Ratio of kaon and pion leptonic decay constants with $N_f = 2+1+1$
Wilson-clover twisted-mass fermions

C. Alexandrou,1,2 S. Bacchio,2 G. Bergner,3 P. Dimopoulos,4 J. Finkenrath,2 R. Frezzotti,5 M. Garofalo,6,7 B. Kostrzewa,8 G. Koutsou,2 P. Labus,9 F. Sanfilippo,10 S. Simula,10 M. Ueding,7 C. Urbach,7 and U. Wenger11

(Extended Twisted Mass Collaboration)

1Department of Physics, University of Cyprus, 20537 Nicosia, Cyprus
2Computation-based Science and Technology Research Center, The Cyprus Institute, 20 Konstantinou Kavafi Street, 2121 Nicosia, Cyprus
3University of Jena, Institute for Theoretical Physics, Max-Wien-Platz 1, D-07743 Jena, Germany
4Dipartimento di Scienze Matematiche, Fisiche e Informatiche, Università di Parma and INFN, Gruppo Collegato di Parma, Parco Area delle Scienze 7/a (Campus), I-43124 Parma Italy
5Dipartimento di Fisica, Università di Roma “Tor Vergata” and INFN, Sezione di Tor Vergata, Via della Ricerca Scientifica 1, I-00133 Roma, Italy
6Dipartimento di Fisica, Università Roma Tre and INFN, Sezione di Roma Tre, Via della Vasca Navale 84, I-00146 Rome, Italy
7HISKP (Theory), Rheinische Friedrich-Wilhelms-Universität Bonn, Nussallee 14-16, 53115 Bonn, Germany
8High Performance Computing and Analytics Lab, Rheinische Friedrich-Wilhelms-Universität Bonn, Friedrich-Hirzebruch-Allee 8, 53115 Bonn, Germany
9Fraunhofer Institute for Industrial Mathematics (ITWM), Fraunhofer-Platz 1, 67663 Kaiserslautern, Germany
10Istituto Nazionale di Fisica Nucleare, Sezione di Roma Tre, Via della Vasca Navale 84, I-00146 Rome, Italy
11Institute for Theoretical Physics, Albert Einstein Center for Fundamental Physics, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland

(Received 20 April 2021; accepted 30 September 2021; published 26 October 2021)

We present a determination of the ratio of kaon and pion leptonic decay constants in isosymmetric QCD (ISOQCD), $f_K/f_\pi$, making use of the gauge ensembles produced by the Extended Twisted Mass Collaboration with $N_f = 2+1+1$ flavors of Wilson-clover twisted-mass quarks, including configurations close to the physical point for all dynamical flavors. The simulations are carried out at three values of the lattice spacing ranging from $\sim 0.068$ to $\sim 0.092$ fm with linear lattice size up to $L \sim 5.5$ fm. The scale is set by the particle data group (PDG) value of the pion decay constant, $f_{\pi}^{\text{ISOQCD}} = 130.4(2)$ MeV, at the ISOQCD pion point, $M_{\pi}^{\text{ISOQCD}} = 135.0(2)$ MeV, obtaining for the gradient-flow scales the values $w_0 = 0.17383(63)$ fm, $\sqrt{t_0} = 0.14436(61)$ fm and $t_0/w_0 = 0.11969(62)$ fm. The data are analyzed within the framework of SU(2) chiral perturbation theory without resorting to the use of renormalized quark masses. At the ISOQCD kaon point $M_K^{\text{ISOQCD}} = 494.2(4)$ MeV we get $(f_K/f_\pi)^{\text{ISOQCD}} = 1.1995(44)$, where the error includes both statistical and systematic uncertainties. Implications for the Cabibbo-Kobayashi-Maskawa matrix element $|V_{us}|$ and for the first-row Cabibbo-Kobayashi-Maskawa unitarity are discussed.

DOI: 10.1103/PhysRevD.104.074520

I. INTRODUCTION

The leptonic decay constants of charged pseudoscalar $(P)$ mesons are the crucial hadronic ingredients necessary for obtaining precise information on the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements describing
the weak mixings among quark flavors [1]. Within the Standard Model (SM) the unitarity of the CKM matrix imposes important constraints on various sums of squares of matrix elements and, therefore, any violation of such constraints would imply the presence of physics beyond the SM. The way the CKM entries can be determined is based on the knowledge of the experimental leptonic decay rates and of the corresponding theoretical calculations. In particular, both the charged kaon and pion leptonic decay widths into muons are known experimentally with a good precision [2], obtaining for their ratio the value

$$\frac{\Gamma(K \to \mu\nu\mu[\gamma])}{\Gamma(\pi \to \mu\nu\mu[\gamma])} = 1.3367(2)_{\pi}(29)_{K}[29],$$

(1)

where \(\gamma\) stands for the contribution of virtual and real photons. On the theoretical side, within the SM the above ratio is given by

$$\frac{\Gamma(K \to \mu\nu\mu[\gamma])}{\Gamma(\pi \to \mu\nu\mu[\gamma])} = \frac{V_{us}}{V_{ud}} \frac{f_K}{f_\pi} \left( \mathcal{M}_{K}^2 - m_u^2 \right)^2 \left( 1 + \delta R_{K}\delta \right).$$

(2)

where \(V_{us}\) and \(V_{ud}\) are the relevant CKM entries, \(m_u\) is the charged pion(kaon) mass, \(m_\mu\) is the muon mass and \(\delta R_{K}\delta\) represents the isospin breaking (IB) corrections due both to the mass difference \((m_d - m_u)\) between the light \(u\) and \(d\) quarks and to the quark electric charges. Finally, in Eq. (2) \(f_K/f_\pi\) is the ratio of kaon and pion leptonic decay constants defined in isosymmetric QCD (ISOQCD), i.e., with \(m_u = m_d\) and zero quark electric charges.

Recently [3,4] the IB correction \(\delta R_{K}\delta\) has been determined using a nonperturbative approach, based on first principles, through QCD + QED simulations on the lattice, obtaining \(\delta R_{K}\delta = -0.0126(14)\). From Eq. (1) one has

$$\left| \frac{V_{us}}{V_{ud}} \frac{f_K}{f_\pi} \right| = 0.27683(29)_{\exp}(20)_{th} = 0.27683(35),$$

(3)

which corresponds to an accuracy of \(\approx 0.13\%\). As is well known [5], the IB correction \(\delta R_{K}\delta\) and the ISOQCD ratio \(f_K/f_\pi\) separately depend on the prescription used to define what is meant by ISOQCD, while only the product \(\left(f_K/f_\pi\right)\sqrt{1 + \delta R_{K}\delta}\) is independent on such prescription. The hadronic prescription adopted in Refs. [3,4] corresponds to

$$M_{\pi}^{\text{ISOQCD}} = 135.0(2)\,\text{MeV},$$

$$M_K^{\text{ISOQCD}} = 494.2(4)\,\text{MeV},$$

$$f_\pi^{\text{ISOQCD}} = 130.4(2)\,\text{MeV},$$

(4)–(6)

while the quantity \((m_d - m_u)\) is obtained from the difference between the experimental charged and neutral kaon masses. The physical pion and kaon masses (4)–(5) are consistent with those recommended by FLAG-3 [6], and the pion decay constant (6), derived according to Ref. [7] adopting the value of the CKM entry \(|V_{ud}|\) from Ref. [8], is used to set the lattice scale.\(^1\)

In this work we present our determination of the leptonic decay constant ratio \(f_K/f_\pi\) at the physical ISOQCD point given by Eqs. (4)–(6), evaluated using the Extended Twisted Mass Collaboration (ETMC) gauge ensembles produced with \(N_f = 2 + 1 + 1\) flavors of Wilson Clover twisted-mass quarks, including configurations close to the physical point for all dynamical flavors [10,11]. The lattice data will be analyzed within the framework of SU(2) chiral perturbation theory (ChPT) without making use of renormalized quark masses.\(^2\) By means of the pion data we determine the gradient-flow (GF) scales \(w_0\) [12], \(\sqrt{f_0}\) [13] and \(t_0/w_0\) adopting the physical value (6) at the pion point (4) to set the lattice scale, obtaining

$$w_0 = 0.17383(63)\,\text{fm},$$

$$\sqrt{f_0} = 0.14436(61)\,\text{fm},$$

$$t_0/w_0 = 0.11969(62)\,\text{fm},$$

(7)–(9)

where the error includes both statistical and systematic uncertainties. Our findings (7)–(8) are a little larger than the MILC results [14] \(w_0 = 0.1714^{+15}_{-12}\,\text{ fm}\) and \(\sqrt{f_0} = 0.1416^{+8}_{-7}\,\text{ fm}\) as well as the high precision QCD (HPQCD) results [15] \(w_0 = 0.1715(9)\,\text{ fm}\) and \(\sqrt{f_0} = 0.10(8)\,\text{ fm}\), both obtained using the hadronic value (6) to set the lattice scale. Within \(\approx 1.5\) standard deviations our result (7) is consistent with the recent, precise Budapest-Marseille-Wuppertal (BMW) determination \(w_0 = 0.17236(70)\,\text{ fm}\), obtained in Ref. [16] using the \(\Omega^+\)-baryon mass to set the lattice scale. Furthermore, the differences with the recent results \(w_0 = 0.1709(11)\,\text{ fm}\) and \(\sqrt{f_0} = 0.1422(14)\,\text{ fm}\), obtained in Ref. [17] using the \(\Omega^+\)-baryon mass to set the lattice scale, are within \(\approx 2\) and \(\approx 1.5\) standard deviations, respectively.

As for the ratio \(f_K/f_\pi\) we determine its value at the physical ISOQCD point (4)–(6) and in the continuum and infinite volume limits, obtaining

\(^1\)In Ref. [4] it has been shown that within the precision of the lattice simulations the prescription given by Eqs. (4)–(6) is equivalent to the Gasser-Rusetsky-Scimemi scheme [9], where the renormalized quark masses and coupling constant in a given short-distance scheme (viz. the \(\overline{\text{MS}}\) scheme) and at a given scale (viz. 2 GeV) are equal in the full QCD + QED and ISOQCD theories. For completeness we mention that in the charm sector the \(D_s\)-meson mass \(M_{D_s}^{\text{ISOQCD}}\) was chosen to be equal to its experimental value \(M_{D_s} = 1969.0(1.4)\,\text{MeV}\) [2].

