contributing at NNLO in the $\delta$ expansion, while they appear at NLO in the conventional $SU(3)$ case.
Next we calculate the $\eta$-$\eta'$ mixing order by order in the $\delta$ expansion. In literature, there are two bases to address the $\eta$-$\eta'$ mixing, namely the singlet-octet basis with $\eta_0$ and $\eta_8$, and the quark-flavor basis with $\eta_q$ and $\eta_s$. The relations between fields in these two bases are

$$
\begin{pmatrix}
\eta_8 \\
\eta_0
\end{pmatrix} = \begin{pmatrix}
\sqrt{\frac{2}{3}} & \sqrt{\frac{1}{3}} \\
\sqrt{\frac{1}{3}} & \sqrt{\frac{2}{3}}
\end{pmatrix}
\begin{pmatrix}
\eta_q \\
\eta_s
\end{pmatrix}.
$$

(6)

In the large-$N_C$ limit where the $U(1)_A$ anomaly is absent, $\eta_q$ and $\eta_s$ are the mass eigenstates and they are generated by the axial-vector currents with the quark flavors $q\bar{q} = (u\bar{u} + d\bar{d})/\sqrt{2}$ and $s\bar{s}$, respectively. The two bases are related to each other through an orthogonal transformation and provide an equivalent description for the $\eta$-$\eta'$ mixing.

As noticed in Refs. [35, 37], when doing the loop calculations with $\eta$ and $\eta'$, it is rather cumbersome to work with the $\eta_0$ and $\eta_8$ states. The reason is that at leading order the Lagrangian in Eq. (1) gives the mixing between $\eta_0$ and $\eta_8$, and the mixing strength is proportional to $m_\Delta^2 - m_\pi^2$, which in the $\delta$ expansion is formally counted as the same order as the diagonal terms in the mass matrix for $\eta_0$ and $\eta_8$. As a result, the insertion of the $\eta_0$-$\eta_8$ mixing in the chiral loops will not increase the $\delta$ order of the loop diagrams. This makes the loop calculation technically much more complicated, as one needs to consider the arbitrary insertions of the $\eta_0$-$\eta_8$ mixing in the chiral loop diagrams. Nevertheless, Refs. [35, 37] provide a simple recipe to handle this problem by expressing the Lagrangian in terms of the $\bar{\eta}$ and $\bar{\eta}'$ states which result from the diagonalization of $\eta_0$ and $\eta_8$ at leading order in $\delta$. The main difference is that the mixing between $\bar{\eta}$ and $\bar{\eta}'$ is now at least a NLO effect in $\delta$, while the $\eta_0$-$\eta_8$ mixing was appearing at LO. The relation between the LO mass eigenstates $\bar{\eta}$ and $\bar{\eta}'$ and the singlet-octet basis is given by the mixing angle $\theta$:

$$
\begin{pmatrix}
\bar{\eta} \\
\bar{\eta}'
\end{pmatrix} = \begin{pmatrix}
c_\theta & -s_\theta \\
s_\theta & c_\theta
\end{pmatrix}
\begin{pmatrix}
\eta_8 \\
\eta_0
\end{pmatrix},
$$

(7)

with $c_\theta = \cos \theta$ and $s_\theta = \sin \theta$. The LO mixing angle $\theta$ and masses of $\bar{\eta}$ and $\bar{\eta}'$ are given by the leading order Lagrangian $\mathcal{L}^{(\delta^0)}$ in Eq. (1) (see e.g. Ref. [35]):

$$
m_\bar{\eta}^2 = \frac{M_0^2}{2} + \frac{m_K^2}{2} - \sqrt{\frac{M_0^4}{4} - \frac{4M_0^2\Delta^2}{3} + 4\Delta^4},
$$

(8)

$$
m_\bar{\eta}'^2 = \frac{M_0^2}{2} + \frac{m_K^2}{2} + \sqrt{\frac{M_0^4}{4} - \frac{4M_0^2\Delta^2}{3} + 4\Delta^4},
$$

(9)

$$
sin \theta = -\left(\sqrt{1 + \frac{(3M_0^2 - 2\Delta^2 + \sqrt{9M_0^4 - 12M_0^2\Delta^2 + 36\Delta^4})^2}{32\Delta^4}}\right)^{-1},
$$

(10)

with $\Delta^2 = m_K^2 - m_\pi^2$. Here $m_K$ and $m_\pi$ denote the LO kaon and pion masses, respectively.

When higher order corrections are taken into account, the LO diagonalized $\bar{\eta}$ and $\bar{\eta}'$ will get mixed again. Up to the NNLO, a general parametrization of the bilinear terms involving the $\bar{\eta}$ and $\bar{\eta}'$ states can be written as

$$
\mathcal{L} = \frac{\delta_1}{2} \partial_\mu \partial_\nu \bar{\eta} \partial^\mu \partial^\nu \bar{\eta} + \frac{\delta_2}{2} \partial_\mu \partial_\nu \bar{\eta}' \partial^\mu \partial^\nu \bar{\eta}' + \frac{\delta_3}{2} \partial_\mu \partial_\nu \bar{\eta} \partial^\mu \partial^\nu \bar{\eta}' + \frac{\delta_4}{2} \partial_\mu \partial_\nu \bar{\eta}' \partial^\mu \partial^\nu \bar{\eta}
$$

$$
+ \frac{1 + \delta_5}{2} \partial_\mu \bar{\eta} \partial^\mu \bar{\eta}' + \frac{1 + \delta_6}{2} \partial_\mu \bar{\eta}' \partial^\mu \bar{\eta} + \delta_7 \partial_\mu \bar{\eta} \partial^\mu \bar{\eta}' + \delta_8 \partial_\mu \bar{\eta}' \partial^\mu \bar{\eta}
$$

$$
+ m_\bar{\eta}^2 + \frac{\delta_{m_\bar{\eta}}}{2} \bar{\eta} \bar{\eta}' - (\frac{m_\bar{\eta}'}{2} + \frac{\delta_{m_\bar{\eta}'}}{2}) \bar{\eta}' \bar{\eta} - \frac{\delta_{m_\bar{\eta}''}}{2} \bar{\eta} \bar{\eta}' - \frac{\delta_{m_\bar{\eta}''}}{2} \bar{\eta}' \bar{\eta}
$$

(11)

where the $\delta_j$'s contain the NLO and NNLO corrections. Here these operators must be understood as the terms of the effective action that provide the pseudoscalar meson self-energies. The higher-derivative terms $\delta_j = 1,2,3$ in the first line of Eq. (11) are exclusively contributed by the $O(p^6)$ operator $C_{12}$ in Eq. (5), which belongs to the NNLO Lagrangian. The remaining $\delta_j$'s receive contributions from the NLO operators in Eq. (4), the NNLO ones in Eq. (5) and the one-loop diagrams, which contribute at NNLO. Their explicit expressions can be found in App. A.

At leading order, there is only the mass mixing term from Eq. (1) whereas at NLO and NNLO one has to deal in addition with the kinematic mixing terms in Eq. (11), apart from the mass mixing. The physical states of $\eta$ and
\( \eta' \) can be obtained from the perturbative-expansion (\( \delta \)-expansion) in three steps: as a first step, we eliminate the higher-derivative terms through the field redefinitions of \( \eta \) and \( \bar{\eta} \); then we transform and rescale the fields resulting from the first step in order to write the kinematic terms in the canonical form; after the preceding two steps, there is only the mass mixing term left, which is straightforward to handle.

In the first step, we make the following field redefinitions for the \( \eta \) and \( \bar{\eta} \) states

\[
\eta \to \eta + \alpha_1 \Box \eta + \alpha_2 \Box \bar{\eta}, \quad \bar{\eta} \to \bar{\eta} + \alpha_2 \Box \eta + \alpha_3 \Box \bar{\eta},
\]

(12)

with the d’Alembert operator \( \Box \equiv \partial_\mu \partial^\mu \). After some algebra manipulations, it is straightforward to obtain

\[
\alpha_1 = -\frac{\delta_1}{2}, \quad \alpha_2 = -\frac{\delta_3}{2}, \quad \alpha_3 = -\frac{\delta_2}{2},
\]

(13)

so that the three higher-derivative terms in Eq. (11) will be eliminated. Notice that the \( \alpha_{1,2,3} \) are NNLO, i.e., \( \mathcal{O}(\delta^2) \).

Substituting the field redefinitions from Eq. (12) into the general mixing structure in Eq. (11) and keeping the terms up to NNLO, the resulting bilinear Lagrangian reads

\[
\mathcal{L} = \frac{1 + \delta_\eta + m_\eta^2 \delta_1}{2} \partial_\mu \partial^\mu \eta + \frac{1 + \delta_{\bar{\eta}} + m_{\bar{\eta}}^2 \delta_2}{2} \partial_\mu \bar{\eta} \partial^\mu \bar{\eta} + \left[ \delta_k + \frac{\delta_3}{2} (m_\eta^2 + m_{\bar{\eta}}^2) \right] \partial_\mu \eta \partial^\mu \bar{\eta}
\]

\[
- \frac{m_\eta^2 + \delta m_2}{2} \eta \eta - \frac{m_{\bar{\eta}}^2 + \delta m_2}{2} \bar{\eta} \bar{\eta} - \delta m_2 \eta \bar{\eta}.
\]

(14)

In the second step, we need to eliminate the kinematic mixing term in Eq. (14), and then to rescale the fields to have them in the canonical forms. This can be done perturbatively. In the final step, we take care of the mass mixing term. The last two steps can be achieved through the following field transformations

\[
\begin{pmatrix} \eta \\ \eta' \end{pmatrix} = \begin{pmatrix} \cos \theta_\delta & -\sin \theta_\delta \\ \sin \theta_\delta & \cos \theta_\delta \end{pmatrix} \begin{pmatrix} 1 + \delta_A & \delta_B \\ \delta_B & 1 + \delta_C \end{pmatrix} \begin{pmatrix} \bar{\eta} \\ \bar{\eta} \end{pmatrix},
\]

(15)

with \( \eta, \eta' \) the physical states and

\[
\begin{align*}
\delta_A &= \frac{\delta_\eta}{2} + \frac{m_\eta^2 \delta_1}{2} - \frac{\delta_{\eta, \text{NLO}}^2}{8} - \frac{\delta_{\bar{\eta}, \text{NLO}}^2}{8}, \\
\delta_B &= \frac{\delta_k}{2} + \frac{\delta_3}{4} (m_\eta^2 + m_{\bar{\eta}}^2) - \frac{\delta_{\eta, \text{NLO}} \delta_{\bar{\eta}, \text{NLO}}}{8} - \frac{\delta_{\eta, \text{NLO}} \delta_{\bar{\eta}, \text{NLO}}}{8}, \\
\delta_C &= \frac{\delta_{\bar{\eta}}}{2} + \frac{m_{\bar{\eta}}^2 \delta_2}{2} - \frac{\delta_{\eta, \text{NLO}}^2}{8} - \frac{\delta_{\bar{\eta}, \text{NLO}}^2}{8},
\end{align*}
\]

(16)

where \( \delta_{\eta, \text{NLO}}, \delta_{\bar{\eta}, \text{NLO}}, \delta_{\bar{\eta}, \text{NLO}} \) stand for the NLO parts of the three quantities respectively. We point out that \( \delta_\eta, \delta_{\bar{\eta}}, \delta_k \) receive both NLO and NNLO contributions, while \( \delta_1, \delta_2, \delta_3 \) are only contributed by the NNLO effect, which is the \( C_{12} \) operator in Eq. (5). Comparing with the NLO results in Eq. (15) from our previous paper [35], we have generalized the expression to the NNLO case in the present Eq. (15).