\(^2\)An analysis of the kaon and pion masses and decay constants in terms of renormalized quark masses is ongoing and will be presented in a forthcoming ETMC publication.
where again the error includes both statistical and systematic uncertainties.

The IB correction $\delta R_{K\pi} = -0.0126(14)$, determined in Refs. [3,4] and adopted in Eqs. (2)-(3), stems from the sum of a QED and a strong IB terms, which are both prescription dependent as well as their sum and the ISOQCD value (10). Within the Gasser-Rusetsky-Seicemi prescription (see footnote 1) they are equal, respectively, to $-0.0062(12)$ and $-0.0064(7)$. Thus, for the ratio of kaon and pion leptonic decay constant including strong IB effects (which is prescription dependent) we get

$$\left(\frac{f_K}{f_\pi}\right)_{\text{ISOQCD}} = 1.1995(44), \quad (10)$$

For comparison, the $N_f = 2 + 1 + 1$ determinations, entering the FLAG-4 average [18], yield the value $(f_K/f_\pi) = 1.1932(19)$ [15,19,20], which is well consistent with our result (11). Once corrected for the strong IB effects obtained in Refs. [15,19,20], the FLAG-4 average becomes $(f_K/f_\pi)_{\text{ISOQCD}} = 1.1966(18)$, which agrees with our finding (10).

Taking the updated value $|V_{ud}| = 0.97370(14)$ from superallowed nuclear beta decays [2,21], Eqs. (3) and (10) yield the following value for the CKM element $|V_{us}|$:

$$|V_{us}| = 0.22472(24)_{\exp}(84)_{\text{th}} = 0.22472(87), \quad (12)$$

which is nicely consistent with the latest estimate $|V_{us}| = 0.2252(5)$ from leptonic modes provided by the PDG [2]. Correspondingly, using $|V_{ab}| = 0.00382(24)$ [2] the first-row CKM unitarity becomes

$$|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 0.99861(48), \quad (13)$$

which would imply a $\approx 3\sigma$ tension with unitarity from leptonic modes.

The plan of the paper is as follows.

In Sec. II some details of the ETMC gauge ensembles and of the simulations are illustrated, while a more complete description is provided in Appendix A. For each gauge ensemble the pion mass and decay constant are extracted from the relevant two-point correlation functions using a single exponential fit in the appropriate regions of large time distances. Alternatively, in Appendix B the extraction of the ground-state properties is performed through the multiple exponential procedure of Ref. [22]. For one gauge ensemble (cA211.12.48), because of a small deviation from maximal twist, the mass and the decay constant are corrected as described in Appendix C. In Sec. III the SU(2) ChPT predictions at next-to-leading order (NLO) for the pion decay constant $f_\pi$, including finite volume effects (FVEs), are presented. For the ensembles cB211.25.XX, sharing the same light-quark mass and lattice spacing and differing only for the lattice size $L$, the FVEs are investigated using both the NLO and the resummed NNLO formulas of Ref. [23]. In Sec. IV, adopting the physical value (6) at the pion point (4), we perform two determinations of the GF scale $\theta_0$ using the data for either $f_\pi$ or the quantity $X$ = $(f_\pi M_\pi^3)^{1/5}$, which is found to be less affected by statistical and systematic errors. The two determinations of $\theta_0$ agree very nicely, but the one based on the quantity $X$ turns out to be more precise by a factor of $\approx 2.5$. In the same way the other two GF scales $\sqrt{\theta_0}$ and $t_0/\theta_0$ are determined in Appendix D, where our calculations of the relative GF scales $\theta_0/a$, $\sqrt{\theta_0}/a$ and $t_0/(\theta_0a)$ at the physical pion point are also described. In Sec. V we analyze the data for the decay constant ratio $f_K/f_\pi$ using SU(2) ChPT. In Sec. VI the implications for $V_{us}$ and the first-row CKM unitarity are discussed. Finally, our conclusions are collected in Sec. VII.

II. ETMC ENSEMBLES

In this work we make use of the gauge ensembles produced recently by ETMC in ISOQCD with $N_f = 2 + 1 + 1$ flavors of Wilson-clover twisted-mass quark and described in Refs. [10,11]. The gluon action is the improved Iwasaki one [24], while the fermionic action includes a Clover term [25] with a coefficient fixed by its estimate in one-loop tadpole boosted perturbation theory [26]. Its inclusion turns out to be very beneficial for reducing cutoff effects, in particular on the neutral pion mass, thereby making numerically stable simulations close to the physical pion point [10].

The Wilson mass counterterms of the two degenerate light quarks as well as of the strange and charm quarks are chosen in order to guarantee automatic $O(a)$ improvement [27,28]. The masses of the strange and charm sea quarks are tuned to their physical values for each ensemble [10,11]. For the valence strange and charm sectors, a mixed action setup employing Osterwalder-Seiler fermions [29], with the same critical mass as determined in the unitary setup, has been adopted in order to avoid any undesired strange-charm quark mixing (through cutoff effects) and to preserve the automatic $O(a)$ improvement of physical observables [30].

Some properties of the ETMC ensembles, which are relevant for this work, are collected in Table I, while the simulation setup is described in detail in Appendix A. With respect to Ref. [11] two other dedicated gauge ensembles, cB211.25.24 and cB211.25.32, have been produced for the investigation of FVEs.

Note that in the case of the ensembles cB211.072.64 and cC211.06.80, corresponding, respectively, to a lattice spacing equal to $a \approx 0.082$ fm and $a \approx 0.069$ fm, the pion mass is simulated quite close to the physical ISOQCD value (4).
For each ensemble we compute the pion correlator given by

$$C_\pi(t) = \frac{1}{L^3} \sum_{x,z} \langle 0 | P_5(x) P_5^\dagger(z) | 0 \rangle \delta_{i(z-t)},$$

(14)

where $P_5(x) = \bar{q}_5(x) \gamma_5 q_5(x)$ is a local interpolating pion field. The Wilson parameters of the two mass-degenerate valence quarks are always chosen to have opposite values. In this way the squared pion mass differs from its continuum counterpart only by terms of $O(a^2 \mu_\pi)$ [27, 28].

At large time distances one has

$$C_\pi(t) \sim a, (T-t) \gg \frac{Z_\pi}{2M_\pi} \left[ e^{-M_\pi t} + e^{-M_\pi (T-t)} \right],$$

(15)

so that the pion mass $M_\pi$ and the matrix element $Z_\pi = |\langle \pi \bar{q}_5 \gamma_5 q_5 | 0 \rangle|^2$ can be extracted from the exponential fit given in the rhs of Eq. (15).

For maximally twisted fermions the value of $Z_\pi$ determines the pion decay constant $f_\pi$ without the need of the knowledge of any renormalization constant [27, 31], namely

$$a f_\pi = 2a \mu_\pi \frac{\sqrt{a^4 Z_\pi}}{a M_\pi \sinh(a M_\pi)}.$$  

(16)

The time intervals $[t_{\min}, t_{\max}]$ adopted for the fit (15) of the pion correlation function (14) as well as the extracted values of the pion mass and decay constant in lattice units are collected in Table II. As anticipated in the Introduction, in this work we will make also use of the data for the quantity $X_\pi$ defined as

$$X_\pi \equiv (f_\pi M_\pi^4)^{1/5},$$

(17)

which turns out to be less affected by lattice artifacts (see below Fig. 1 and later Sec. III C). The values of $X_\pi$ in lattice

| Ensemble     | $\beta$  | $V/a^4$   | $a/(\text{fm})$ | $a \mu_\pi$ | $M_\pi$ (MeV) | $L$(fm) | $M_\pi L$ | Conf |
|--------------|----------|-----------|-----------------|-------------|---------------|---------|----------|------|
| cA211.53.24  | 1.726    | $24^3 \times 48$ | 0.0947(4)       | 0.00530     | 346.4(1.6)    | 2.27    | 3.99     | 628  |
| cA211.40.24  | 1.726    | $24^3 \times 48$ | 0.00400         | 301.2(2.1)  | 2.27          | 3.47    | 662      |
| cA211.30.32  | 32$^3$   | 0.03030    | 261.1(1.1)      | 3.03         | 4.01          | 1237    |
| cA211.12.48  | 48$^3$   | 0.01120    | 167.1(0.8)      | 4.55         | 3.85          | 322     |
| cB211.25.24  | 1.778    | $24^3 \times 48$ | 0.0816(3)       | 0.00250     | 259.2(3.0)    | 1.96    | 2.57     | 500  |
| cB211.25.32  | 32$^3$   | 0.00250    | 253.3(1.4)      | 2.61         | 3.35          | 400     |
| cB211.25.48  | 48$^3$   | 0.00250    | 253.0(1.0)      | 3.92         | 5.02          | 314     |
| cB211.14.64  | 64$^3$   | 0.01400    | 189.8(0.7)      | 5.22         | 5.02          | 437     |
| cB211.072.64 | 64$^3$   | 0.00072    | 136.8(0.6)      | 5.22         | 3.62          | 374     |
| cC211.06.80  | 1.836    | $80^3 \times 160$ | 0.0694(3)       | 0.00060     | 134.2(0.5)    | 5.55    | 3.78     | 401  |
units are shown in the last column of Table II. The statistical errors of the lattice data lie in the range $0.1 \pm 1.1\%$ for the pion mass, in the range $0.2 \pm 0.8\%$ for the pion decay constant and in the range $0.1 \pm 0.9\%$ for the quantity $X_{\pi}$. We stress that in the case of the four ensembles $cA11.12.48$, $cB11.14.64$, $cB211.072.64$ and $cC211.06.80$ (which correspond to $M_{\pi} \lesssim 190$ MeV) the statistical errors of $aX_{\pi}$ turn out to be less than half of those of $af_{\pi}$.

An alternative way to extract the pion mass and decay constant is the ordinary differential equation (ODE) procedure of Ref. [22]. The results obtained by applying this method to the pion correlation function (14) are collected in Appendix B and found to be totally consistent with the findings of the single exponential fit (15) of Table II.

In the case of the ensemble $cA11.12.48$ due to a small deviation from maximal twist a correction needs to be applied. According to Appendix C the squared pion mass is left uncorrected, while for the pion decay constant $f_{\pi}$ we use the following formula

$$f_{\pi}|_{\text{corrected}} = f_{\pi} K_{\epsilon},$$

where

$$K_{\epsilon} = \sqrt{1 + (Z_{A} m_{\text{PCAC}}/\mu_{\epsilon})^{2}}.$$  

$\epsilon$ is the bare untwisted partially conserved axial current (PCAC) mass, $Z_{A}$ is the renormalization constant of the axial current and $\mu_{\epsilon}$ is the bare twisted mass of the light valence quarks. For the ensemble $cA11.12.48$ one has $Z_{A} \approx 0.75$ and $m_{\text{PCAC}}/\mu_{\epsilon} \approx -0.21(5)$ [11].

The statistical accuracy of the correlator (14) is significantly improved by using the so-called one-end stochastic method [32], which includes spatial stochastic sources at a single time slice randomly chosen. Statistical errors are evaluated using the jackknife procedure.

The results obtained for the pion decay constant $w_{0}f_{\pi}$ and for the quantity $w_{0}X_{\pi}$ [see Eq. (17)], are shown in Fig. 1 for all the gauge ensembles. By comparing the results corresponding to the ensembles $cB211.25.XX$ the FVEs are clearly visible in the case of $f_{\pi}$, while they are almost absent in the case of $X_{\pi}$. Moreover, also discretization effects in $X_{\pi}$ turn out to be smaller than those present in $f_{\pi}$.

III. THE PION DECAY CONSTANT $f_{\pi}$ WITHIN SU(2) CHPT

Within SU(2) ChPT [33] the pion decay constant $f_{\pi}$ is given at NLO by

$$f_{\pi} = f [1 - 2 \xi_{\epsilon} \log(\xi_{\epsilon}) + 2 A_{1} \xi_{\epsilon}],$$

(20)

where

$$\xi_{\epsilon} = \frac{2 B m_{\epsilon}}{(4 \pi f)^{2}}$$

(21)

with $m_{\epsilon} = m_{u} = m_{d}$ being the renormalized light-quark mass. In Eqs. (20)–(21) $B$ and $f$ are the LO SU(2) ChPT low-energy constants (LECs), while the coefficient $A_{1}$ is related to the NLO LEC $\tilde{\xi}_{4}^{\text{phys}}$ by

$$\tilde{\xi}_{4}^{\text{phys}} = A_{1} + 2 \log \left( \frac{4 \pi f}{M_{\pi}^{\text{SMQCD}}} \right).$$

(22)

For the squared pion mass one has at NLO

$$M_{\pi}^{2} = 2 B m_{\epsilon} [1 + \xi_{\epsilon} \log(\xi_{\epsilon}) + C_{1} \xi_{\epsilon}].$$

(23)
where the coefficient $C_1$ is related to the NLO LEC $\tilde{\varepsilon}_3^{\text{phys}}$ by

$$\tilde{\varepsilon}_3^{\text{phys}} = -C_1 + 2 \log \left( \frac{4\pi f}{M_{\text{ISOQCD}}} \right).$$  \hfill (24)

### A. Finite volume effects within NLO SU(2) ChPT

The structure of FVEs on the pion decay constant can be studied using SU(2) ChPT [33]. At NLO FVEs come entirely from the discretized sum over periodic momenta of the loop contributions. For a finite spatial volume $V = L^3$ one has

$$f_\pi(L) = f_\pi(L \to \infty) \left[ 1 + \Delta FVE(L) \right],$$  \hfill (25)

where $f_\pi(L \to \infty)$ is given by Eq. (20). The correction term $\Delta FVE(L)$ can be obtained from the chiral log in Eq. (20) via the following replacement

$$\xi f \log(\xi f) \to \xi f \tilde{g}_1(\lambda),$$  \hfill (26)

where $\lambda \equiv \sqrt{2 m f} L = \sqrt{\varepsilon f} 4\pi f L$ and

$$\tilde{g}_1(\lambda) = 4 \sum_{n=1}^\infty \frac{m(n)}{\sqrt{n} \lambda} K_1(\sqrt{n} \lambda),$$  \hfill (27)

with $K_1$ being a Bessel function of the second kind and $m(n)$ the multiplicities of a three-dimensional vector $\vec{n}$ having integer norm $n$ [i.e., $m(n) = \{6, 12, 8, 6, \ldots \}$]. At sufficiently large values of $\lambda$ the Bessel function can be replaced by its asymptotic expansion, which leads to

$$\tilde{g}_1(\lambda) \approx 4 \sqrt{\pi} \sum_{n=1}^\infty \frac{m(n)}{(\sqrt{n} \lambda)^{3/2}} e^{-\sqrt{n} \lambda}.$$  \hfill (28)

Thus, within NLO SU(2) ChPT the quantity $\Delta FVE(L)$ is given by

$$\Delta FVE(L) = -2\varepsilon f \tilde{g}_1(\lambda).$$  \hfill (29)

In the case of the squared pion mass one gets

$$M_\pi^2(L) = M_\pi^2(L \to \infty) \left[ 1 - \frac{1}{4} \Delta FVE(L) \right]^2,$$  \hfill (30)

where $M_\pi^2(L \to \infty)$ is given by Eq. (23).

### B. FVEs for the ensembles cB211.25.XX

In this section we study the FVEs on the pion mass and decay constant corresponding to the three ensembles cB211.25.XX of Table I, which share the same light-quark mass and lattice spacing but differ only for the lattice size $L$. We consider SU(2) ChPT both at NLO, i.e., the Gasser-Leutwyler (GL) formulas (25) and (30), and at NNLO + resummation, i.e., the Colangelo-Dürr-Haefeli (CDH) formulas [23]. The latter ones read as

$$f_\pi(L) = f_\pi(\infty) \left\{ 1 - 2\varepsilon f \tilde{g}_1(M_\pi L) + 2\varepsilon f^2 \left[ C_{f_s}^{(1)} \tilde{g}_1(M_\pi L) + C_{f_s}^{(2)} \tilde{g}_2(M_\pi L) + S_{f_s}^{(4)} \right] \right\},$$  \hfill (31)

$$M_\pi(L) = M_\pi(\infty) \left\{ 1 + \frac{1}{2} \varepsilon f \tilde{g}_1(M_\pi L) - \varepsilon f^2 \left[ C_{M_s}^{(1)} \tilde{g}_1(M_\pi L) + C_{M_s}^{(2)} \tilde{g}_2(M_\pi L) + S_{M_s}^{(4)} \right] \right\},$$  \hfill (32)

where $\tilde{g}_1$ is defined in Eq. (27), while

$$\tilde{g}_2(\lambda) \equiv 4 \sum_{n=1}^\infty \frac{m(n)}{\sqrt{n} \lambda} K_2(\sqrt{n} \lambda)$$  \hfill (33)

and

$$C_{f_s}^{(1)} = \frac{7}{9} + 2\varepsilon f + \frac{4}{3} \varepsilon f^2 - 3\varepsilon f^3,$$  \hfill (34)

$$C_{f_s}^{(2)} = \frac{112}{9} - \frac{8}{3} \varepsilon f + \frac{32}{3} \varepsilon f^2,$$  \hfill (35)

$$C_{M_s}^{(1)} = -\frac{55}{18} + 4\varepsilon f + \frac{8}{3} \varepsilon f^2 - \frac{5}{2} \varepsilon f^3 - 2\varepsilon f^4,$$  \hfill (36)

$$C_{M_s}^{(2)} = C_{f_s}^{(2)} = \frac{112}{9} - \frac{8}{3} \varepsilon f - \frac{32}{3} \varepsilon f^2,$$  \hfill (37)

with $\tilde{g}_i$ being NLO LECs that have a logarithmic pion mass dependence

$$\tilde{\varepsilon}_i = \varepsilon_i^{\text{phys}} + 2 \log \left( \frac{M_{\text{ISOQCD}}}{M_{\pi}} \right).$$  \hfill (38)

Finally, in Eqs. (31)–(32) the NNLO terms $S_{f_s}^{(4)}$ and $S_{M_s}^{(4)}$ are defined in the Appendix A of Ref. [23], but useful approximate analytic formulas are given by [23]

$$S_{f_s}^{(4)} = \left( \frac{4}{3} s_0 - \frac{13}{6} s_1 \right) \tilde{g}_1(M_\pi L)$$

$$- \left( \frac{40}{3} s_0 - 4 s_1 - \frac{8}{3} s_2 - \frac{13}{3} s_3 \right) \tilde{g}_2(M_\pi L),$$  \hfill (39)

$$S_{M_s}^{(4)} = \frac{13}{3} s_0 \tilde{g}_1(M_\pi L) - \left( \frac{40}{3} s_0 + \frac{32}{3} s_1 + \frac{26}{3} s_2 \right) \tilde{g}_2(M_\pi L),$$  \hfill (40)
with
\[ s_0 = 2 - \frac{\pi}{2}, \quad s_1 = \frac{\pi}{4} - \frac{1}{2}, \quad s_2 = \frac{1}{2} - \frac{\pi}{2}, \quad s_3 = \frac{3\pi}{16} - \frac{1}{2}. \]

The expansion variable \( \xi_\pi \) is defined as [23]
\[ \xi_\pi \equiv \frac{M_\pi^2}{(4\pi f_\pi)^2}. \]  
(42)

Different choices of the expansion variable are possible: one can replace \( f_\pi \) with the LO LEC \( f \) and/or replace \( M_\pi^2 \) with \( 2B_m \) (and correspondingly \( M_\pi L \) with \( \sqrt{2B_m L} \)) in the arguments of the functions \( \tilde{g}_1 \) and \( \tilde{g}_2 \). At NLO (i.e., for the GL formula) the above changes are equivalent, since any difference represents a NNLO effect. Instead, in the CDH formula additional terms appear at NNLO, which can be found in Ref. [34]. Here we consider only the alternative definition
\[ \xi_\pi \rightarrow \frac{M_\pi^2}{(4\pi f_\pi)^2}, \]  
(43)

which requires the addition to the rhs of Eq. (31) of the term \( f_\pi(\infty)\{8\tilde{g}_1^2 \tilde{g}_1(M_\pi L)\} \) and to the rhs of Eq. (31) of the term \( M_\pi(\infty)\{-2\tilde{g}_1^2 \tilde{g}_1(M_\pi L)\} \).