In the practical calculation, it is more often to use the inverse of the relations in Eq. (15), where the perturbative expansion leads to

\[
\begin{pmatrix} \bar{\eta} \\ \bar{\eta}' \end{pmatrix} = \begin{pmatrix} 1 + \delta_A' & \delta_B' \\ \delta_B' & 1 + \delta_C' \end{pmatrix} \begin{pmatrix} \cos \theta_\delta & \sin \theta_\delta \\ -\sin \theta_\delta & \cos \theta_\delta \end{pmatrix} \begin{pmatrix} \eta \\ \eta' \end{pmatrix},
\]

(17)

with

\[
\begin{align*}
\delta_A' &= \frac{\delta_\eta}{2} - \frac{m_\eta^2 \delta_1}{2} + \frac{3 \delta_{\eta, \text{NLO}}^2}{8} + \frac{3 \delta_{\bar{\eta}, \text{NLO}}^2}{8}, \\
\delta_B' &= -\frac{\delta_k}{2} - \frac{\delta_3}{4} (m_\eta^2 + m_{\bar{\eta}}^2) + \frac{3 \delta_{\eta, \text{NLO}} \delta_{\bar{\eta}, \text{NLO}}}{8} + \frac{3 \delta_{\eta, \text{NLO}} \delta_{\bar{\eta}, \text{NLO}}}{8}, \\
\delta_C' &= -\frac{\delta_{\bar{\eta}}}{2} - \frac{m_{\bar{\eta}}^2 \delta_2}{2} + \frac{3 \delta_{\eta, \text{NLO}}^2}{8} + \frac{3 \delta_{\bar{\eta}, \text{NLO}}^2}{8},
\end{align*}
\]

(18)

The \( \theta_\delta \) appearing in Eqs. (15) and (17) is determined through

\[
\tan \theta_\delta = \frac{\delta \eta}{m_\eta^2 - m_{\bar{\eta}}^2},
\]

(19)
Combining Eqs. (7) and (15), it is straightforward to derive the relations between the four parameters in the two-mixing-angle scheme in Eq. (21) and the \( \chi \)PT LECs:

\[
\begin{align*}
F_8^2 &= F^2 \left\{ \cos(\theta + \theta_5) + \delta_B \sin(\theta - \theta_5) + \delta_A \cos \theta \cos \theta_5 - \delta_C \sin \theta \sin \theta_5 \right\}^2 \\
&+ \left\{ \sin(\theta + \theta_5) + \delta_B \cos(\theta - \theta_5) + \delta_A \cos \theta \sin \theta_5 + \delta_C \sin \theta \sin \theta_5 \right\}^2, \\
F_0^2 &= F^2 \left\{ -\sin(\theta + \theta_5) + \delta_B \cos(\theta - \theta_5) - \delta_A \sin \theta \cos \theta_5 - \delta_C \cos \theta \sin \theta_5 \right\}^2 \\
&+ \left\{ \cos(\theta + \theta_5) - \delta_B \sin(\theta - \theta_5) - \delta_A \sin \theta \sin \theta_5 + \delta_C \cos \theta \cos \theta_5 \right\}^2, \\
\tan \theta_8 &= \frac{\sin(\theta + \theta_5) + \delta_B \cos(\theta - \theta_5) + \delta_A \cos \theta \sin \theta_5 + \delta_C \sin \theta \cos \theta_5}{\cos(\theta + \theta_5) + \delta_B \sin(\theta - \theta_5) + \delta_A \cos \theta \cos \theta_5 - \delta_C \sin \theta \sin \theta_5}, \\
\tan \theta_0 &= \frac{-\sin(\theta + \theta_5) + \delta_B \cos(\theta - \theta_5) - \delta_A \sin \theta \cos \theta_5 - \delta_C \cos \theta \sin \theta_5}{\cos(\theta + \theta_5) - \delta_B \sin(\theta - \theta_5) - \delta_A \sin \theta \sin \theta_5 + \delta_C \cos \theta \cos \theta_5},
\end{align*}
\]

where the \( \chi \)PT LECs are implicitly included in \( \theta, \theta_5, \delta_A, \delta_B \) and \( \delta_C \). Since \( \theta_5, \delta_A, \delta_B, \delta_C \sim O(\delta) \) or \( O(\delta^2) \), at LO one has \( F_8 = F_0 = F \) and one mixing-angle \( \theta_8 = \theta_0 = \theta \).

The relations between the physical \( \eta, \eta' \) states and the quark-flavor basis is commonly parametrized as

\[
\begin{pmatrix}
\eta \\
\eta'
\end{pmatrix} = \frac{1}{F} \begin{pmatrix}
F_q \cos \phi_q & -F_q \sin \phi_q \\
F_q \sin \phi_q & F_q \cos \phi_q
\end{pmatrix} \begin{pmatrix}
\eta_q \\
\eta_s
\end{pmatrix}.
\]
Combining Eqs. (6), (7) and (15), it is straightforward to obtain the parameters in Eq. (23):

\[
F_q^2 = \frac{2F_0^2 + F_8^2 - 2\sqrt{2}F_0 F_8 \sin(\theta_0 - \theta_s)}{3},
\]

\[
F_s^2 = \frac{F_0^2 + 2F_8^2 + 2\sqrt{2}F_0 F_8 \sin(\theta_0 - \theta_s)}{3},
\]

\[
\tan \phi_q = \frac{\sqrt{2}F_0 \cos \theta_8 + F_0 \sin \theta_0}{\sqrt{2}F_0 \sin \theta_0 - F_8 \cos \theta_s},
\]

\[
\tan \phi_s = \frac{\sqrt{2}F_0 \cos \theta_0 + F_8 \sin \theta_8}{\sqrt{2}F_8 \sin \theta_8 - F_0 \cos \theta_0}
\]

where at LO in the \(\delta\)-expansion one has \(F_q = F_s = F\) and \(\phi_q = \phi_s = \theta_{3\mu} - \theta\), with the ideal mixing \(\theta_{3\mu} = -\arcsin \sqrt{2/3}\).

### C. Insights into previous studies of the \(\eta\)-\(\eta'\) mixing

In the previous subsection we have performed the full computation of the mixing up to NNLO in the \(\delta\) expansion. It is interesting to make a brief summary of the assumptions made in previous works, where plenty of mixing formalisms have been proposed to address the \(\eta\)-\(\eta'\) system [8, 24, 32–35, 38, 39]. In Ref. [38], only the lowest order in the quark masses and \(1/N_C\), i.e. the LO contributions in the \(\delta\) expansion, were taken into account. Even though it provided a reasonable first approximation, it failed to give an accurate description of the experimentally observed mass ratio \(m_{\eta'}^2/m_{\eta}^2\). The \(O(p^2)\) contributions were studied up to NLO in \(1/N_C\) in Ref. [39] (including the terms in Eq. (1) and \(A_1\) and \(A_2\) in Eq. (4)), perfectly explaining the experimental value of \(m_{\eta'}^2/m_{\eta}^2\). However, it turned out to be inadequate to give a proper value for the \(\eta\)-\(\eta'\) mixing angle. On the other hand, the authors in Refs. [32–34] went up to NLO in the \(p^2\) expansion but keeping just the LO in \(1/N_C\) (including the terms in Eq. (1) and \(L_5\) and \(L_8\) in Eq. (4)). Both the \(\eta\)-\(\eta'\) mixing angle and the ratio \(F_K/F_\pi\) were qualitatively reproduced in this case. The full set of NLO contributions in the \(\delta\)-expansion (i.e., the effects up to NLO both in \(1/N_C\) and \(p^2\)) was analyzed in Ref. [8], together with the mixing angles and the \(\pi, K, \eta\) and \(\eta'\) axial-vector decay constants. In Ref. [35], the contributions from the tree-level resonance exchanges and partial NNLO effects, e.g. the loop diagrams, were considered for the masses of \(\eta\) and \(\eta'\). In this work, we generalize the discussions up to the full NNLO study in the \(\delta\)-expansion and confront our theoretical expressions with the very recent lattice simulation data and the phenomenological inputs from the two-mixing-angle scheme.

Reference [24] introduced a quark-model inspired approach to the \(\eta\)-\(\eta'\) mixing, which is commonly referred as the FKS formalism and used in many phenomenological analyses [40]. The essence of the FKS formalism is the assumption that the axial decay constants in the quark-flavor basis takes the same mixing pattern as the states

\[
\begin{pmatrix}
F_\eta^q \\
F_\eta^s
\end{pmatrix} = \begin{pmatrix}
\cos \phi - \sin \phi \\
\sin \phi & \cos \phi
\end{pmatrix} \begin{pmatrix}
F_q \\
0
\end{pmatrix} \begin{pmatrix}
0 \\
F_s
\end{pmatrix},
\]

where the decay constants are defined as the matrix elements of the axial currents

\[
\langle 0|A_\mu(0)|P(k)\rangle = i\sqrt{2}F_\mu k_\mu, \quad (a = q, s; P = \eta, \eta'),
\]

\[
A_\mu^q = \frac{1}{\sqrt{2}}(\bar{u}\gamma_\mu \gamma_5 u + \bar{d}\gamma_\mu \gamma_5 d), \quad A_\mu^s = \bar{s}\gamma_\mu \gamma_5 s.
\]

From another point of view, the pattern of Eq. (25) employed in the FKS formalism relies on the assumption that there is no mixing between the decay constants of the flavor states \(\eta_q\) and \(\eta_s\). In the \(\chi PT\) framework, the physical masses and decay constants can be obtained from the bilinear parts of nonet fields in the effective action with the correlation function of two axial currents. Since the correlation function is the second derivative with respect to the axial-vector external source \(a_\mu\) and \(a_\mu\) always appears in the Lagrangian together with the partial derivative \(\partial_\mu\) as shown in Eq. (2), the absence of the mixing for the \(\eta_q\) and \(\eta_s\) decay constants in Eq. (25) implies that there are no kinematic mixing terms for the quark-flavor states \(\eta_q\) and \(\eta_s\) in the FKS formalism. In fact, the assumption in Ref. [32] is in accord with the FKS formalism. This can be simply demonstrated by expanding the chiral operators considered in Refs. [32, 33], i.e. those in Eq. (1) and \(L_5, L_8\) \(^1\) in Eq. (4), up to quadratic terms in \(\eta_q\) and \(\eta_s\). No

\(^1\) Our \(L_5\) and \(L_8\) operators correspond to the \(A_2\) and \(A_1\) terms in Refs. [32–34], respectively. The \(\Lambda\) term in the previous references
kinematic mixing terms for the $\eta_8$ and $\eta_s$ fields result from these chiral operators. This also confirms the finding in Ref. [34] that only when the NLO of $1/N_C$ operator is excluded the FKS formalism is recovered with their chiral Lagrangian calculations.

Since general terms up to NNLO in $\delta$ expansion are kept in our discussion, unlike in the previous works [8, 32–35, 38, 39] where different assumptions, such as the preference of the higher order $p^2$ and $1/N_C$ effects, are made, it is important and interesting for us to justify these assumptions in later discussions.

D. Masses and decay constants of pion and kaon up to NNLO in $\delta$ expansion

The NLO expression of the pion decay constant in the $\delta$ expansion reads

$$F_\pi = F\left(1 + 4L_5 \frac{m_\pi^2}{F^2}\right), \quad (27)$$

or, up to the precision considered, one can also use the physical $F_\pi$ in the expression inside brackets,

$$F_\pi = F\left(1 + 4L_5 \frac{m_\pi^2}{F_\pi^2}\right). \quad (28)$$

The differences between Eqs. (27) and (28) are NNLO effects. We mention that at a given order there is always ambiguity in choosing the renormalized quantities in the higher order expressions. In contrast, there is formally no ambiguity in the expressions in terms of the quantity $F$, which is the pNGB axial decay constant in the chiral and large $N_C$ limits. For example, if we limit our analysis up to NLO, formally, it is equally good to use $F_\pi$ or $F_K$ in the denominators of the NLO part in Eq. (28), since the difference is beyond the NLO precision. A typical solution in the chiral study is to express the quantities, such as $m_\pi$, $F_\pi$, $m_K$, $F_K$, in terms of the renormalized $F_\pi$ in the higher order corrections, as done in the two-loop calculations in $SU(3)$ $\chi$PT [41]. We follow this rule throughout the current work to estimate the uncertainty due to the truncation of the $\delta$ expansion when one works at a given order in perturbation theory. Notice that the mass $m_\pi$ in the above equations stands for its LO expression, which is linearly dependent on the up/down-quark mass, and the expressions beyond the leading order will be given below.

Similarly up to NNLO, we can either use $F$ or $F_\pi$ in the NLO and NNLO expressions for other quantities such as $F_K$ and the $\delta_i$’s in Eq. (11). In the NNLO expressions, the difference between using $F$ or $F_\pi$ in the denominators is a next-to-next-to-next-to-leading order effect ($N^3$LO). Since in this work we study lattice simulation data up to pion mass of 500 MeV, the convergence of the chiral series is expected to be much slower than that in the physical case with $m_\pi = 135$ MeV. Therefore it is a priori not trivial to judge whether the two approaches–using $1/F^2$ and $1/F_\pi^2$–are numerically equivalent or the lattice data prefer one of them. Indeed in Ref. [42], it is already noticed that to use $F$ or $F_\pi$ could cause some noticeable effects. We will use the difference between both approaches as an estimate of the truncation error at a given order in $\delta$.