The GL formula corresponds to Eqs. (31)–(32) with all \( C_s \) and \( S_s \) set equal to zero. The CDH formula requires the knowledge of the values of the four NLO LECs \( \tilde{g}_i^{\text{phys}} \) with \( i = 1, \ldots, 4 \).

In Figs. 2 and 3 we compare the FVEs on the pion mass and decay constant for the three ensembles cB211.25.XX of Table I, evaluated using the GL and CDH formulas and assuming, respectively, the two definitions (42) and (43) for the expansion variable \( \xi_\pi \).

In the case of the CDH formula we adopt the following values of the NLO LECs: \( \tilde{\xi}_1^{\text{phys}} = -0.4, \tilde{\xi}_2^{\text{phys}} = 4.3, \tilde{\xi}_3^{\text{phys}} = 3.2 \) and \( \tilde{\xi}_4^{\text{phys}} = 4.4 \) (see Ref. [34]). The CDH results depend on such a choice and the sensitivity to the specific value of \( \tilde{\xi}_2^{\text{phys}} \) is illustrated in both figures by the green triangles.

It can be seen that the GL formula applied to both the pion mass and decay constant works quite well for \( M_\pi L \gtrsim 3 \), particularly in the case of the definition (43) of the expansion variable \( \xi_\pi \). The above condition is satisfied by all ETMC ensembles of Table I except the ensemble cB211.25.24.

C. FVEs for the quantity \( X_\pi \)

The interesting feature of the quantity \( X_\pi \), given by Eq. (17), is the absence of NLO chiral logs in its SU(2) ChPT expansion [see Eqs. (20) and (23)] when expressed in terms of quark masses. This implies the absence of FVEs at NLO, which in turn is also the origin of the small FVEs observed in the right panel of Fig. 1. This point is better elucidated in Fig. 4, where the results corresponding to the three ensembles cB211.25.XX differing only in the lattice size \( L \) are shown.

IV. DETERMINATION OF THE GF SCALE \( w_0 \) FROM THE PION DATA

Let us now apply the SU(2) ChPT predictions for interpolating the pion data to the physical pion mass and for extrapolating them to the continuum and infinite

![Graph showing values of the pion mass and decay constant](https://example.com/graph.png)

**FIG. 2.** Values of the pion mass (left panel) and pion decay constant (right panel) in lattice units for the three ensembles cB211.25.XX of Table I. The red circles represent the data versus \( M_\pi L \). The expansion variable \( \xi_\pi \) is given by Eq. (42). The blue squares correspond to the data corrected by the GL formula, while black diamonds represent the data corrected by the CDH formula, adopting for the NLO LECs the values \( \tilde{\xi}_1^{\text{phys}} = -0.4, \tilde{\xi}_2^{\text{phys}} = 4.3, \tilde{\xi}_3^{\text{phys}} = 3.2 \) and \( \tilde{\xi}_4^{\text{phys}} = 4.4 \). The green triangles correspond to the CDH correction assuming \( \tilde{\xi}_2^{\text{phys}} = 3.3 \). The horizontal dotted lines are the values of the pion mass and decay constant in the infinite volume limit.
volume limits. The goal is to determine the GF scale $w_0$ adopting the physical value (6) at the pion point (4) without resorting to the use of the renormalized light-quark mass. In the next two subsections we separately analyze the pion decay constant $f_\pi$ and the quantity $X_\pi$, respectively.

A. Determination of $w_0$ using the data for $f_\pi$

Using the simulated values $aM_\pi$ and $af_\pi$ in lattice units we evaluate the expansion variable $\xi_\pi$, defined (from now on) as

$$\xi_\pi \equiv \frac{(aM_\pi)^2}{(4\pi af_\pi)^2} = \frac{M_\pi^2}{(4\pi f_\pi)^2},$$

which depends on neither $w_0$ nor $w_0/a$. Then, for each gauge ensemble we calculate the FVE correction $\Delta_{FVE}(L)$ as

$$\Delta_{FVE}(L) = -2\xi_\pi g_1(M_\pi L)$$

and we reexpress the quantity $\xi_\pi$ [see Eq. (21)] in terms of the pion mass in the infinite volume limit [see Eq. (30)]

$$\xi_\pi \to \xi \equiv \frac{M_\pi^2(L \to \infty)}{(4\pi f_\pi)^2} = \frac{w_0 M_\pi^2}{(4\pi w_0 f_\pi)^2} \left[1 - \frac{1}{4} \Delta_{FVE}(L)^2\right],$$

where only the knowledge of $w_0/a$ is required to calculate the pion mass in units of $w_0$ and the free parameter becomes $w_0 f_\pi$.

We correct the data of the pion decay constant $w_0 f_\pi(L)$ for FVEs [see Eq. (25)], namely

$$w_0 f_\pi(L \to \infty) = \frac{w_0 f_\pi(L)}{1 + \Delta_{FVE}(L)^2}.\quad (47)$$

Analogously, for the pion mass $w_0 M_\pi(L)$ one has

$$w_0 M_\pi(L \to \infty) = \frac{w_0 M_\pi(L)}{1 - \frac{1}{4} \Delta_{FVE}(L)^2}.\quad (48)$$

The data for $w_0 f_\pi(L \to \infty)$ are fitted in terms of the variable $\xi$ [see Eq. (46)] using the following functional form

$$w_0 f_\pi(L \to \infty) = w_0 f \left[1 - 2\xi \log(\xi) + 2A_1\xi + A_2\xi^2\right] + \frac{a^2}{w_0^2} \left[D_0 + D_1\xi^2\right],$$

where with respect to a pure NLO ansatz we have added a possible higher-order term quadratic in $\xi$ as well as discretization effects proportional to $a^2$ and $a^2 M_\pi^2$.

The free parameters appearing in Eq. (49) are $w_0 f$, $A_1$, $A_2$, $D_0$, $D_1$, and their values are obtained from a standard $\chi^2$ minimization. From the value of $w_0 f$ the GF scale $w_0$ can be determined as follows. Let us consider the physical value of the variable (44), namely

FIG. 3. The same as in Fig. 2, but adopting the alternative definition (43) for the expansion variable $\xi_\pi$ and assuming $f = 122.5$ MeV and $a = 0.080$ fm.

FIG. 4. Values of $w_0 X_\pi$ versus $M_\pi L$ for the three ensembles cB211.25.XX differing only for the lattice size L. The dashed line indicates the simple exponential fit of the form $A[1 + Be^{-M_\pi L}/(M_\pi L)^{3/2}]$. 

C. ALEXANDROU et al. PHYS. REV. D 104, 074520 (2021)
Using Eq. (49) in the continuum limit the physical value of the variable \(46\), namely \(\xi^{\text{ISOQCD}} = (M_{\pi}^{\text{ISOQCD}}/4\pi f_{\pi})^2\), can be obtained by solving the relation

\[
d = \frac{1}{2}\log(\xi_{\text{ISOQCD}}) + A_1(\xi_{\text{ISOQCD}}^2) + A_2(\xi_{\text{ISOQCD}})^2. \tag{51}
\]

In this way the value of the LEC \(f\) in physical units is given by \(f = M_{\pi}^{\text{ISOQCD}}/(4\pi \sqrt{\xi^{\text{ISOQCD}}})\) and, therefore, \(w_0\) can be determined using the value of \(w_0 f\).

We start by considering a pure NLO fit, i.e., \(A_2 = 0\), including only the discretization effect proportional to \(a^2\), i.e., \(D_1 = 0\) in Eq. (49), and we apply it to all pion data up to \(M_\pi \approx 350\) MeV. The discretization coefficient \(D_0\) turns out to be quite small, \(D_0 = -0.05(4)\), and the corresponding \(\chi^2/(\text{d.o.f.})\) is equal to \(\chi^2/(\text{d.o.f.}) \approx 1.5\) for ten data points and three parameters. For the GF scale \(w_0\) we get \(w_0 = 0.1712(14)\) fm, which exhibits a \(\approx 0.8\%\) accuracy. However, a drastic improvement in the quality of the fit is obtained by including the discretization term proportional to \(a^2 M_\pi^2\), i.e., \(D_1 \neq 0\). This leads to \(\chi^2/(\text{d.o.f.}) \approx 0.2\), obtaining for \(w_0\) the value

\[
w_0 = 0.1740(15)\ \text{fm}, \tag{52}
\]

with \(f = 124.4(6)\) MeV and \(\xi_{\text{phys}} = 3.24(29)\) [see Eq. (22)]. The quality of the above fit is illustrated in Fig. 5. The result (52) is confirmed by a NLO fit without the discretization effects proportional to \(a^2 M_\pi^2\) (i.e., \(D_1 = 0\)), but limited to pion masses below \(\approx 190\) MeV (four data points and three parameters). In this case one gets \(w_0 = 0.1736(15)\) fm, \(f = 122.8(4)\) MeV, \(\xi_{\text{phys}} = 4.06(18)\) and \(\chi^2/(\text{d.o.f.}) \approx 0.1\).

In order to investigate systematic effects we include the quadratic term proportional to \(\ln(a)\), obtaining \(w_0 = 0.1736(16)\) fm, \(f = 123.4(7)\) MeV, \(\xi_{\text{phys}} = 3.26(30)\) and \(\chi^2/(\text{d.o.f.}) \approx 0.2\), and we check also the impact of FVEs by multiplying the correction \(\Delta F^{\text{FVE}}(L)\) of Eq. (45) by a factor \(\kappa_{\text{FVE}}\) used as a further free parameter in the NLO fit. The factor \(\kappa_{\text{FVE}}\) turns out to be consistent with unity, \(\kappa_{\text{FVE}} = 1.20(18)\), and we get \(w_0 = 0.1743(16)\) fm, \(f = 124.5(6)\) MeV, \(\xi_{\text{phys}} = 3.18(30)\), and \(\chi^2/(\text{d.o.f.}) \approx 0.1\).