We take the pion decay constant as an example to illustrate the differences of using $F$ and $F_\pi$ in the higher order expressions. Using $F$ in the higher order corrections, its expression reads

$$F_\pi = F\left[1 + 4L_5 \frac{m_\pi^2}{F^2} + 4L_4 \frac{m_\pi^2 + 2m_K^2}{F^2} + (24L_5^2 - 64L_5L_8) \frac{m_\pi^4}{F^4} + (8C_{14} + 8C_{17}) \frac{m_\pi^4}{F^2}\right.\nonumber$$

$$\left. + \frac{A_0(m_\pi^2)}{16\pi^2 F^2} + \frac{A_0(m_K^2)}{32\pi^2 F^2}\right]. \quad (29)$$

The one-point loop function $A_0(m^2)$ is calculated in dimensional regularization within the $\overline{MS} - 1$ scheme [1] and it reads

$$A_0(m^2) = -m^2 \ln \frac{m^2}{\mu^2}, \quad (30)$$

corresponds to our $A_2$ operator in Eq. (4). The $\Lambda$ term, though introduced from the beginning in these references, is dropped in their later discussions, since it is $1/N_C$ suppressed.
with the renormalization scale $\mu$ fixed at 770 MeV throughout. Using Eq. (27) to replace $F$ by $F_\pi$ in the NLO and NNLO corrections, the resulting form is

$$F_\pi = F \left[ 1 + 4 L_5 \frac{m^2_\pi}{F^2_\pi} + 4 L_4 \frac{m^2_\pi + 2m^2_K}{F^2_\pi} + (56L^2_5 - 64L_5L_8) \frac{m^4_\pi}{F^4_\pi} + (8 C_4 + 8 C_{17}) \frac{m^4_\pi}{F^4_\pi} + \frac{A_0(m^2_\pi)}{16\pi^2 F^2_\pi} \right].$$  (31)

In the $\delta$ expansion, the expressions for a physical quantity with $F$ or $F_\pi$ in the higher order chiral corrections differ only for the $L_5 L_{j=5,8}$ and $L_5 A_{j=1,2}$ terms, since the differences by replacing $F$ by $F_\pi$ are originated from the NLO expressions of $F_\pi$ in Eq. (28) and we only retain terms up to NLO in this work. It is clear that the difference between Eqs. (29) and (31) is the $L^2_5$ term. Notice that in the $\delta$ expansion scheme, the terms like $L_5 L_4$ are N^3LO and will be dropped throughout the article.

The corresponding expression for the kaon decay constant when one uses $F$ to express the NLO and NNLO corrections reads

$$F_K = F \left[ 1 + 4 L_5 \frac{m^2_K}{F^2_\pi} + 4 L_4 \frac{m^2_\pi + 2m^2_K}{F^2_\pi} + (24L^2_5 - 64L_5L_8) \frac{m^4_\pi}{F^4_\pi} + 8 C_4 \frac{m^4_\pi - 2m^4_K m^2_\pi + m^4_K}{F^4_\pi} + 8 C_{17} \frac{m^2_\pi(2m^2_K - m^2_\pi)}{F^2_\pi} + \frac{A_0(m^2_\pi)}{128\pi^2 F^2_\pi} + \frac{3 A_0(m^2_K)}{64\pi^2 F^2_\pi} \right].$$  (32)

On the other hand, expressing the NLO and NNLO contributions in terms of $F_\pi$ yields

$$F_K = F \left[ 1 + 4 L_5 \frac{m^2_K - m^2_\pi}{F^2_\pi} + 4 L_4 \frac{m^2_\pi + 2m^2_K}{F^2_\pi} + 8 L_5 \frac{3m^4_\pi + 4m^2_K m^2_\pi - 64L_5L_8 \frac{m^4_\pi}{F^4_\pi}}{F^4_\pi} + 8 C_4 \frac{2m^4_\pi - 2m^4_K m^2_\pi + m^4_K}{F^4_\pi} + 8 C_{17} \frac{m^2_\pi(2m^2_K - m^2_\pi)}{F^2_\pi} \right].$$  (33)

The expanded expression for the ratio of $F_K/F_\pi$ in terms of $F$ up to NNLO in $\delta$ expansion, takes the form

$$\frac{F_K}{F_\pi} = 1 + 4 L_5 \frac{m^2_K - m^2_\pi}{F^2_\pi} + 8 L_5 \frac{3m^4_\pi + 2m^2_K m^2_\pi - 5m^4_\pi}{F^4_\pi} + 64 L_5 L_8 \frac{m^4_\pi - m^4_K}{F^4_\pi} + 16 C_4 \frac{m^4_\pi - 2m^4_K m^2_\pi + m^4_K}{F^4_\pi} + 16 C_{17} \frac{m^2_\pi(2m^2_K - m^2_\pi)}{F^2_\pi} \right].$$  (34)

When expressing the previous result in terms of $F_\pi$, it reads

$$\frac{F_K}{F_\pi} = 1 + 4 L_5 \frac{m^2_K - m^2_\pi}{F^2_\pi} + 8 L_5 \frac{3m^4_\pi + 2m^2_K m^2_\pi - 5m^4_\pi}{F^4_\pi} + 64 L_5 L_8 \frac{m^4_\pi - m^4_K}{F^4_\pi} + 16 C_4 \frac{m^4_\pi - 2m^4_K m^2_\pi + m^4_K}{F^4_\pi} + 16 C_{17} \frac{m^2_\pi(2m^2_K - m^2_\pi)}{F^2_\pi} \right].$$  (35)

which differs from Eq. (34) in the $L^2_5$ term.

The pion squared mass up to NNLO is given by

$$m^2_\pi = \frac{m^2_\pi}{m^2_\pi + m^2_\pi^{NLO} + m^2_\pi^{2NNLO}},$$  (36)
with
\[ m^2_\pi = 2B\hat{m}, \]
\[ m^2_{\pi,NLO} = \frac{8(2L_8 - L_5)m^4_\pi}{F^2}, \]
\[ m^2_{\pi,NNLO} = \frac{8(2L_6 - L_4)m^2_\pi(2m^2_K + m^2_\pi)}{F^2} - \frac{64(L_5^2 - 6L_5L_8 + 8L_8^2)m^6_\pi}{F^4}
- \frac{16(2C_{12} + C_{14} + C_{17} - 3C_{19} - 2C_{31})m^6_\pi}{F^4}
+ \frac{m^2_\pi(c^2_0 - 2\sqrt{2}c_0s_\theta + 2s^2_0)A_0(m^2_\eta)}{96\pi^2 F^2}
+ \frac{m^2_\pi(2c^2_3 + 2\sqrt{2}c_0s_\theta + s^2_0)A_0(m^2_\eta)}{96\pi^2 F^2}
- \frac{m^2_\pi A_0(m^2_\eta)}{32\pi^2 F^2}. \]

When expressing the renormalized \( m_\pi \) in terms of \( F_\pi \), the only differences are the \( L_5L_8 \) and \( L_8^2 \) terms in Eq. (39) and the other parts are the same as in Eq. (36) with the explicit replacement of \( F \) by \( F_\pi \) in Eq. (27). Therefore we only give the different parts for simplicity when expressing in terms of \( F_\pi \) and they read
\[ m^2_{\pi,(F_\pi),L_5L_8,L_8^2} = \frac{128(4L_5L_8 - L_8^2)m^6_\pi}{F^4}. \]

The mass squared for kaon up to NNLO is provided by
\[ m^2_K = m^2_K + m^2_{K,NLO} + m^2_{K,NNLO}, \]
with
\[ m^2_K = B(\hat{m} + m_\pi), \]
\[ m^2_{K,NLO} = \frac{8(2L_8 - L_5)m^4_K}{F^2}, \]
\[ m^2_{K,NNLO} = \frac{8(2L_6 - L_4)m^2_K(2m^2_K + m^2_\pi)}{F^2} - \frac{64(L_5^2 - 6L_5L_8 + 8L_8^2)m^6_K}{F^4}
- \frac{32C_{12}m^6_K}{F^4} + \frac{32C_{31}m^6_\pi}{F^4}
+ \frac{16C_{17}m^2_K m^2_\pi(-2m^2_K + m^2_\pi)}{F^4}
- \frac{16C_{14}m^2_K(2m^4_K - 2m^2_K m^2_\pi + m^4_\pi)}{F^4}
+ \frac{48C_{19}m^2_K(2m^4_K - 2m^2_K m^2_\pi + m^4_\pi)}{F^2}
- \frac{4c^2_3m^2_K + 2\sqrt{2}c_0s_\theta(2m^2_K - m^2_\pi) + (3m^2_\eta + m^2_\eta)s^2_0A_0(m^2_\eta)}{192\pi^2 F^2}
- \frac{4c^2_3m^2_K + 2\sqrt{2}c_0s_\theta(2m^2_K - m^2_\pi) + (3m^2_\eta + m^2_\eta)s^2_0A_0(m^2_\eta)}{192\pi^2 F^2}. \]

When expressing Eq. (41) in terms of \( F_\pi \), the differences are the \( L_5L_8 \) and \( L_8^2 \) terms in Eq. (43) and the new expressions are
\[ m^2_{K,(F_\pi),L_5L_8,L_8^2} = \frac{64L_5^2m^6_K(m^2_K + m^2_\pi)}{F^4} + \frac{128L_5L_8m^4_K(3m^2_K + m^2_\pi)}{F^4}. \]

Notice that the masses of pion and kaon appearing in NLO and NNLO parts in the above equations correspond to the renormalized quantities, instead of their LO expressions. In addition, this gives the quark mass ratio relation \( m_q/m_\pi = \frac{2m^2_K/m^2_\pi - 1}{m^2_K/m^2_\pi} \).

When performing the chiral extrapolation of the lattice data, instead of the renormalized \( m^2_K \) as in the previous equations, it is convenient to use the LO kaon mass squared in the higher order corrections. In this way, we do not need to iteratively solve Eq. (41) in order to give the value of \( m_K \) for a given \( m_\pi \). The result in terms of \( \overline{m}_K \) in the NLO and NNLO expressions becomes
\[ m^2_{K,Lat} = \frac{m^2_K + m^2_{K,-NLO} + m^2_{K,-NNLO}}{F^4}, \]
with

\[
m_{K}^{2, \text{Lat-NLO}} = \frac{8(2L_{5} - L_{5})\overline{m}_{K}^{2}}{F^{2}},
\]

\[
m_{K}^{2, \text{Lat-NNLO}} = \frac{8(2L_{6} - L_{6})\overline{m}_{K}^{2}(2\overline{m}_{K}^{2} + m_{\eta}^{2}) + 64(L_{5}^{2} - 2L_{5}L_{6})\overline{m}_{K}^{4} - 32C_{12}\overline{m}_{K}^{2} + 32C_{31}\overline{m}_{K}^{2}}{F^{4}}
\]

\[
+ 16C_{17}\overline{m}_{K}^{3}(2\overline{m}_{K}^{2} - 2\overline{m}_{K}^{2}m_{\eta}^{2} + m_{\eta}^{4})
\]

\[
+ 48C_{19}\overline{m}_{K}^{2}(2\overline{m}_{K}^{2} - 2\overline{m}_{K}^{2}m_{\eta}^{2} + m_{\eta}^{4})
\]

\[
\frac{[\epsilon_{6}^{2}m_{\eta}^{2} + 2\sqrt{2}c_{9}s_{0}(-2m_{\eta}^{2} + m_{\eta}^{2} - 4m_{K}^{2}m_{\eta}^{2})A_{0}(m_{\eta}^{2})]}{192\pi^{2}F^{2}}
\]

\[
\frac{[-4\epsilon_{6}^{2}\overline{m}_{K}^{2} + 2\sqrt{2}c_{9}s_{0}(2\overline{m}_{K}^{2} - m_{\eta}^{2} + (3m_{\eta}^{2} + m_{\eta}^{2})s_{0}^{2}]A_{0}(m_{\eta}^{2})]}{192\pi^{2}F^{2}}.
\]

When expressing Eq. (46) in terms of \(F_{\pi}\), the differences are the \(L_{5}L_{8}\) and \(L_{5}^{2}\) terms in Eq. (48) and the new expressions are

\[
m_{K}^{2, \text{Lat,}(F_{\pi}),L_{5}L_{8},L_{5}^{2}} = \frac{64L_{5}(L_{5} - 2L_{5})\overline{m}_{K}^{2}(m_{\eta}^{2} - m_{\eta}^{2})}{F_{\pi}^{2}}.
\]

When confronting with the lattice data, we only consider the simulated points with the physical strange-quark mass, i.e. the lattice ensembles that when extrapolating to the physical pion masses lead simultaneously to physical kaon masses. In this case, we can express the LO kaon mass squared as

\[
\overline{m}_{K}^{2} = B(m_{s}^{\text{Phy}} + \tilde{m}) = B(m_{s}^{\text{Phy}} + \tilde{m}^{\text{Phy}}) - B\tilde{m}^{\text{Phy}} + B\tilde{m} = \overline{m}_{K}^{2, \text{Phy}} - \frac{\overline{m}_{\pi}^{2, \text{Phy}}}{2} + \frac{\overline{m}_{\eta}^{2}}{2},
\]

where \(\overline{m}_{\pi}^{2, \text{Phy}}\) and \(\overline{m}_{\eta}^{2, \text{Phy}}\) can be obtained through Eqs. (36) and (41) by substituting the physical masses of \(\pi, K, \eta, \eta', \pi', \) and \(\eta, \eta'\) in the NLO and NNLO expressions. For \(\overline{m}_{\eta}^{2}\), which varies in the lattice simulation, we can extract its value by using Eq. (36). In this case, \(m_{\pi}^{\text{Phy}}\) in Eq. (36) takes the value from lattice simulation, and \(m_{K}, m_{\eta}, m_{\eta'}, m_{\eta''}\), which only appear in the NNLO part, can be approximated by their LO expressions.