After averaging the above results our determinations of \(w_0\), \(f\) and \(\xi_{\text{phys}}\) based on the analysis of \(f_{\pi}\) are

\[
w_0 = 0.17390(157)_{\text{stat+fit}}(30)_{\text{syst}}(160)\ \text{fm}, \tag{53}
\]

\[
f = 124.0(6)_{\text{stat+fit}}(7)_{\text{syst}}(9)\ \text{MeV}, \tag{54}
\]

\[
\xi_{\text{phys}} = 3.44(27)_{\text{stat+fit}}(36)_{\text{syst}}(45), \tag{55}
\]

FIG. 5. Values of the pion decay constant \(w_0 f_\pi\) corrected for FVEs according to Eq. (47) (open markers) and compared to the results of the NLO ChPT fit corresponding to \(A_2 = 0\) in Eq. (49) applied to all data points \((M_\pi \lesssim 350\) MeV). The solid line represents the results of the fit in the continuum limit, while the dashed lines correspond to the fit evaluated at each value of \(\beta\). The cross represents the result at the physical pion point (4) corresponding to the value \(w_0 = 0.1740(15)\) fm, obtained as described in the text.
where \( (\Delta)^{\text{stat+fit}} \) incorporates the uncertainties induced by both the statistical errors and the fitting procedure itself, \( (\Delta)^{\text{syst}} \) corresponds to the uncertainty related to chiral interpolation, discretization and finite-volume effects, while the last error is their sum in quadrature. More precisely, the various systematic uncertainties are estimated by considering the results obtained with \( A_2 = 0 \) or \( A_2 \neq 0 \) in the case of the chiral extrapolation, with \( D_1 \neq 0 \) or \( D_1 = 0 \) (but limited to \( M_x < 190 \) MeV) for the discretization effects and with \( \kappa_{\text{FVE}} = 1 \) or \( \kappa_{\text{FVE}} \neq 1 \) for the FVEs.

### B. Determination of the GF scale \( w_0 \) using the data for \( X_x \)

In this section we illustrate the results of the analysis of the lattice data for the quantity \( w_0 X_x \) adopting the following fitting function

\[
w_0 X_x = (w_0 f_x) \left\{ (4\pi)^4 \xi^2 [1 - 2\xi \log(\xi) + 2A_1 \xi + A_2 \xi^2 + a^2 (D_0 + D_1(\xi))] \right\}^{1/5} \cdot (1 + F_{\text{FVE}} \xi^2 e^{-M_x L}/(M_x L)^{3/2}),
\]

where the variable \( \xi \) is defined by Eq. (46), given in terms of the pion mass corrected for the FVEs using the GL formula (45), and the coefficient \( A_1 \) is related to the LEC \( \bar{\varepsilon}_4^{\text{phys}} \) by Eq. (22). In Eq. (56) we have taken into account that the FVEs on \( X_x \) start only at NNLO, i.e., at order \( O(\xi^2) \). Their impact is obtained by including \( (F_{\text{FVE}} 
eq 0) \) or by excluding \( (F_{\text{FVE}} = 0) \) the higher order FVEs. Moreover, the NLO chiral log is present only because we employ meson masses and it would disappear if the light-quark mass would be instead considered (in this case the linear coefficient \( A_1 \) provides directly the difference \( \bar{\varepsilon}_4^{\text{phys}} - \bar{\varepsilon}_3^{\text{phys}} \)).

We have performed several fits similar to those adopted in Sec. IV A and the corresponding results are collected in Table III. The quality of the NLO fit with \( D_1 \neq 0 \) is illustrated in Fig. 6, where it is also clearly visible the presence of discretization effects proportional to \( a^2 M_x^2 \), as already observed in the case of \( w_0 f_x \) (see Fig. 5). We stress that for both quantities, \( w_0 f_x \) and \( w_0 X_x \), the inclusion of a discretization term proportional to \( a^2 M_x^2 \) leads to higher values of \( w_0 \). This result is reassuringly confirmed also by a NLO fit without such a discretization term (i.e., \( D_1 = 0 \)), but limited to pion masses below \( \approx 190 \) MeV (see the fourth row of Table III).

By averaging the last four results of Table III one has

\[
w_0 = 0.17383(57)_{\text{stat+fit}}^{(26)}_{\text{syst}} \text{ fm}, \quad (57)
\]

\[
f = 124.0(1.2)_{\text{stat+fit}}^{(0.7)}_{\text{syst}} \text{ MeV}, \quad (58)
\]

\[
\bar{\varepsilon}_4^{\text{phys}} = 3.37(27)_{\text{stat+fit}}^{(38)}_{\text{syst}} \text{ MeV}. \quad (62)
\]

TABLE III. Results for \( w_0 \) obtained by fitting the lattice data for \( w_0 X_x \) using Eq. (56) and adopting the ISOQCD values (4) and (6) for fixing the lattice scale at the physical pion point.

| \( A_2 \neq 0 \) | \( D_1 \neq 0 \) | \( F_{\text{FVE}} \neq 0 \) | Range of \( M_x \) | \( w_0 \) (fm) | \( f \) (MeV) | \( \bar{\varepsilon}_4^{\text{phys}} \) | \( \chi^2/(d.o.f.) \) |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| No              | No              | No              | \(< 350 \) MeV   | 0.17213 (47)    | 122.4 (0.7)     | 4.23 (9)        | 0.26            |
| No              | Yes             | No              | \(< 350 \) MeV   | 0.17394 (58)    | 124.4 (1.2)     | 3.24 (29)       | 0.03            |
| No              | No              | No              | \(< 190 \) MeV   | 0.17343 (53)    | 122.8 (1.0)     | 4.04 (16)       | 0.05            |
| Yes             | Yes             | No              | \(< 350 \) MeV   | 0.17378 (56)    | 124.3 (1.3)     | 3.27 (30)       | 0.04            |
| No              | Yes             | Yes             | \(< 350 \) MeV   | 0.17415 (61)    | 124.6 (1.3)     | 3.15 (35)       | 0.02            |

\[
\bar{\varepsilon}_4^{\text{phys}} = 3.43(28)_{\text{stat+fit}}^{(36)}_{\text{syst}} \text{ MeV}. \quad (59)
\]

which nicely agree with the corresponding results obtained by the analysis of \( f_x \) given by Eqs. (53)–(55). Note that the determination of \( w_0 \) using \( X_x \) is more precise than the one from \( f_x \) by a factor equal to \( \approx 2.5 \).

Our result (57) is slightly larger than both the MILC result \( w_0 = 0.1714^{+2.15}_{-1.7_{\text{hf}}} \) fm from Ref. [14] and the HPQCD result \( w_0 = 0.1715(9) \) fm from Ref. [15], obtained using the values (6) to set the lattice scale. Within \( \pm 1 \) standard deviations it is consistent with the recent, precise BMW determination \( w_0 = 0.17236(70) \), obtained in Ref. [16] using the \( \Omega^- \)-baryon mass to set the lattice scale. Furthermore, the difference with the recent result \( w_0 = 0.1709(11) \) fm, obtained in Ref. [17] using the \( \Omega^- \)-baryon mass to set the lattice scale, is within \( \pm 2 \) standard deviations.

In Appendix D 2 the procedure used in this section to determine the GF scale \( w_0 \) is repeated in the case of the scales \( \sqrt{t_0} \) and \( t_0/w_0 \), obtaining

\[
\sqrt{t_0} = 0.14436(54)^{\text{stat+fit}}_{\text{syst}}(30)_{\text{[61]}} \text{ fm}, \quad (60)
\]

\[
f = 124.1(1.2)^{\text{stat+fit}}_{\text{syst}}(0.7)_{\text{[1.4]}} \text{ MeV}, \quad (61)
\]

\[
\bar{\varepsilon}_4^{\text{phys}} = 3.37(27)^{\text{stat+fit}}_{\text{syst}}(38)_{\text{[47]}}, \quad (62)
\]
Our finding (60) is larger than the MILC result \( t_0/w_0 = 0.1416^{+0.08}_{-0.08} \) fm from Ref. [14] and the HPQCD result \( t_0 = 0.1420(8) \) fm from Ref. [15], while within \( \pm 1.5 \) standard deviations it is consistent with the recent result \( t_0 = 0.1422(14) \) fm from Ref. [17].

The values of the lattice spacing corresponding to the three GF scales are collected in Table XII of Appendix D.

\[
V. \text{SU(2) CHPT ANALYSIS OF } f_K/f_\pi
\]

The kaon correlator

\[
C_K(t) = \frac{1}{L^3} \sum_{x,z} \langle 0 | \bar{q}_s(x) \gamma_5 q_r(x) \bar{q}_r(z) \gamma_5 q_s(z) | 0 \rangle \delta_{t(t_x-t_z)}.
\]

\[(66)\]

Our finding (60) is larger than the MILC result \( t_0/w_0 = 0.11969(52)_{\text{stat}+\text{fit}}(33)_{\text{syst}} \) fm, [62]

\[
t_0/w_0 = 0.11969(52)_{\text{stat}+\text{fit}}(33)_{\text{syst}} \text{ fm}, \quad (63)
\]

\[
f = 124.2(1.4)_{\text{stat}+\text{fit}}(0.8)_{\text{syst}} \text{ MeV}, \quad (64)
\]

\[
\bar{\epsilon}_{s}^{\text{phys}} = 3.31(27)_{\text{stat}+\text{fit}}(40)_{\text{syst}} \text{MeV}, \quad (65)
\]
At large time distances one has
\[
 C_K(t) \approx \frac{Z_K}{a^2 M_K^2} e^{-M_K t} + e^{-M_K(T-t)}, \quad (67)
\]
which allows the extraction of the kaon mass \(M_K\) and the matrix element \(Z_K = \langle \langle K | q_1 T \gamma_5 q_2 | 0 \rangle \rangle^2\) from the exponential fit given in the rhs of Eq. (67). The kaon decay constant \(f_K\) is given by
\[
 a f_K = (a \mu_\pi + a \mu_s) \sqrt{\frac{a^4 Z_K}{a M_K \sinh(a M_K)}} \quad (68)
\]
and, using the pion data (16) for \(f_\pi\), the ratio \(f_K/f_\pi\) is evaluated at each simulated strange bare quark mass. The time intervals \([t_{\text{min}}, t_{\text{max}}]\) adopted for the fit (67) of the kaon correlation function (66) are the same as those used for the case of the pion correlator, collected in Table II.