III. PHENOMENOLOGICAL DISCUSSIONS

The big challenge in the present general discussions on the \(\eta-\eta'\) mixing is the determination of the unknown LECs in Eqs. (1), (4) and (5). The recent lattice simulations on the light pseudoscalar mesons provide us valuable sources to constrain these free parameters. The considered lattice simulations include the mass dependences of the masses of \(\eta, \eta'\) [9–13] and kaon [16, 17], and the \(\pi, K\) decay constants [16, 17] and their ratios [18]. Moreover, relevant phenomenological results and experimental data will be also included to constrain the LECs.

Since we do not consider isospin violating effects, we will take the values for the physical pion and kaon masses in the isospin limit from Ref. [43], where the corrections from the electromagnetic contributions are removed,

\[
m_{\pi} = 134.8 \text{ MeV}, \quad m_{K} = 494.2 \text{ MeV}.
\]

These values will be used in later chiral extrapolations, while for the physical masses of \(\eta\) and \(\eta'\) and the decay constants of pion and kaon, we will take their world-average values from Ref. [44].

In order to show the results step by step, we present the discussions in the following sections split in three steps: we consider fits performed at leading order, next-to-leading order and next-to-next-to-leading order.

A. Leading-order analyses

At leading order, the \(\eta-\eta'\) mixing is described by one free parameter, namely the singlet \(\eta_{0}\) mass \(M_{0}\) in Eq. (1) and the explicit expressions for the masses and mixing angle are given in Eqs. (8), (9) and (10). At this order, the \(\pi, K\) decay constants are degenerate and given by their chiral and large \(N_{c}\) limits, i.e. \(F_{\pi} = F_{K} = F\). Therefore we shall not take the lattice simulations of the decay constants into account for the LO discussion as they clearly show...
the need of higher order corrections for a suitable description. Also at leading order, \( F \) will not enter the masses and mixing angle, as shown in Eqs. (8), (9) and (10). As a result of this, we do not need to distinguish the two situations with \( F \) or \( F_* \) discussed previously in the expressions of different observables. Apart from the lattice simulation data, we also fit the physical values of the \( \eta \) and \( \eta' \) masses. Nonetheless, fitting the physical masses with the experimental precision at the level of several hundred-thousandth is too ambitious. Since the ultimate goal of the present work is the NNLO study, the ballpark estimate of our theoretical uncertainty, starting from the N\(^3\)LO part, should be around 3\%. This value is obtained from the general rule that each higher order correction in \( \delta \) expansion, either the SU(\( 3 \))-flavor breaking or the 1/\( N_C \) effect, is around 30\%. In fact, the estimated three-percent uncertainty is also similar to the typical error bars reported in many lattice simulations, in the range from 3 \( -\) 10\% [9–13]. Consistently, we assign a 1\% uncertainty to the physical values of \( m_\eta \) and \( m_\eta' \) in the fits.

The value of the singlet mass \( M_0 \) from the LO fit is
\[
M_0 = (835.7 \pm 7.5) \text{ MeV}.
\] (52)

The physical masses for the \( \eta, \eta' \) and their LO mixing angle \( \theta \) from the fit are found to be
\[
m_\eta = (494.4 \pm 1.3) \text{ MeV}, \quad m_\eta' = (969.8 \pm 5.8) \text{ MeV}, \quad \theta = -18.9^\circ \pm 0.3^\circ.
\] (53)

The resulting plots can be seen in Fig. 1. We verify that if the physical masses are excluded in the fit, \( M_0 = 813 \pm 11 \text{ MeV} \) results. If we only include the physical masses and exclude the lattice simulation data in the fit, \( M_0 = 859 \pm 11 \text{ MeV} \) is obtained. These determinations of \( M_0 \) lie within the broad range summarized in Ref. [40] and are quite close with the commonly used values of \( M_0 = 850 \text{ MeV} \) [40]. Taking into account the large uncertainties of the lattice simulation data, especially for \( m_\eta' \), and the concise formalism of the LO mixing, it is impressive that the lattice simulation data can already be qualitatively described with the LO analysis, as shown in Fig. 1. This also indicates that the higher order mixing effects can only give moderate corrections to the masses of \( \eta \) and \( \eta' \).

Nevertheless, in order to describe the lattice data more accurately, specially the \( \eta \) masses, the chiral corrections beyond the leading order are needed. For the physical masses, it has also been shown that the LO description fails to explain the mass ratio of \( \eta \) and \( \eta' \) accurately enough [38]. Therefore it is essential to generalize the discussions to NLO and NNLO in order to achieve a precise description both for lattice simulations and physical data.

### B. Next-to-leading order analyses

At next-to-leading order, in addition to the parameter \( M_0 \) at leading order, there are five additional free parameters: the decay constant \( F \) at chiral and large \( N_C \) limits, and the four NLO LECs \( L_5, L_8, \Lambda_1 \) and \( \Lambda_2 \) in Eq. (4). At this order, as well as at next-to-next-to-leading order, one can rewrite the chiral expansion of the observables in various equivalent ways up to the perturbative order in \( \delta \) under consideration. In the following discussion, we will perform two types of fits: one using \( F \) in the theoretical NLO and NNLO expressions and the other employing \( F_* \), as discussed in Sect. II.D. Since the differences of the theoretical expressions used in the two types of fits are beyond the considered precision, the variances of the outputs from the two fits can be considered as systematic errors from the theoretical models by neglecting higher order contributions. In the following, we will explicitly present the fit results by using \( F \) in the theoretical expressions, which is the most straightforward option, as discussed in Sect. II.D. The outputs of the fits with the theoretical formulas expressed in terms of \( F_* \) will be used to estimate the systematic errors: the difference between the central values of the two types of fits will be used to estimate the truncation uncertainty due to working just up to a given order in the \( \delta \)-expansion, providing the second error for each quantity in the following tables.

In Refs. [32–34], it is argued that at each chiral order, the leading \( N_C \) effects are dominant, or in other words that the \( \Lambda_1 \) and \( \Lambda_2 \) terms are assumed to be much less irrelevant than the \( L_5 \) and \( L_8 \) terms in the NLO \( \delta \) expansion. This assumption has been more or less confirmed when focusing on the masses of \( \eta, \eta' \) and the LO mixing angle at the physical points [32–34]. In Ref. [35], the local higher order LECs were estimated by the tree-level resonance exchanges and it was found that with those LECs \( \Lambda_2 \) seems to be more important than \( \Lambda_1 \) when focusing on the physical masses for \( \eta \) and \( \eta' \). It is interesting to check how these assumptions work when including the lattice simulations and the phenomenological results of the two-mixing-angle parameters, which are not considered in Refs. [32–35]. Different sets of fits to the lattice data and phenomenological inputs from the two-mixing-angle scheme are performed either by fixing \( \Lambda_{i=1,2} \) to zero or releasing their values, in order to reexamine the assumptions. Interestingly we do not find qualitative changes between the fits with fixed \( \Lambda_{i=1,2} = 0 \) and the ones with free values for these parameters. This tells us that indeed the \( \Lambda_1 \) and \( \Lambda_2 \) terms do not significantly improve the fit results, even after taking into account the lattice simulations. Nevertheless, we find that these two terms are quite important to reproduce the phenomenological mixing angles \( \theta_0 \) and \( \theta_5 \) in the fits where \( M_0 \) is fixed at its LO value. If \( M_0 \) is released in the fits we find that including
FIG. 1: The masses of $\eta$ and $\eta'$ from the LO fit. The left most two points correspond to the physical masses. The remaining lattice simulation data are taken from Refs. [12, 13] (ETMC), [11] (UKQCD), [10] (RBC/UKQCD), [9] (HSC), where we only take into account the simulation points with $m_\pi < 500$ MeV. The shade area surrounding each curve stands for the statistical uncertainty from the fit.

$\Lambda_1$ and $\Lambda_2$ improves the descriptions of $m_{\eta'}$ from lattice simulations. Therefore, we will not further discuss fits with $\Lambda_1$ and $\Lambda_2$ set to zero in the following. Instead, we focus on the results given in Table I with all the four NLO LECs in the fits, namely $L_5$, $L_8$, $\Lambda_1$ and $\Lambda_2$ in Eq. (4).

For the parameter $M_0$, we take two strategies to estimate its value in NLO analysis. In one of them we fix $M_0 = 835.7$ MeV from its LO determination (NLOFit-A) and in the other case we free its value for the NLO fit (NLOFit-B). These two NLO fits are given in Table I. The first error bar for each fitted parameter corresponds to the statistical one from the fits and the second error bar is estimated from the variation of the fits between those using $F_\pi$ and $F_\pi$ in the NLO (and later also NNLO) theoretical expressions. From the two fits shown in Table I, one can see that freeing $M_0$ in the fits barely changes the fit quality with respect to the cases when its value is fixed, although there are slight variations in the determinations of $M_0$ and $\Lambda_2$.