As in the case of the pion data (see Sec. II), due to a small deviation from maximal twist, a correction should be applied to observables of the ensemble cA211.12.48. We use the following formula (see Appendix B)
\[
 f_K^{\text{corrected}} \simeq f_K \cdot K_f, \quad (69)
\]
with
\[
 K_f = \frac{1}{\cos[(\theta_3 + \theta_4)/2]}, \quad (70)
\]
where, we remind,
\[
 \frac{1}{\cos(\theta_4)} \equiv K_i = \sqrt{1 + (Z_A m_{\text{PCAC}}/\mu_i)^2}, \quad (71)
\]
m_{\text{PCAC}} is the bare untwisted PCAC mass, \(Z_A\) is the renormalization constant of the axial current and \(\mu_i\) is the bare twisted mass of the valence quarks. In the degenerate case \(m_i = m_\pi\) one gets \(K_f = K_\pi\), i.e., Eq. (19), while for \(m_i \gg m_\pi\) one has \(K_f \simeq 1/\cos(\theta_4/2)\).

Since the LECs of the SU(2) ChPT depend on the value of the (renormalized) strange quark mass \(m_s\), we need to interpolate the ratio \(f_K/f_\pi\) at an approximately fixed value of \(m_s\). To this end we take advantage of the fact that the meson mass combination \(2M_K^2 - M_\pi^2\) is proportional to \(m_s\) at LO in ChPT. Thus, for each gauge ensemble, adopting a simple quadratic spline, the lattice data for \(f_K/f_\pi\) are interpolated at a reference kaon mass given by
\[
 M_K^{\text{ref}} = \sqrt{(M_K^{\text{ISOQCD}})^2 + \left(\frac{M_\pi^2 - (M_K^{\text{ISOQCD}})^2}{2}\right)^2}, \quad (72)
\]
with \(M_K^{\text{ISOQCD}}\) and \(M_\pi^{\text{ISOQCD}}\) chosen as in Eqs. (4) and (5), respectively. The physical units for \(M_\pi\) (and consequently for \(M_K^{\text{ref}}\)) are obtained by using the results for the lattice spacing given in Table XII of Appendix D 2 for each choice of the GF scale. In what follows we make use of our determination (57) of the GF scale \(\Lambda_0\). In this way the renormalized strange quark mass \(m_s^{\text{ref}}\) corresponding to \(M_K^{\text{ref}}\) is kept close to its physical value.

The results obtained for the ratio \(f_K/f_\pi\) interpolated at the kaon reference mass (72) are shown in Fig. 7 for all the ETMC gauge ensembles. The statistical errors of the data lie in the range \(0.1 \pm 0.6\%\).

We now apply the correction for FVEs using the GL formula and the expansion variable \(\xi_\pi\) defined as

![FIG. 7. Values of the ratio \(f_K/f_\pi\) interpolated at the kaon reference mass (72) versus the squared pion mass. The vertical dotted line indicates the location of the physical ISOQCD point (4). For the ensemble cA211.12.48 the corrected value of the ratio \(f_K/f_\pi\), obtained using Eqs. (18) and (69), is considered.](074520-12)
ansatz $R$ at the physical ISOQCD point, given by Eqs. (4) and (5), obtained using the fitting function (77).

Finally, in terms of the variable $\xi$, defined in Eq. (46), the data for $(f_K/f_\pi)(L \to \infty)$ are fitted using the following ansatz

$$\frac{f_K}{f_\pi}(L \to \infty) = R_0 \left[ 1 + \frac{5}{4} \xi \log(\xi) + R_1 \xi + R_2 \xi^2 \right]$$

where with respect to the well-known SU(2) ChPT prediction at NLO a quadratic term in $\xi$ as well as discretization effects proportional to $a^2$ and $a^2 M^2$ have been added.

The free parameters appearing in Eq. (77) are $R_0, R_1, R_2, \tilde{D}_0, \tilde{D}_1$, and their values are obtained by a straightforward $\chi^2$-minimization procedure. We have performed several fits based on Eq. (77) and the results for the ratio $(f_K/f_\pi)^{\text{ISOQCD}}$ at the physical pion point (4) are collected in Table IV.

The quality of the NLO fit with $R_2 = \tilde{D}_1 = 0$ is illustrated in Fig. 8. It can be seen that FVEs are properly taken care of and that discretization effects are quite small. As a check of the impact of FVEs we multiply the GL correction in Eq. (76) by a factor $\kappa_{\text{FVE}}$, which is treated as a further free parameter in the NLO fit. The factor $\kappa_{\text{FVE}}$ turns out to be consistent with unity, $\kappa_{\text{FVE}} = 1.19(24)$, and the NLO result $(f_K/f_\pi)^{\text{ISOQCD}} = 1.1995(35)$ is reassuringly confirmed.

Putting together all the various results we obtain

$$\left(\frac{f_K}{f_\pi}\right)_{\text{ICQD}}^{\text{stat}+\text{fit}} = 1.1995(44)_{\text{stat}+\text{fit}}(7)_{\text{syst}}[44],$$

where we remind that $(\cdot)_{\text{stat}+\text{fit}}$ incorporates the uncertainties induced by both the statistical errors and the fitting procedure itself. Adopting the results of the ODE procedure (see Appendix B) for the extraction of the pion and kaon masses and decay constants the analysis of the ratio $f_K/f_\pi$ yields

$$\left(\frac{f_K}{f_\pi}\right)_{\text{ICQD}}^\text{NLO} = 1.1994(43)_{\text{stat}+\text{fit}}(7)_{\text{syst}}[43],$$

which compares very well with the finding (78).

![Figure 8](link)

**FIG. 8.** Values of the ratio $f_K/f_\pi$ corrected for FVEs according to Eq. (76) (open markers) compared to the results of the NLO fit corresponding to $R_2 = \tilde{D}_1 = 0$ in Eq. (77) applied to all pion masses ($M_\pi \lesssim 350$ MeV). The solid line represents the results of the fit in the continuum limit, while the dashed lines correspond to the fit evaluated at each value of $\beta$. The cross represents the result at the physical pion point (4).
The present result (78) improves drastically the precision of the previous \( N_f = 2 + 1 + 1 \) ETMC determination \( \langle f_K/f_\pi \rangle_{\text{ISOQCD}} = 1.188(15) \) [19] by a factor of \( \approx 3.5 \) reaching the level of \( \approx 0.4\% \). For comparison, the \( N_f = 2 + 1 + 1 \) determinations, entering the FLAG-4 average [18] and corrected for strong IB effects, yield a consistent value within the uncertainties, namely \( \langle f_K/f_\pi \rangle_{\text{ISOQCD}} = 1.1966(18) \) [15,19,20]. Our finding (78) is also in good agreement with the recent determination \( \langle f_K/f_\pi \rangle_{\text{ISOQCD}} = 1.1964(44) \) obtained in Ref. [35] adopting the same ISOQCD prescription in a mixed-action approach (domain-wall valence quarks with staggered sea quarks).

VI. IMPLICATIONS FOR \( V_{us} \) AND THE FIRST-ROW CKM UNITARITY

Inserting our ISOQCD result (78) into Eq. (3) the ratio of the CKM entries \( V_{us} \) and \( V_{ud} \) is given by

\[
\left| \frac{V_{us}}{V_{ud}} \right| = 0.23079(24) \exp(87)_{\text{th}} = 0.23079(90). \tag{80}
\]

Using the value \( |V_{ud}| = 0.97370(14) \) from superallowed nuclear beta decays [2,21], which updates the old result \( V_{ud} = 0.97420(21) \) from Ref. [8], Eq. (3) yields the following value for the CKM element \( |V_{us}| \):

\[
|V_{us}| = 0.22472(24) \exp(84)_{\text{th}} = 0.22472(87). \tag{81}
\]

which is in good agreement with the latest estimate \( |V_{us}| = 0.2252(5) \) from leptonic modes provided by the PDG [2].

Using the values \( |V_{ub}| = 0.00382(24) \) [2] and \( |V_{ud}| = 0.97370(14) \) [2,21] our result (81) implies for the unitarity of the first-row of the CKM matrix the value

\[
|V_{ud}|^2 + |V_{us}|^2 + |V_{ab}|^2 = 0.99861(48), \tag{82}
\]

which in turn would imply a \( \approx 3\sigma \) tension with unitarity from leptonic modes. Had we used the result \( V_{ud} = 0.97420(21) \) from Ref. [8] the first-row CKM unitarity would be fulfilled within one standard deviation, i.e., within a precision of \( \approx 0.5 \) permil.

Another source of information on \( V_{us} \) is represented by the semileptonic \( K_{l3} \) decay. In this case the relevant hadronic quantity is the vector form factor at zero momentum transfer \( f_+(0) \). From the high-precision experimental data on \( K_{l3} \) decays one has \( V_{us}f_+(0) = 0.2165(4) \) [36].

Using the ETMC determination \( f_+(0) = 0.9709(46) \) obtained with Wilson twisted-mass quarks in Ref. [37], one gets the semileptonic result \( V_{us} = 0.2230(11) \) to be compared with the leptonic one given in Eq. (81). The above finding is combined with Eq. (80) to obtain the red ellipse in Fig. 9, which represents a 68% likelihood contour. For comparison the blue ellipse corresponds to the FLAG-4 contour for \( N_f = 2 + 1 + 1 \) [18], defined by the bands corresponding to \( V_{us} = 0.2231(7) \) and \( V_{us}/V_{ud} = 0.2313(5) \). The two determinations of \( V_{ud} \) obtained in Refs. [8] and [21] are also shown. Finally, the dotted line represents the correlation between \( V_{us} \) and \( V_{ud} \) when the CKM matrix is taken to be unitary.

VII. CONCLUSIONS

We have presented a determination of the ratio of kaon and pion leptonic decay constants in ISOQCD, \( f_K/f_\pi \), adopting the gauge ensembles produced by ETMC with \( N_f = 2 + 1 + 1 \) flavors of Wilson-clover twisted-mass quarks, including configurations close to the physical point for all dynamical flavors.
The simulations are carried out at three values of the lattice spacing ranging from \(\sim 0.068\) to \(\sim 0.092\) fm with linear lattice size up to \(L \sim 5.5\) fm. The scale is set using the value the pion decay constant \(f_\pi^{\text{ISOQCD}} = 130.4(2)\) MeV taken from Ref. [7]. Two observables, \(f_\pi\) and \((f_\pi M_\pi^2)^{1/3}\), have been analyzed within the framework of SU(2) ChPT without making use of renormalized quark masses. The latter quantity is found to be marginally affected by lattice artifacts and provides a precise determination of the GF scales, namely \(w_0 = 0.17383(63)\) fm, \(\sqrt{t_0} = 0.14436(61)\) fm and \(t_0/w_0 = 0.11969(62)\) fm.