Concerning the results of the LECs in Table I, the resulting values for $F$ from the two fits are quite compatible and close to the physical pion decay constant. For $\Lambda_1$ and $\Lambda_2$, their values are poorly known in literature and it is helpful to compare our values with their following estimate for their ranges: we take the LO determination $M_0 = 835.7$ MeV, and we then separately include the $\Lambda_1$ and $\Lambda_2$ terms in the $\eta$-$\eta'$ mixing and vary their values to obtain new results for $m_\eta$ and $m_{\eta'}$ with the physical $m_\pi$. Since the $\Lambda_1$ and $\Lambda_2$ terms are NLO $1/N_C$ effects, it is reasonable to assume that their corrections to $m_\eta^2$ or $m_{\eta'}^2$ should be at most around 30% of the LO results. In this way we can set up conservative and rough estimates for the ranges of $\Lambda_1$ and $\Lambda_2$, which are found to be

$$|\Lambda_1| < 0.4, \quad |\Lambda_2| < 0.7.$$  \hspace{1cm} (54)

The resulting magnitudes of $\Lambda_1$ in our fits are tiny and consistent with zero, as shown in Table I. For the parameter
Table I: Parameters from the NLO fits. The details of the different fits are explained in the text. The column with 835.7* for $M_0$ (in MeV) denotes the fit results where the value of $M_0$ has been fixed to its LO determination. The first error bar for each parameter is directly given by the fits and the second one corresponds to the systematical error due to working up to a given order in $\delta$ and neglecting higher corrections. The way to estimate the systematical error is explained in detail in the text.

| Parameters | Inputs | NLOFit-A | NLOFit-B |
|------------|--------|----------|----------|
| $F_0$ (MeV) | $118.0 \pm 16.5$ | $104.7 \pm 3.0 \pm 0.1$ | $99.7 \pm 3.6 \pm 1.6$ |
| $F_8$ (MeV) | $133.7 \pm 11.1$ | $113.2 \pm 0.3 \pm 4.4$ | $113.5 \pm 0.3 \pm 4.2$ |
| $\theta_0$ (Degree) | $-11.0 \pm 3.0$ | $-7.4 \pm 2.1 \pm 1.1$ | $-10.6 \pm 2.4 \pm 0.1$ |
| $\theta_8$ (Degree) | $-26.7 \pm 5.4$ | $-21.6 \pm 2.2 \pm 3.8$ | $-25.4 \pm 2.6 \pm 2.3$ |
| $m_s/m$   | $27.5 \pm 3.0$ | $22.6 \pm 0.8 \pm 0.6$ | $21.8 \pm 0.6 \pm 1.3$ |
| $F_0$ (MeV) | $106.0 \pm 11.1^*$ | $94.0 \pm 1.9 \pm 1.6$ | $90.5 \pm 2.4 \pm 0.3$ |
| $F_8$ (MeV) | $143.8 \pm 16.5^*$ | $122.2 \pm 1.2 \pm 5.2$ | $120.9 \pm 1.2 \pm 5.5$ |
| $\theta_0$ (Degree) | $34.5 \pm 5.4^*$ | $40.3 \pm 3.1 \pm 3.5$ | $35.0 \pm 3.8 \pm 1.6$ |
| $\theta_8$ (Degree) | $36.0 \pm 4.2^*$ | $39.8 \pm 1.8 \pm 2.1$ | $37.2 \pm 1.8 \pm 1.1$ |

Table II: Outputs from the NLO fits. Notice that $F_0$, $F_8$, $\theta_0$ and $\theta_8$ are not the phenomenological inputs in the fits. They are related to $F_0$, $F_8$, $\theta_0$ and $\theta_8$ through Eq. (24). The phenomenological values for the mixing parameters are taken from Ref. [15]. We have tripled the error bands here in order to make a conservative estimate. The input for $m_s/m$ is taken from the FLAG working group in Ref. [43] and we assign the 10% error bar as done in Ref. [20]. For the error bars of each quantity, we follow the same treatment as in Table I.

C. NLO fits focusing on the masses

In this section, we present another kind of NLO fits by focusing on the masses of $\eta, \eta', K$ and excluding the decay constants $F_\pi, F_K$ and their ratio. This kind of discussion is well motivated, since it is known that the NNLO
FIG. 2: The masses of $\eta$ and $\eta'$ from the NLO and NNLO fits. The left most two points correspond to the physical masses. The remaining lattice simulation data are taken from Refs. [12, 13] (ETMC), [11] (UKQCD), [10] (RBC-UKQCD), [9] (HSC), where we only take into account the points with $m_\pi < 500$ MeV. The shaded area around each curve stands for the statistical error band from the fits.

corrections in $\delta$ counting, such as the LEC $L_4$, are important to simultaneously describe $F_\pi$ and $F_K$ [20–22]. But this LEC is absent in NLO study. We have also provided another independent confirmation on this finding in Fig. 4, where one can see that the decay constants of pion and kaon are poorly reproduced at next-to-leading order in $\delta$ expansion. When only focusing on the $\eta, \eta'$ and $K$ masses and the ratio $m_\pi/\hat{m}$ at next-to-leading order the parameter $F$ cannot be resolved, because it always appears in the form $L_5/F^2$ or $L_8/F^2$. We will fix its value to $F = 90$ MeV, close to the values given in Table I. For the mixing parameters we consider the mixing angles of $\theta_0$ and $\theta_8$, but exclude the constants $F_0$ and $F_8$. This is because $F_0$ and $F_8$ are dependent on the parameter $F$ and should be determined together with $F_\pi$ and $F_K$. For simplicity in later discussion, we call the fits performed in this section as the mass-focusing type throughout.

As in the previous section, we present the fits with $F$ in the denominators of the theoretical expressions (e.g. Eq. (27)) and use the fits with $F_\pi$ to estimate the systematic errors. For each case, we perform the fits either by fixing $M_0$ at its LO determination (NLOFit-C) or by freeing its value (NLOFit-D). The fitted parameters are given in Table III and the $m_\pi/\hat{m}$ ratio and mixing angles are given in Table IV. The resulting figures from NLOFit-C and NLOFit-D are quite similar and we explicitly show one set of them, e.g. NLOFit-D in Figs. 2 and 3 for the $\eta(\prime)$ and kaon masses, respectively.

A significant difference between the results in Table I and the mass-focusing fits in Table III is that much larger statistical error bars are obtained in the latter case, especially for the LECs $L_5$ and $L_8$, as they are constrained by fewer data. Likewise, there are large systematical errors for the values of $L_5$ and $L_8$ in Table III, indicating a larger truncation uncertainty due to higher orders. We do not see a significant improvement when freeing the value of $M_0$ in the fits.
FIG. 3: Kaon mass from the NLO and NNLO fits. The lattice simulation data are taken from RBC and UKQCD [16, 17]. Only the unitary points simulated with the physical strange quark mass are included. The shaded area around each curve stands for the statistical error band from the fits.

| Parameters | Inputs | NLOFit-C | NLOFit-D |
|------------|--------|----------|----------|
| $\chi^2/(d.o.f)$ | 165.6/(44-4) | 165.4/(44-5) |
| $M_0$ (MeV) | 835.7* | 813.2±43.1 |
| $10^4 \times L_5$ | 1.34±0.58±0.69 | 1.51±0.68±0.91 |
| $10^4 \times L_8$ | 0.86±0.28±0.33 | 0.95±0.34±0.45 |
| $A_1$ | -0.04±0.04±0.00 | -0.09±0.10±0.08 |
| $A_2$ | 0.17±0.19±0.25 | 0.18±0.19±0.11 |

TABLE III: Parameters from the mass-focusing NLO fits. The meaning of different notations to label different fits are explained in detail in the text. The coupling $F$ is fixed at 90 MeV in these fits. The meaning of the errors is the same as in Table I.

| Parameters | Inputs | NLOFit-C | NLOFit-D |
|------------|--------|----------|----------|
| $\theta_0$ (Degree) | -11.0 ±3.0 | -10.4±2.4±3.5 | -11.0±2.7±2.1 |
| $\theta_8$ (Degree) | -26.7 ±5.4 | -24.5±2.5±3.6 | -26.7±3.5±7.1 |
| $m_s/\bar{m}$ | 27.5±3.0 | 23.5±0.3±0.1 | 23.4±0.4±0.1 |
| $\theta_3$ (Degree) | 34.5 ±5.4 | 36.3±1.6±0.4 | 34.1±3.6±4.1 |
| $\theta_5$ (Degree) | 36.0 ±4.2 | 37.4±0.9±0.3 | 36.3±1.8±2.1 |

TABLE IV: The outputs from the mass-focusing NLO fits. See Table II for the phenomenological inputs. The meaning of the errors is the same as in Table I.
D. Next-to-next-to-leading order analyses

From the NLO discussions in the previous two sections, we observe that the phenomenological results and the lattice simulations on $\eta$ and $\eta'$ states can be reasonably reproduced. This is an important improvement comparing with the LO study, since at this order we only have the conventional one-mixing-angle formalism. The two-mixing-angle formalism only shows up beyond LO. However, observing $m_K, F_\pi, F_K$ and their ratio in Figs. 3, 4 and 5, it is clear that the NLO analysis is still inadequate. We need to include higher order contributions beyond NLO in order to further improve the descriptions. Moreover, the chiral logarithms predicted by $\chi$PT at one loop start at NNLO in the $\delta$ expansion. Due to their importance in other observables, we consider it is relevant to discuss the impact of these chiral logs.

As in the NLO case, we perform two types of fits, using the NLO and NNLO theoretical expressions given in terms of $F$ and $F_\pi$ for various observables. We explicitly present the fit results with $F$ in the theoretical expressions and use the alternative fits expressed in terms of $F_\pi$ to estimate the systematical errors, due to working up to NNLO in $\delta$ and neglecting higher orders. According to the Lagrangian in Eq. (5), eleven additional unknown LECs appear at NNLO and there will be seventeen parameters in total for the NNLO study. At the present precision of the lattice simulations and phenomenological inputs, it is impossible to obtain sensible and stable fits if we free all of the seventeen parameters. Therefore, we need to take other independent determinations for some of the LECs in order to proceed the NNLO study.

We mention that the state-of-art determinations of the $O(p^4)$ LECs in $SU(3)$ $\chi$PT suffer uncertainties from the many poorly known $O(p^6)$ LECs $[20, 21]$. Because of the large number of barely known $O(p^6)$ LECs, it is rather difficult to get conclusive results in the present two-loop $SU(3)$ $\chi$PT studies $[20, 21]$. In the present work, there are five $O(p^6)$ LECs, i.e. $C_{12}, C_{14}, C_{17}, C_{19}, C_{31}$, in Eq. (5) and we cannot make precise determinations of these $C_i$
parameters here. Maybe when taking into account the scattering data, one can make more stringent constraints on the $C_i$ LECs in $U(3)\chi$PT. But this is beyond the scope of current work. Instead we take the $C_i$ values from the Dyson-Schwinger–like approach given in Ref. [49], where all of the $O(p^6)\; C_i$ at leading $N_C$ are predicted. In order to show the dependences of the final results on the $C_i$ values, we also perform other fits by using their updated determinations [50]. Like in Ref. [20], we multiply the $O(p^6)\; C_i$ from Refs. [49, 50] by a global factor $\alpha$ and consider $\alpha$ as a free parameter in the fits. In this way, we partially compensate the large uncertainties of the $C_i$ parameters.

For the operators of $v_2^{(2)}$, $L_{18}$ and $L_{25}$ in Eq. (5), they are not present in $SU(3)\chi$PT and purely contribute to the $\eta-\eta'$ mixing, being irrelevant to the pion and kaon observables. Since the $\eta-\eta'$ mixing parameters have already been satisfactorily described in the NLO fits, we do not further include $v_2^{(2)}$, $L_{18}$ and $L_{25}$ at NNLO study. Their inclusion in the present analysis tend to make the fit unstable. Clearly studying more $\eta^{(')}$ related observables it would be possible to extract these parameters but this is out of the reach of the present analysis. A global fit is too unconstrained, being unstable and producing values of the latter couplings compatible with zero within uncertainties. Then we are left with three $O(N_0^0, p^4)$ operators: $L_4$, $L_6$ and $L_7$, which have corresponding parts in $SU(3)\chi$PT. Since $U(3)$ and $SU(3)\chi$PT contain different dynamical degrees of freedom, the corresponding LECs from the two theories can be different. A typical example is the $L_7$ parameter in $SU(3)\chi$PT, which is demonstrated to be dominated by the singlet $\eta_0$ state [1]. Since in $U(3)\chi$PT the singlet $\eta_0$ has been explicitly introduced, the value of $L_7$ in this theory can be totally different from $L_7^{SU(3)}$ in $SU(3)$ case. While for other $O(p^4)$ LECs, such as $L_{i=4,5,6,8}$, the differences between $U(3)$ and $SU(3)\chi$PT are not expected to be as large as the $L_7$ case, since they do not receive the tree-level contributions from the $\eta_0$ state.

Another subtlety to take into account is that $m_\eta$ and $m_{\eta'}$ appear in the chiral loops and, at the same time, the
final expressions of \( m_\eta \) and \( m_\eta' \) depend on the these loops as well. In order to avoid making the complicated iterative procedure to obtain the \( \eta-\eta' \) mixing parameters, we use the LO formulas for \( m_\eta \) and \( m_\eta' \) in the chiral loops. The differences caused by this simple treatment and the strict iterative procedure are beyond the NNLO precision in \( \delta \) expansion, since the chiral loops themselves are already NNLO. Our simple solution is also justified by the fact that the LO description of \( m_\eta \) and \( m_\eta' \) is in qualitative agreement with the lattice simulation data, as shown in Sect. III A. Since the qualitative agreement between the LO formulas and the lattice simulation data requires the value of \( M_0 \) to be around 835.7 MeV, as given in Eq. (52), we fix \( M_0 = 835.7 \) MeV in the following discussions. This also helps to stabilize the NNLO fits, with its many free parameters. Other useful criteria to discriminate reasonable fits are the a priori ranges estimated in Eq. (54), since the fits with large magnitudes of \( \Lambda_1 \) and \( \Lambda_2 \) imply unphysically large corrections to the \( \eta-\eta' \) mixing parameters and the breakdown of the \( \delta \) expansion. In the following we only present the fit results that are consistent with Eq. (54).

With all of the above setups, the values of parameters from the NNLO fits are summarized in Table V. The fits labeled by NNLOFit-A and NNLOFit-B correspond to using different values of the \( \mathcal{O}(p^6) \) LECs. For NNLOFit-A, the \( C_i \) values are taken from Ref. [49]:

\[
C_{12} = -0.34 , \quad C_{14} = -0.83 , \quad C_{17} = 0.01 , \quad C_{19} = -0.48 , \quad C_{31} = -0.63 ,
\]

which are given in units of \( 10^{-3} \text{GeV}^{-2} \). For NNLOFit-B, we take their updated \( \mathcal{O}(p^6) \) \( C_i \) values from Ref. [50]:

\[
C_{12} = -0.34 , \quad C_{14} = -0.87 , \quad C_{17} = 0.17 , \quad C_{19} = -0.27 , \quad C_{31} = -0.46 ,
\]

which are given in units of \( 10^{-3} \text{GeV}^{-2} \).

It is clear that the parameters resulting from fits with different \( C_i \) inputs slightly differ from one another. We remind that the first error bar for each parameter in Table V corresponds to the statistical one directly from the fits and the second error bar stands for the systematical one, which is estimated, as usual, from the variation of the parameter from the NNLO fits with the theoretical expressions in terms of \( F \) and those expressed as functions of \( F_\pi \).

At NNLO, one has the contributions from the chiral loops and the \( \mathcal{O}(p^6) \) LECs, which make our determinations in Table V closer to the recent two-loop results of the \( SU(3) \) \( \chi \text{PT} \) LECs, comparing with the NLO determinations in Table I. Some typical trends of the values of parameters from the NLO study in Table I to the NNLO one in Table V are summarized now. The axial-vector decay constant \( F \) at leading \( N_C \) and chiral limit is reduced at NNLO, which is mainly due to the inclusion of \( L_4 \). Our conclusion is based on the fact that strong correlations between \( F \) and \( L_4 \) always appear, which has been confirmed in previous study [22, 23]. For \( L_5 \) and \( L_8 \), we find that their values are obviously reduced compared to the NLO determination and become closer to the two-loop results in Ref. [21].

As mentioned in the former reference, the discussions in the two-loop \( SU(3) \) \( \chi \text{PT} \) are sensitive to the value of the \( 1/N_C \) suppressed LEC \( L_4 \). The present study provides an independent determination for this parameter and for the \( 1/N_C \) suppressed LEC \( L_6 \) as well. We mention that our determinations of \( L_4 \) have opposite signs with respect to that in Ref. [21], which may be the source of the smaller \( F \) obtained in that reference. Fit solutions with larger \( \Lambda_1 \) and \( \Lambda_2 \) than those in Eq. (54) (out of the an a priori range (54)) are discarded: they are not considered as reasonable physical solutions and will not be discussed any further. According to the values of \( \alpha \) in the two fits, it seems that our study somewhat prefers smaller magnitudes of the \( \mathcal{O}(p^6) \) LECs than those from the Dyson-Schwinger approach given in Refs. [49, 50] and also prefers a global change of sign with respect to Eqs. (55) and (56). We have investigated the impact of fitting \( \alpha \) but releasing one of the \( \mathcal{O}(p^6) \) LECs as an independent parameter (e.g., \( C_{14} \)), but no definitive conclusion could be extracted. These puzzles cannot be resolved here and it is definitely interesting and necessary to further investigate the values of the \( \mathcal{O}(p^6) \) LECs in the future.

The various plots from the NNLO fits are shown in Fig. 2 for \( m_\eta \) and \( m_\eta' \), Fig. 3 for \( m_K \), Fig. 4 for \( F_\pi \) and \( F_K \), and Fig. 5 for the ratio \( F_K/F_\pi \), together with the NLO results and the lattice simulation data and experimental inputs. The shaded area surrounding each curve represents the statistical error band. The figures from NNLOFit-B are compatible with those from NNLOFit-A within the uncertainties, so we only show the results for the former in Figs. 2, 3, 4 and 5. We observe slight improvements of the masses for \( \eta, \eta' \) and significant improvements for \( m_K, F_\pi, F_K \) and the ratios of \( F_K/F_\pi \), comparing with the curves from NLO study. Moreover the \( \chi^2 \) for the NNLO fits are greatly reduced compared with \( \chi^2 \) for the NLO ones, indicating that the NNLO corrections are important at the present level of precision and essential to simultaneously describe the lattice simulation data and experimental inputs of the light pseudoscalar mesons \( \pi, K, \eta \) and \( \eta' \).
\[
\begin{array}{|c|c|c|}
\hline
\text{Parameters} & \text{Inputs} & \text{NNLOFit-A} & \text{NNLOFit-B} \\
\hline
F_0 (\text{MeV}) & 118.0 \pm 16.5 & 106.6 \pm 1.6 \pm 2.2 & 103.0 \pm 1.6 \pm 0.2 \\
F_8 (\text{MeV}) & 133.7 \pm 11.1 & 123.5 \pm 1.2 \pm 2.5 & 119.7 \pm 1.1 \pm 0.5 \\
\theta_0 (\text{Degree}) & -11.0 \pm 3.0 & -7.9 \pm 1.1 \pm 2.0 & -8.1 \pm 1.1 \pm 2.4 \\
\theta_8 (\text{Degree}) & -26.7 \pm 5.4 & -27.2 \pm 1.1 \pm 2.0 & -25.3 \pm 1.0 \pm 1.0 \\
m_s/\hat{m} & 27.5 \pm 3.0 & 28.6 \pm 0.8 \pm 4.2 & 34.0 \pm 0.7 \pm 2.0 \\
F_5 (\text{MeV}) & 106.0 \pm 11.4* & 92.5 \pm 1.1 \pm 0.9 & 91.4 \pm 1.0 \pm 0.3 \\
F_6 (\text{MeV}) & 143.8 \pm 16.5* & 134.3 \pm 1.5 \pm 2.0 & 128.7 \pm 1.6 \pm 2.4 \\
\theta_6 (\text{Degree}) & 34.5 \pm 5.4* & 35.4 \pm 1.4 \pm 0.8 & 36.8 \pm 0.7 \pm 0.3 \\
\theta_8 (\text{Degree}) & 36.0 \pm 4.2* & 37.2 \pm 1.0 \pm 0.9 & 38.4 \pm 1.2 \pm 2.4 \\
\hline
\end{array}
\]

TABLE VI: Parameters from the NNLO fits. In all of these fits, \( M_0 \) is fixed at 835.7 MeV from its LO determination. The precise details of the different notations are explained in the text. The meaning of the error bars is the same as in Table I.

IV. CONCLUSIONS

In this article we have performed a thorough study on the \( \eta-\eta' \) mixing, and axial-vector decay constants for the pion and kaon, up to next-to-next-to-leading order in \( \delta \) expansion within \( U(3) \) chiral perturbation theory. We have carried on a detailed scrutiny and discussions of our results, which have been carefully compared to other works in literature for the \( \eta-\eta' \) mixing. A general mixing formalism, including the higher-derivative terms and kinematic mixing cases, has been addressed in detail. The connections between the mixing parameters from the popular two-mixing-angle scheme and the low energy constants from chiral perturbation theory have been established, both for the singlet-octet basis and the quark-flavor basis.

The considered quantities, including the masses of \( \eta, \eta' \) and \( K \), the quark mass ratio of \( m_s/\hat{m} \), the parameters in the two-mixing-angle scheme and the \( \pi, K \) decay constants have been confronted with recent lattice simulations and phenomenological inputs. We find that the next-to-leading-order fits yield satisfactory descriptions for the masses of the three pseudoscalar mesons as functions of \( m_s^2 \) and the four mixing parameters \((F_0, F_8, \theta_0, \theta_8)\), producing in addition reasonable values of low energy constants. Nonetheless, when the \( \pi \) and \( K \) decay constants are included together with the masses and mixing parameters in the fits, the next-to-leading-order analyses are inadequate and it is necessary to step into the next-to-next-to-leading-order study. Using the \( \mathcal{O}(p^6) \) LECs determinations from a Dyson-Schwinger-like approach \([49, 50]\) multiplied by a global factor, we are able to achieve a reasonable description for all of the physical quantities considered above and the resulting values for the leading \( N_C \mathcal{O}(p^3) \) low energy constants \( L_5 \) and \( L_8 \) turn to be compatible with the very recent two-loop determinations in Ref. [21]. Therefore we conclude that the large \( N_C \ U(3) \) chiral perturbation theory offers a concise theoretical framework that is able to simultaneously reproduce accurately the general \( \eta-\eta' \) mixing and to provide sophisticated enough expressions to describe the chiral extrapolations of the \( \pi \) and \( K \) decay constants and masses.

Finally, it is worthy to remark that some of the parameters in our best analysis (NNLOFit-B) in Table V have been determined with relatively small errors. For instance, the NLO parameters \( \Lambda_{1,2} \), which are fitted up to \( \mathcal{O}(N_C^{-2}) \) in the NNLO analysis, become

\[
\Lambda_1 = -0.002 \pm 0.061 \pm 0.088, \quad \Lambda_2 = -0.20 \pm 0.12 \pm 0.06.
\]
The NNLO fit also determines some $U(3)$ NNLO couplings with relatively high precision. NNLOFit-B yields

$$10^3 \times L_4 = -0.18 \pm 0.06 \pm 0.14, \quad 10^3 \times L_6 = 0.08 \pm 0.03 \pm 0.07, \quad 10^3 \times L_7 = 0.34 \pm 0.05 \pm 0.10. \quad (58)$$

Even though the error estimates in the present article must be considered with some caution, as some lattice systematic uncertainties escape our control, this hints the potentiality of this $U(3)$ $\chi$PT framework. We hope these results may encourage future lattice analyses along this line.

**Acknowledgments**

We thank Shao-Zhou Jiang for communication on the updated values of the $O(p^6)$ LECs. This work is supported in part by the National Natural Science Foundation of China (NSFC) under Grant No. 11105038, the grants from the Education Department of Hebei Province under contract No. YQ2014034, the grants from the Department of Human Resources and Social Security of Hebei Province with contract No. C201400323, and the Doctor Foundation of Hebei Normal University under Contract No. L2010B04, the Spanish Government (MINECO) and the European Commission (ERDF) [FPA2010-17747, FPA2013-44773-P, FPA2013-40483-P, SEV-2012-0249 (Severo Ochoa Program), CSD2007-00042 (Consolider Project CPAN), the Comunidad de Madrid [HEPHACOS S2009/ESP-1473], and EPOS network of the European Community Research Integrating Activity “Study of Strongly Interacting Matter” (HadronPhysics3, Grant No. 283286). J.J. Sanz-Cillero wants to thank the Center for Future High Energy Physics and the Institute of High Energy Physics in Beijing for their hospitality.