As for the decay constant ratio \(f_K/f_\pi\) we get at the physical ISOQCD point, defined by Eqs. (4)–(6), the result

\[
\left( \frac{f_K}{f_\pi} \right)^{\text{ISOQCD}} = 1.1995(44),
\]

(83)

where the error includes both statistical and systematic uncertainties in quadrature. Our result (83) agrees nicely with the recent \(N_f = 2 + 1 + 1\) determinations, entering the FLAG-4 average [18] and corrected for strong IB effects, namely \(\left( f_K/f_\pi \right)^{\text{ISOQCD}} = 1.1966(18)\) [15,19,20].

Taking the updated value \(|V_{ud}| = 0.97370(14)\) from superallowed nuclear beta decays [2], Eqs. (3) and (83) yield the following value for the CKM element \(|V_{us}|\):

\[
|V_{us}| = 0.22472(24)_{\text{exp}}(84)_{\text{th}} = 0.22472(87),
\]

(84)

which is nicely consistent with the latest estimate \(|V_{us}| = 0.2252(5)\) from leptonic modes provided by the PDG [2]. Correspondingly, using \(|V_{ab}| = 0.00382(24)\) [2] the first-row CKM unitarity becomes

\[
|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 0.99861(48),
\]

(85)

which would imply a \(\approx 3\sigma\) tension with unitarity from leptonic modes.

We thank all the ETMC members for a very productive collaboration. We warmly thank G. C. Rossi for his continuous support and for a careful reading of the manuscript. We are grateful to B. Józ for his kind support with our refactoring and extension of the QPhiX lattice QCD library. We acknowledge PRACE (Partnership for Advanced Computing in Europe) for awarding us access to the high-performance computing system Marconi and Marconi100 at CINECA (Consorzio Interuniversitario per il Calcolo Automatico dell’Italia Nord-orientale) under the Grants No. Pra17-4394, No. Pra20-5171, and No. Pra22-5171, and CINECA for providing us CPU time under the specific initiative INFN-LQCD123. We also acknowledge PRACE for awarding us access to HAWK, hosted by HLRS, Germany, under the Grant with Acid No. 33037. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. ([38]) for funding the project pr74yo by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre ([39]). Some of the ensembles for this study were generated on Jureca Booster [40] and Juwels [41] at the Jülich Supercomputing Centre (JSC) and we gratefully acknowledge the computing time granted there by the John von Neumann Institute for Computing (NIC). The project has received funding from the Horizon 2020 research and innovation program of the European Commission under the Marie Skłodowska-Curie Grant Agreement No. 642069 (HPC-LEAP) and under Grant Agreement No. 765048 (STIMULATE). The project was funded in part by the NSFC (National Natural Science Foundation of China) and the DFG (Deutsche Forschungsgemeinschaft, German Research Foundation) through the Sino-German Collaborative Research Center Grant No. TRR110 “Symmetries and the Emergence of Structure in QCD” (NSFC Grant No. 12070131001, DFG Project-ID 196253076—TRR 110). R. F. acknowledges the University of Rome Tor Vergata for the support granted to the project PLNUGAMMA. F. S. and S. S. are supported by the Italian Ministry of Research (MIUR) under Grant No. PRIN 20172LNEEZ. F. S. is supported by INFN under Grant73/CALAT. P. D. acknowledges support form the European Unions Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 813942 (EuroPLEX) and from INFN under the research project INFN-QCDLAT. S. B. and J. F. are supported by the H2020 project PRACE6-IP (Grant Agreement No. 82376) and the COMPLEMENTARY/0916/0015 project funded by the Cyprus Research Promotion Foundation. The authors acknowledge support from project NextQCD, cofunded by the European Regional Development Fund and the Republic of Cyprus through the Research and Innovation Foundation (EXCELLENCE/0918/0129).

**APPENDIX A: ALGORITHMIC DETAILS AND PARAMETERS FOR THE ETMC GAUGE ENSEMBLES**

In Appendix A, we present the algorithmic setup employed for the generation of our ensembles of gauge configurations, while the simulation parameters are given in Table V.

1. Integrator setups

In the generation of gauge ensembles via the hybrid Monte Carlo algorithm, the effective lattice action can be represented by a sum over monomials corresponding to different contributions to the partition function as defined below. In the integration of the equations of motion, the forces contributed by the different monomials differ by orders of magnitude, allowing them to be integrated on different time scales accordingly, as detailed in Table VI below.
We define below the different types of monomials that we employ in our effective lattice action to simulate QCD using $N_f = 2 + 1 + 1$ twisted mass clover fermions.

**Gauge** $\text{gau}(\beta, c_1)$

$$\frac{\beta}{3} \sum_x \left( c_0 \sum_{\gamma = 1}^4 \{1 - \text{ReTr}(U_{1 \times 1}^{1 \times 1})\} + c_1 \sum_{\gamma = 1}^4 \{1 - \text{ReTr}(U_{1 \times 2}^{1 \times 2})\} \right).$$

(A1)

with $c_0 = (1 - 8c_1)$, for the Iwasaki action used here [24], $c_1 = -0.331$.

**Degenerate determinant** $\text{det}(\rho)$ The action contribution of a degenerate doublet of clover-improved twisted mass quarks is given by

$$S[\gamma, \tilde{\chi}, U] = \sum_x \left\{ \tilde{\chi}(x) \left[ 1 + 2k c_{SW} T + 2 i k \mu \gamma_5 \tau^3 \right] \chi(x) - \kappa \tilde{\chi}(x) \sum_{\mu=1}^4 U_{\mu}(x) (r - \gamma_\mu) \chi(x + a\mu) \ight. \\
+ U_\mu^\dagger(x - a\bar{\mu}) (r + \gamma_\mu) \chi(x - a\bar{\mu}) \right\} \\
= \sum_{x,y} \tilde{\chi}(x) M_{xy} \chi(y),$$

(A2)

in the twisted basis and in the hopping parameter normalization, where $T$ is the clover term. In our simulations we use the conventional value $r = 1$.

For convenience, we define $\tilde{\mu} \equiv 2\kappa\mu$ and absorb $2k c_{SW}$ into $T$, defining the two-flavor operator

$$Q = \gamma_5 M = \begin{pmatrix} Q_+ \\ Q_- \end{pmatrix},$$

(A3)

and the Hermitian operator $Q_{sw} = \gamma_5 D_{sw}$, where in turn $D_{sw}$ is the clover-improved Wilson Dirac operator. We then have $Q_{\pm} = Q_{sw} \pm i\tilde{\mu}$, such that $Q_+ = 0$ and $Q_+ Q_- = Q_{sw}^2 + 2\tilde{\mu}^2$. The contribution to the partition function of the mass-degenerate (light) twisted mass quark doublet is thus given by $\text{det}(Q_+, Q_-) = \text{det}(Q_{sw}^2 + 2\tilde{\mu}^2)$. An even-odd Schur decomposition of the submatrices $Q_{\pm}$ then gives

$$Q_{\pm} = \gamma_5 \begin{pmatrix} 1 + T_{ee} \pm i\tilde{\mu} \gamma_5 & M_{eo} \\ M_{oe} & 1 + T_{oo} \pm i\tilde{\mu} \gamma_5 \end{pmatrix}$$

$$= \gamma_5 \begin{pmatrix} M_{o\pm}^e & M_{eo} \\ M_{oe} & M_{oo}^e \end{pmatrix}$$

$$= \begin{pmatrix} \gamma_5 M_{o\pm}^e & 0 \\ \gamma_5 M_{oe} & 1 \end{pmatrix} \begin{pmatrix} 1 & (M_{o\pm}^e)^{-1} M_{eo} \\ 0 & \gamma_5 (M_{oo}^e - M_{oe} (M_{o\pm}^e)^{-1} M_{eo}) \end{pmatrix},$$

(A4)

from which we obtain $\hat{Q}_{\pm}$ defined only on the odd sites of the lattice.

---

**Table V.** Simulation parameters for the ensembles used for this study. Please refer to Appendix A1 for details on the integrator setup.

| Ensemble   | $\beta$ | $c_{SW}$ | $\kappa$ | $V/a^4$ | $a_{\mu}$ | $a_{\mu}$ | $a_{\delta}$ | $\lambda_{\min}$ | $\lambda_{\max}$ |
|------------|--------|---------|---------|--------|---------|---------|---------|--------|---------|
| cA211.53.24 | 1.726  | 1.74    | 0.1400645 | 243 $\times$ 48 | 0.00530 | 0.1408  | 0.1521  | 0.0000376 | 4.7      |
| cA211.40.24 | 2.43   | 1.74    | 0.0900400 | 243 $\times$ 48 | 0.00400 | 0.1248664 | 0.131052 | 0.0000344 | 4.3      |
| cA211.30.32 | 1.726  | 1.69    | 0.1394267 | 243 $\times$ 48 | 0.00250 | 0.1246864 | 0.131052 | 0.0000344 | 4.3      |
| cA211.12.48 | 1.400650 | 483 $\times$ 96 | 0.00120 |
| cB211.25.24 | 1.778  | 1.69    | 0.1394267 | 243 $\times$ 48 | 0.00250 | 0.1246864 | 0.131052 | 0.0000344 | 4.3      |
| cB211.25.32 | 1.778  | 1.69    | 0.1394267 | 243 $\times$ 48 | 0.00250 | 0.1246864 | 0.131052 | 0.0000344 | 4.3      |
| cB211.25.48 | 1.778  | 1.69    | 0.1394267 | 243 $\times$ 48 | 0.00250 | 0.1246864 | 0.131052 | 0.0000344 | 4.3      |
| cB211.14.64 | 1.778  | 1.69    | 0.1394267 | 243 $\times$ 48 | 0.00250 | 0.1246864 | 0.131052 | 0.0000344 | 4.3      |
| cB211.07.64 | 1.778  | 1.69    | 0.1394267 | 243 $\times$ 48 | 0.00250 | 0.1246864 | 0.131052 | 0.0000344 | 4.3      |
| cC211.06.80 | 1.836  | 1.6452  | 0.13875285 | 803 $\times$ 160 | 0.00060 | 0.106586 | 0.107146 | 0.0000376 | 4.7      |
|             |        |         |          |        |         |         |         |        |         |
### TABLE VI. Integrators setup used for the ensembles analyzed in this study. The number of timescales and trajectory length, \( \tau \), used for each ensemble are indicated in the respective headers.