**Appendix A: Higher order corrections to the $\pi$ and $\pi'$ bilinear terms**

In the following we provide the explicit expressions of the $\delta_i$’s in Eq. (11). When expressing the results in terms of $F$, they take the form

$$\delta_1 = \frac{32C_{12}}{3F^2} [c_\phi(m^2_K - m^2_\pi) + 4\sqrt{2}c_\phi s_\theta (m^2_K - m^2_\pi) + s^2_\phi (2m^2_K + m^2_\pi)], \quad (A1)$$

$$\delta_2 = \frac{32C_{12}}{3F^2} [c_\phi(2m^2_K + m^2_\pi) - 4\sqrt{2}c_\phi s_\theta (m^2_K - m^2_\pi) + s^2_\phi (4m^2_K - m^2_\pi)], \quad (A2)$$

$$\delta_3 = -\frac{64C_{12}}{3F^2} (m^2_K - m^2_\pi) (\sqrt{2}c^2_\phi - c_\phi s_\theta - \sqrt{2}s^2_\phi), \quad (A3)$$

$$\delta_\pi = \frac{8L_5 [c_\phi (4m^2_K - m^2_\pi) + 4\sqrt{2}c_\phi (m^2_K - m^2_\pi) s_\theta + (2m^2_K + m^2_\pi) s^2_\phi]}{3F^2} + \frac{c^2_\phi A_0 (m^2_K)}{16\pi^2 F^2} + \frac{8L_4 (2m^2_K + m^2_\pi)}{F^2} + \frac{8L_1 s_\theta [2\sqrt{2}c_\phi (m^2_K - m^2_\pi) + (2m^2_K + m^2_\pi)s_\phi]}{F^2} + \frac{64L_5 (L_5 - 2L_8)}{3F^4} [c_\phi (4m^4_K - m^4_\pi) + 4\sqrt{2}c_\phi (m^4_K - m^4_\pi) s_\theta + (4m^4_K - 4m^2_K m^2_\pi + 3m^4_\pi) s^2_\phi], \quad (A4)$$

$$\delta_\pi' = \frac{8L_5 [c_\phi (2m^2_K + m^2_\pi) + 4\sqrt{2}c_\phi (-m^2_K - m^2_\pi) s_\theta + (4m^2_K - m^2_\pi) s^2_\phi]}{3F^2} + \frac{s^2_\phi A_0 (m^2_K)}{16\pi^2 F^2} + \frac{8L_4 (2m^2_K + m^2_\pi)}{F^2} + \frac{8L_1 c_\phi (2m^2_K + m^2_\pi) + 2\sqrt{2}(-m^2_K + m^2_\pi) s_\phi]}{F^2} + \frac{64L_5 (L_5 - 2L_8)}{3F^4} [c_\phi (2m^4_K + m^4_\pi) + 4\sqrt{2}c_\phi (-m^4_K + m^4_\pi) s_\theta + (4m^4_K - m^4_\pi) s^2_\phi], \quad (A5)$$
\[
\delta_k = -\frac{16L_5(m_K^2 - m_\pi^2)(\sqrt{2}c_\theta^2 - c_\theta s_\theta - \sqrt{2}s_\theta^2)}{3F^2} - c_\theta s_\theta A_1 \\
+ \frac{c_\theta s_\theta A_0(m_K^2)}{16\pi^2 F^2} - \frac{8L_1\sqrt{2}c_\theta^2(m_K^2 - m_\pi^2) + c_\theta(2m_K^2 + m_\pi^2)s_\theta + \sqrt{2}(-m_K^2 + m_\pi^2)s_\theta^2}{F^2} \\
- \frac{128L_5(L_5 - 2L_8)(m_K^4 - m_\pi^4)(\sqrt{2}c_\theta^2 - c_\theta s_\theta - \sqrt{2}s_\theta^2)}{3F^4} \\
- \frac{64(C_{14} + C_{17})m_K^2(m_K^2 - m_\pi^2)(\sqrt{2}c_\theta^2 - c_\theta s_\theta - \sqrt{2}s_\theta^2)}{3F^2},
\]
(A6)

\[
\delta_{\phi^2} = \frac{16L_5^4}{3F^4} [c_\theta(8m_K^2 - 8m_\pi^2m_\pi^2 + 3m_\pi^4) + 8\sqrt{2}c_\theta m_K^2(m_K^2 - m_\pi^2)s_\theta + (4m_\pi^4 - 4m_K^2m_\pi^2 + 3m_\pi^4)s_\theta^2] \\
+ \frac{2}{3} s_\theta[2\sqrt{2}c_\theta(m_K^2 - m_\pi^2) + (2m_K^2 + m_\pi^2)s_\theta]A_2 \\
+ \frac{1}{16\pi^2} \frac{1}{18F^2} \left[ c_\theta^2(16m_K^2 - 7m_\pi^2) + 4\sqrt{2}c_\theta^3(8m_K^2 - 5m_\pi^2)s_\theta + 12c_\theta^2(4m_K^2 - m_\pi^2)s_\theta^2 \\
+ 16\sqrt{2}c_\theta(m_K^2 - m_\pi^2)s_\theta + 2(2m_K^2 + m_\pi^2)s_\theta^2 \right] A_0(m_\pi^2) \\
+ \frac{4m_K^2 - m_\pi^2}{18F^2} \left[ c_\theta^2 m_\pi^2 + 2\sqrt{2}c_\theta(m_K^2 - m_\pi^2)s_\theta - 4m_K^2 s_\theta^2 \right] A_0(m_K^2) + \frac{m_\pi^2(c_\theta^2 - 2\sqrt{2}c_\theta s_\theta + 2s_\theta^2)}{2F^2} A_0(m_\pi^2) \\
- \frac{16L_5 s_\theta[4\sqrt{2}c_\theta m_K^2(m_K^2 - m_\pi^2) + (4m_K^2 - 4m_\pi^2m_\pi^2 + 3m_\pi^4)s_\theta]}{F^2} + 6(2m_K^2 + m_\pi^2)s_\theta^2 \phi(2) \\
+ \frac{16L_6(2m_K^2 + m_\pi^2)[c_\theta^2(4m_K^2 - m_\pi^2) + 4\sqrt{2}c_\theta(2m_K^2 - m_\pi^2)m_\pi^2 - m_\pi^4)s_\theta + (2m_K^2 + m_\pi^2)^2 s_\theta^2]}{3F^2} \\
+ \frac{16L_7[8c_\theta^2(m_K^2 - m_\pi^2)^2 + 4\sqrt{2}c_\theta(2m_K^2 - m_\pi^2)m_\pi^2 + m_\pi^4)s_\theta + (2m_K^2 + m_\pi^2)^2 s_\theta^2]}{3F^2} \\
+ \frac{256L_5 - 2L_8) L_8}{3F^4} \left[ c_\theta^2(8m_K^2 - 4m_K^2m_\pi^2 - 4m_\pi^4 + 3m_\pi^6) + 4\sqrt{2}c_\theta m_K^2(2m_K^2 - m_\pi^2)m_\pi^2 + m_\pi^4)s_\theta + (4m_\pi^6 - 2m_K^2m_\pi^2 - 2m_\pi^4 + 3m_\pi^6)s_\theta^2 \right] \\
+ \frac{16L_5 - 2L_8) A_2 s_\theta[2\sqrt{2}c_\theta(m_K^4 - m_\pi^4) + (2m_K^2 + m_\pi^2)s_\theta]}{3F^2} \\
+ \frac{16(3C_{19} + 2C_{31})}{3F^2} \left[ c_\theta(16m_K^6 - 24m_K^4m_\pi^2 + 12m_K^2m_\pi^4 - m_\pi^6) \\
+ 4\sqrt{2}c_\theta(4m_K^6 - 6m_K^4m_\pi^2 + 3m_K^2m_\pi^4 + m_\pi^6)s_\theta + (8m_K^6 - 12m_K^4m_\pi^2 + 6m_K^2m_\pi^4 + m_\pi^6)s_\theta^2 \right],
\]
(A7)
\[ \delta m_{\pi}^2 = \frac{2}{3} c_\theta [c_\theta (2m_K^2 + m_\pi^2) + 2\sqrt{2}(-m_K^2 + m_\pi^2)s_\theta] A_2 \\
+ \frac{16L_6}{3F^2} [c_\theta^2 (4m_K^4 - 4m_K^2m_\pi^2 + 3m_\pi^4) + 8\sqrt{2}c_\theta m_\pi^2 (-m_K^2 + m_\pi^2)s_\theta + (8m_K^4 - 8m_K^2m_\pi^2 + 3m_\pi^4)s_\theta^2] \\
+ \frac{1}{16\pi^2} \left\{ \frac{(4m_K^2 - m_\pi^2)(2c_\theta^2 - 2\sqrt{2}c_\theta s_\theta - 3c_\theta^2 s_\theta^2 + 2\sqrt{2}c_\theta s_\theta^3 + 2s_\theta^4)}{18F^2} A_0(m_\pi^2) \right\} \\
+ \frac{1}{18F^2} \left[ 2c_\theta^2 (2m_K^2 + m_\pi^2) - 16\sqrt{2}c_\theta (m_K^2 - m_\pi^2)s_\theta + 12c_\theta (4m_K^2 - m_\pi^2)s_\theta^2 \right] \\
- 4\sqrt{2}c_\theta (8m_K^2 - 5m_\pi^2)s_\theta^2 + (16m_K^2 - 7m_\pi^2)s_\theta^4 \right\} A_0(m_\pi^2) \\
- \frac{-4c_\theta^2 m_K^2 + 2\sqrt{2}c_\theta (2m_K^2 - m_\pi^2)s_\theta + m_\pi^2 s_\theta^2}{3F^2} A_0(m_\pi^2) + \frac{m_\pi^2 (2c_\theta^2 + 2\sqrt{2}c_\theta s_\theta + s_\theta^2)}{2F^2} A_0(m_\pi^2) \\
- \frac{16L_2c_\theta [c_\theta (4m_K^4 - 4m_K^2m_\pi^2 + 3m_\pi^4) + 4\sqrt{2}m_\pi^2 (-m_K^2 + m_\pi^2)s_\theta + 6c_\theta (2m_K^2 + m_\pi^2)s_\theta^2]}{F^2} \\
+ \frac{16L_7 [c_\theta^2 (2m_K^2 + m_\pi^2)^2 + 4\sqrt{2}c_\theta (-2m_K^2 + m_K^2m_\pi^2 + m_\pi^4)s_\theta + 8(m_K^2 - m_\pi^2)^2 s_\theta^2]}{3F^2} \\
+ \frac{16L_6 (2m_K^2 + m_\pi^2)[c_\theta^2 (2m_K^2 + m_\pi^2) + 4\sqrt{2}c_\theta (-m_K^2 + m_\pi^2)s_\theta + (4m_K^2 - m_\pi^2)s_\theta^2]}{3F^2} \\
+ \frac{256(L_5 - 2L_8)L_8}{3F^2} [c_\theta^2 (4m_K^4 - 2m_K^2m_\pi^2 - 2m_\pi^2 m_K^4 + 3m_\pi^6) + \\
4\sqrt{2}c_\theta m_K^4 (-2m_K^2 + m_K^2m_\pi^2 + m_\pi^4)s_\theta + (8m_K^2 - 4m_K^2m_\pi^2 - 4m_K^2m_\pi^4 + 3m_\pi^6)s_\theta^2] \\
+ \frac{16(L_5 - 2L_8)A_2 c_\theta [2\sqrt{2}s_\theta (-m_K^4 + m_\pi^4) + (2m_K^4 + m_\pi^4)c_\theta]}{3F^2} \\
+ \frac{16(3C_{10} + 2C_{31})}{3F^2} \left[ c_\theta^2 (8m_K^6 - 12m_K^2m_\pi^2 + 6m_\pi^2 m_K^4 + m_\pi^6) \\
- 4\sqrt{2}c_\theta (4m_K^6 - 6m_K^2m_\pi^2 + 3m_\pi^2 m_K^4 - m_\pi^6) s_\theta + (16m_K^6 - 24m_K^4 m_\pi^2 + 12m_K^2 m_\pi^4 - m_\pi^6)s_\theta^2 \right], \quad (A8) \]
where \( \delta_{m^2} \) is:

\[
\delta_{m^2} = - \frac{64 L_5 s m^2_K (m^2_K - m^2_\pi) (\sqrt{2} c^2_\theta - c_\theta s_\theta - \sqrt{2} s^2_\theta)}{3 F^2} \\
- \frac{2}{3} \left[ \frac{\sqrt{2} c^2_\theta (m^2_K - m^2_\pi) + c_\theta (2 m^2_K + m^2_\pi) s_\theta + \sqrt{2} (-m^2_K + m^2_\pi) s^2_\theta}{A_2} \right] \\
- \frac{1}{288 \pi^2 F^2} \left[ \frac{\sqrt{2} c^2_\theta (8 m^2_K - 5 m^2_\pi) + c^2_\theta (8 m^2_K + m^2_\pi) s_\theta + 3 \sqrt{2} c^2_\theta (-4 m^2_K + m^2_\pi) s^2_\theta}{A_0 (m^2_\pi)} + 4 c_\theta (5 m^2_K + 2 m^2_\pi) s^3_\theta + 4 \sqrt{2} (-m^2_K + m^2_\pi) s^3_\theta \right] \\
+ \frac{4 \sqrt{2} c^2_\theta (4 m^2_K - m^2_\pi) s_\theta + c_\theta (8 m^2_K + m^2_\pi) s_\theta + \sqrt{2} (-8 m^2_K + 5 m^2_\pi) s^4_\theta}{A_0 (m^2_\pi)} + 4 c^3_\theta s_\theta (-5 m^2_K + 2 m^2_\pi) \right] A_0 (m^2_\pi) \\
+ 6 \sqrt{2} c^2_\theta (2 m^2_K - m^2_\pi) + c_\theta (4 m^2_K + m^2_\pi) + \sqrt{2} (-2 m^2_K + m^2_\pi) A_0 (m^2_\pi) \\
+ 9 m^2_\pi (-\sqrt{2} c_\theta + c_\theta s_\theta + \sqrt{2} s_\theta) A_0 (m^2_\pi) \right] \\
- \frac{64 L_5 s m^2_K (m^2_K - m^2_\pi) (\sqrt{2} c^2_\theta - c_\theta s_\theta - \sqrt{2} s^2_\theta)}{3 F^2} \\
+ \frac{16 L_5 s}{F^2} \left[ \frac{\sqrt{2} c^2_\theta (2 m^2_K - m^2_\pi - m^2_\pi - m^2_\pi) + c_\theta (4 m^2_K + 4 m^2_\pi + 3 m^2_\pi) + 2 \sqrt{2} m^2_K (-m^2_K + m^2_\pi) s_\theta}{A_0 (m^2_\pi)} \right] \\
- \frac{16 L_5}{3 F^2} \left[ \frac{2 \sqrt{2} c^2_\theta (2 m^2_K - m^2_\pi - m^2_\pi) + c_\theta (-4 m^4 + 20 m^2_K m^2_\pi - 7 m^2_\pi) s_\theta + 2 \sqrt{2} (-m^2_K + m^2_\pi + m^2_\pi) s_\theta}{A_0 (m^2_\pi)} \right] \\
- \frac{512 (L_5 - 2 L_8) L_8 s m^2_K (2 m^4_K - m^2_K m^2_\pi - m^2_\pi) (\sqrt{2} c_\theta - c_\theta s_\theta - \sqrt{2} s_\theta)}{3 F^2} \\
+ \frac{16 (L_5 - 2 L_8) A_2 [\sqrt{2} c^2_\theta (m^2_K - m^2_\pi) + c_\theta (2 m^2_K + m^2_\pi) s_\theta + \sqrt{2} (-m^2_K + m^2_\pi) s^2_\theta]}{3 F^2} \\
- \frac{32 (3 C_{19} + 2 C_{31}) (4 m^4_K - 6 m^4_K m^2_\pi + 3 m^2_K m^4_\pi - m^2_\pi) (\sqrt{2} c_\theta - c_\theta s_\theta - \sqrt{2} s_\theta)}{3 F^2} .
\]

When expressing the above results in terms of \( F^\pi \) from Eq. (27), the terms with \( L_8^2 \) and \( L_5 L_8 \) can be different from the expressions in \( F^\pi \) and the other parts remain the same, apart from the obvious replacement of \( F^\pi \) by \( F^\pi \). Therefore, for the expressions of \( \delta_i \) expressed in \( F^\pi \), we only give the parts that are different from those in terms of \( F^\pi \):

\[
\delta^{(F^\pi), L_5^2}_{F^\pi} = \frac{128 L_5^2}{3 F^4} \left[ c^2_\theta (2 m^4_K + 2 m^2_K m^2_\pi - m^4_\pi) + 2 \sqrt{2} c_\theta (m^2_K + m^2_\pi - m^2_\pi) + s^2_\theta (m^2_K + m^2_\pi + m^2_\pi) \right] ,
\]

\[
\delta^{(F^\pi), L_5 L_8}_{F^\pi} = \frac{128 L_5 L_8}{3 F^4} \left[ c^2_\theta (16 m^6_K - 16 m^2_K m^2_\pi + 9 m^6_\pi) + 16 \sqrt{2} c_\theta m^2_K (m^4_K - m^4_\pi) s_\theta \right. \\
+ (8 m^6_K - 8 m^4_K m^2_\pi + 9 m^6_\pi) s^2_\theta \right] ,
\]

\[
\delta^{(F^\pi), L_8^2}_{F^\pi} = \frac{128 L_8^2}{3 F^4} \left[ c^2_\theta (8 m^6_K - 8 m^2_K m^2_\pi + 9 m^6_\pi) - 16 \sqrt{2} c_\theta m^2_K (m^4_K - m^4_\pi) s_\theta \right. \\
+ (16 m^6_K - 16 m^4_K m^2_\pi + 9 m^6_\pi) s^2_\theta \right] ,
\]

\[\text{(A9)}\]
\[ \delta_{m^2, L_5 L_8} = -\frac{1024 L_5 L_8 m_K^2 (m_K^2 - m^2) (\sqrt{2} c_0^2 - c_0 s_0 - \sqrt{2} s_0^2)}{3 F_\pi^2}. \]  
(A15)

In order to obtain the full expressions for the \( \delta_i \)'s given in terms of \( F_\pi \) one has to make use of Eq. (27) up to the precision required. Taking \( \delta_k \) for example, its final expression in terms of \( F_\pi \) is

\[
\delta_k = -\frac{16 L_5 (m_K^2 - m^2) (\sqrt{2} c_0^2 - c_0 s_0 - \sqrt{2} s_0^2)}{3 F_\pi^2} - c_0 s_0 A_1 + \frac{8 L_{18} (\sqrt{2} c_0^2 (m_K^2 - m^2) + c_9 (2 m_K^2 + m^2) s_0 + \sqrt{2} (-m_K^2 + m^2) s_0^2)}{16 m^2 F_\pi^2} \\
+ \frac{256 L_5 L_8 (m_K^2 - m^2) (\sqrt{2} c_0^2 - c_0 s_0 - \sqrt{2} s_0^2) - 128 L_5^2 (m_K^2 + m_K^2 m^2 - 2 m_1^2) (\sqrt{2} c_0^2 - c_0 s_0 - \sqrt{2} s_0^2)}{3 F_\pi^2} - \frac{64 (C_{14} + C_{17}) m_K^2 (m_K^2 - m^2) (\sqrt{2} c_0^2 - c_0 s_0 - \sqrt{2} s_0^2)}{3 F_\pi^2},
\]

(A16)

which differs from Eq. (A6) in the \( L_5^2 \) term. For \( \delta_1, \delta_2 \) and \( \delta_3 \), their expressions are the same regardless of whether \( F \) or \( F_\pi \) is chosen up to next-to-next-to-leading order.

[1] J. Gasser and H. Leutwyler, Ann. Phys. 158 (1984) 142; Nucl. Phys. B 250 (1985) 465.
[2] G. 't Hooft, Nucl. Phys. B 72 (1974) 461; 75 (1974) 461; E. Witten, Nucl. Phys. B 160 (1979) 57.
[3] P. Di Vecchia and G. Veneziano, Nucl. Phys. B 171 (1980) 253; C. Rosenzweig, J. Schechter and T. Trahem, Phys. Rev. D 21 (1980) 3388; E. Witten, Ann. Phys. 128 (1980) 363.
[4] H. Leutwyler, Nucl. Phys. Proc. Suppl. 64 (1998) 223.
[5] R. Kaiser and H. Leutwyler, hep-ph/9806336.
[6] R. Kaiser and H. Leutwyler, Eur. Phys. J. C17 (2000) 623.
[7] P. Herrera-Siklody, J. I. Latorre, P. Pascual and J. Taron, Nucl. Phys. B497 (1997) 345.
[8] P. Herrera-Siklody, J. I. Latorre, P. Pascual and J. Taron, Phys. Lett. B419 (1998) 326.
[9] J. J. Dudek, R. G. Edwards, B. Joo, M. J. Peardon, D. G. Richards and C. E. Thomas, Phys. Rev. D 83, 111502 (2011).
[10] N. H. Christ, C. Dawson, T. Izubuchi, C. Jung, Q. Liu, R. D. Mawhinney, C. T. Sachrajda and A. Soni et al., Phys. Rev. Lett. 105, 241601 (2010).
[11] E. B. Gregory et al. [UKQCD Collaboration], Phys. Rev. D 86, 014504 (2012).
[12] C. Michael et al. [European Twisted Mass Collaboration], PoS LATTICE 2013, 253 (2014).
[13] C. Michael et al. [ETM Collaboration], Phys. Rev. Lett. 111, no. 18, 181602 (2013).
[14] Y. H. Chen, Z. H. Guo and H. Q. Zheng, Phys. Rev. D 85, 054018 (2012).
[15] Y. H. Chen, Z. H. Guo and B. S. Zou, Phys. Rev. D 91, 044010 (2015).
[16] Y. Aoki et al. [RBC and UKQCD Collaborations], Phys. Rev. D 83 (2011) 074508.
[17] R. Arthur et al. [RBC and UKQCD Collaborations], Phys. Rev. D 87 (2013) 094514.
[18] S. Durr et al., Phys. Rev. D 81 (2010) 054507.
[19] S. Descotes-Genon, L. Girlanda and J. Stern, JHEP 0001, 041 (2000).
[20] J. Bijnen and I. Jemos, Nucl. Phys. B 854, 631 (2012).
[21] J. Bijnen and G. Ecker, Ann. Rev. Nucl. Part. Sci. 64, 149 (2014).
[22] G. Ecker, P. Masjuan and H. Neufeld, Eur. Phys. J. C 74, no. 2, 2748 (2014).
[23] Z. H. Guo and J. J. Sanz-Cillero, Phys. Rev. D 89, no. 9, 094024 (2014).
[24] T. Feldmann, P. Kroll and B. Stech, Phys. Rev. D 58, 114006 (1998).
[25] R. Escribano and J. M. Frere, J. High Energy Phys. 0506, 029 (2005).
[26] C. E. Thomas, J. High Energy Phys. 0710, 026 (2007).
[27] G. Li, Q. Zhao, and C. H. Chang, J. Phys. G 35, 055002 (2008).
[28] R. Escribano, P. Masjuan and P. Sanchez-Puertas, Phys. Rev. D 89, no. 3, 034014 (2014).
[29] E. Witten, Nucl. Phys. B 156, 269 (1979); S. Coleman and E. Witten, Phys. Rev. Lett. 45, 100 (1980); G. Veneziano, Nucl. Phys. B 159, 213 (1979).
[30] S. Z. Jiang, F. J. Ge and Q. Wang, Phys. Rev. D 89, 074048 (2014).
[31] N. Beisert and B. Borasoy, Eur. Phys. J. A 11, 329 (2001).
[32] J.-M. Gerard and E. Kou, Phys. Lett. B 616, 85 (2005).
[33] C. Degrandi and J.-M. Gerard, JHEP 0905, 043 (2009).
[34] V. Mathieu and V. Vento, Phys. Lett. B 688, 314 (2010).
[35] Z. -H. Guo and J. A. Oller, Phys. Rev. D 84 (2011) 034005.
[36] J. Bijnen, G. Colangelo and G. Ecker, JHEP 02 (1999) 020.
[37] Z. -H. Guo, J. A. Oller and J. Ruiz de Elvira, Phys. Rev. D 86 (2012) 054006.
[38] H. Georgi, Phys. Rev. D 49, 1666 (1994).
[39] S. Peris, Phys. Lett. B 324, 442 (1994).
[40] For a recent review: T. Feldmann, Int. J. Mod. Phys. A 15 (2000) 159.
[41] G. Amoros, J. Bijnens and P. Talavera, Nucl. Phys. B 568, 319 (2000).
[42] V. Bernard and E. Passemas, JHEP 1004, 001 (2010).
[43] S. Aoki, Y. Aoki, C. Bernard, T. Blum, G. Colangelo, M. Della Morte, S. Durr and A. X. El Khadra et al., Eur. Phys. J. C 74, no. 9, 2890 (2014).
[44] K. A. Olive et al. [Particle Data Group Collaboration], Chin. Phys. C 38, 090001 (2014).
[45] R. Escribano, P. Masjuan and J. J. Sanz-Cillero, JHEP 1105, 094 (2011).
[46] G. Ecker et al., Nucl. Phys. B321, 311 (1989).
[47] A. Pich, I. Rosell and J.J. Sanz-Cillero, JHEP 0701 (2007) 039. J. J. Sanz-Cillero and J. Trnka, Phys. Rev. D 81 (2010) 056005.
[48] M. Jamin, J. A. Oller and A. Pich, Nucl. Phys. B 587, 331 (2000).
[49] S. Z. Jiang, Y. Zhang, C. Li and Q. Wang, Phys. Rev. D 81, 014001 (2010).
[50] S. Z. Jiang, Z. L. Wei, Q. S. Chen and Q. Wang, arXiv:1502.05087 [hep-ph].