| Id     | Type     | \( N_s \) | \( \lambda \) | Monomials                                  |
|--------|----------|-----------|-------------|--------------------------------------------|
| cA211.53.24, 5 timescales, \( \tau = 1.0 \) | 2MN       | 1         | 0.193       | \( \text{gau}(\beta, c_1) \)               |
|        |          | 1         | 0.195       | \( \text{det}(0.1) \)                      |
|        |          | 2         | 0.197       | \( \text{detrat}(0.02,0.1), \text{rat}(0.5) \) |
|        |          | 3         | 0.200       | \( \text{detrat}(0.003,0.02), \text{rat}(6,7) \) |
|        |          | 4         | 0.205       | \( \text{detrat}(0,0.003), \text{rat}(8,9) \) |
| cA211.40.24, 5 timescales, \( \tau = 1.0 \) | 2MN       | 1         | 0.193       | \( \text{gau}(\beta, c_1) \)               |
|        |          | 1         | 0.195       | \( \text{det}(0.1) \)                      |
|        |          | 2         | 0.197       | \( \text{detrat}(0.02,0.1), \text{rat}(0.5) \) |
|        |          | 3         | 0.200       | \( \text{detrat}(0.003,0.02), \text{rat}(6,7) \) |
|        |          | 4         | 0.205       | \( \text{detrat}(0,0.003), \text{rat}(8,9) \) |
| cA211.30.32, 5 timescales, \( \tau = 1.0 \) | 2MN       | 1         | 0.193       | \( \text{gau}(\beta, c_1) \)               |
|        |          | 1         | 0.195       | \( \text{det}(0.1) \)                      |
|        |          | 2         | 0.197       | \( \text{detrat}(0.02,0.1), \text{rat}(0.5) \) |
|        |          | 3         | 0.200       | \( \text{detrat}(0.003,0.02), \text{rat}(6,7) \) |
|        |          | 4         | 0.205       | \( \text{detrat}(0,0.003), \text{rat}(8,9) \) |
| cA211.12.48, 6 timescales, \( \tau = 1.0 \) | 2MN       | 1         | 0.185       | \( \text{gau}(\beta, c_1) \)               |
|        |          | 1         | 0.190       | \( \text{det}(0.16) \)                     |
|        |          | 2         | 0.195       | \( \text{detrat}(0.03,0.16), \text{rat}(0.2) \) |
|        |          | 3         | 0.200       | \( \text{detrat}(0.006,0.03), \text{rat}(3,4) \) |
|        |          | 4         | 0.205       | \( \text{detrat}(0.001,0.006), \text{rat}(5,6) \) |
|        |          | 5         | 0.210       | \( \text{detrat}(0,0.001), \text{rat}(7,9) \) |

---

### RATIO OF KAON AND PION LEPTONIC DECAY CONSTANTS

PHYS. REV. D 104, 074520 (2021)
The light quark determinant can then be reexpressed as
\[ \det(Q_+ Q_-) = \det(M_{ee}^+ M_{ee}^-) \cdot \det(\hat{Q}_+ \hat{Q}_-). \] (A6)

In order to implement mass preconditioning, the \( \hat{Q}_\pm \) can be shifted by a constant through the addition of a further twisted mass: \( \hat{W}_\pm(\rho) = \hat{Q}_\pm \pm i\rho \). Such that \( \hat{W}_+ \hat{W}_- = \hat{Q}_+ \hat{Q}_- + \rho^2 \). It should be noted that this shift is applied to the even-odd-preconditioned operator, such that the factor \( M_{ee}^- \) remains independent of \( \rho \) since its inverse is nontrivial.

In terms of pseudofermion fields, one thus obtains a contribution to the partition function
\[ \int \mathcal{D}\phi_1 \mathcal{D}\phi_2 \exp \left\{ -\phi_1^\dagger \hat{W}_- \frac{1}{\hat{Q}_-} \hat{W}_+ \phi_2 \right\}. \] (A8)
which we refer to as the determinant ratio.

Determinant ratio \([\text{det}((\rho_1, \rho_2, \ldots, \rho_n)]\) Equation (A8) generalizes to the introduction of multiple shifts \( \rho_1, \rho_2, \ldots, \rho_n \) to contributions of the form:
\[ \int \mathcal{D}\phi_1 \mathcal{D}\phi_i \times \exp \left\{ -\phi_1^\dagger \hat{W}_- (\rho_i) \frac{1}{\hat{W}_+(\rho_i) \hat{W}_-(\rho_i)} \hat{W}_+ (\rho_i) \phi_i \right\}. \] (A9)

The pseudofermion fields \( \phi_i \) are defined only on the odd sites of the lattice and are generated from a random spinor field \( R_i \), sampled from a normalized Gaussian distribution at the beginning of each molecular dynamics trajectory. In the case of the determinant, we have \( \phi_1 = \hat{Q}_+ R_i \), while in the case of the determinant ratio we have \( \phi_1 = (\hat{W}_+(\rho_i))^{-1} \hat{W}_+(\rho_i) R_i \).

The complete mass-preconditioned contribution with \( n \) shifts is thus given by
\[ \text{det}(\rho_n) \cdot \text{det}((\rho_{n-1}, \rho_n) \cdot \text{det}((\rho_{n-2}, \rho_{n-1}) \cdot \ldots \cdot \text{det}(0, \rho_1), \] (A10)
where the last factor has the form of Eq. (A8) with the target twisted quark mass in \( \hat{Q}_+ \). In general, the different contributions are integrated on multiple timescales because their contributions to the force differ by orders of magnitude.

Rational approximation partial fraction \([\text{rat}(n_x, n_y)]\) The Dirac operator for the nondegenerate flavor doublet employed in the strange-charm sector reads
\[ D_h(\bar{\mu}, \bar{c}) = D_{sw} \cdot 1_f + i\bar{\mu} \gamma_5 \tau_j^a - \bar{c} \tau_j^1, \] (A11)
with the property
\[ D_h^\dagger = \tau_j^a D_h \gamma_5 \tau_j^a. \] (A12)
Equivalently, as used (without the clover term) in Ref. [42], one may write
\[ D_h'(\mu_x, \mu_y) = D_{sw} \cdot 1_f + i\mu_x \gamma_5 \tau_j^1 + \mu_y \delta^a, \] (A13)
which is related to \( D_h \) by \( D_h' = (1+i\tau_j^a D_h (1-i\tau_j^a))/2 \) and \( (\mu_x, \mu_y) \to (\bar{\mu}, \bar{c}) \).

As before, we define \( Q_h = \gamma_5 D_h \) and the implementation of even-odd preconditioning translates straightforwardly from the mass-degenerate case, although the construction of \( M_{ee}^- \) has to take into account the additional (off-diagonal) flavor structure.

The operator \( \hat{Q}_h \), defined only on the odd sites, has the property \( \hat{Q}_h \gamma_j^a = \tau_j^a \hat{Q}_h \gamma_j^a \) and the nondegenerate quark doublet contributes a factor
\[ \text{det}(Q_h) \propto \text{det}(\hat{Q}_h), \] (A14)
to the partition function, which we simulate via
\[ [\text{det}(\hat{Q}_h)]^{1/2} \approx \text{det}(R^{-1}). \] (A15)
where we made use of the shorthand notation \( \hat{Q}_h = \hat{Q}_h \gamma_j^a \hat{Q}_h \gamma_j^a \).

We use a rational approximation of order \( N \) (see Refs. [43–45])
\[ R(\hat{Q}_h) = A \prod_{k=1}^N \frac{\hat{Q}_h^2 + a_{2i}}{\hat{Q}_h^2 + a_{2i-1}} \approx \frac{1}{\sqrt{\hat{Q}_h^2}}. \] (A16)
For this, we employ the Zolotarev solution [46] for the optimal approximation to \( 1/\sqrt{\gamma} \), where the coefficients \( a_i \) satisfy the property
\[ a_1 > a_2 > \ldots > a_{2N} > 0. \] (A17)
The amplitude \( A \), the coefficients \( a_i \) and the maximal deviation of the rational approximation \( \delta = \max_y |1 - \sqrt{\gamma} R(y)| \) are computed analytically at given order \( N \) and lower bound \( c < y < 1 \). These are
\[ a_i = \text{cs}^2(i \cdot v, \sqrt{1-e}) \quad \text{with} \quad v = \frac{K(\sqrt{1-e})}{2N+1}, \]

\[ A = \frac{2}{1 + \sqrt{1 - d^2}} \prod_{j=1}^{N} s_{2j-1} \quad \text{with} \quad s_j = \text{sn}^2(i \cdot v, \sqrt{1-e}), \]

\[ \delta = \frac{d^2}{1 + \sqrt{1 - d^2}} \quad \text{with} \quad d = (1-e)^{2(N-1)} \prod_{j=1}^{N} s_{(2j-1)}, \]

where \( \text{sn}(u, k) \) and \( \text{cs}(u, k) = \text{cn}(u, k)/\text{sn}(u, k) \) are Jacobi elliptic functions and \( K(k) \) is the complete elliptic integral. In all our simulations we use \( N = 10 \) and \( e = \lambda_{\text{min}}/\lambda_{\text{max}} \) where \( \lambda_{\text{min}} \) and \( \lambda_{\text{max}} \) are, respectively, the lower and upper bound of the eigenvalues of \( \hat{Q}_h^2 \). In order to have all the eigenvalues \( \lambda \) in the range \( e < \lambda < 1 \), we rescale \( \hat{Q}_h^2 \) with \( \lambda_{\text{max}} \). The values of \( \lambda_{\text{min}} \) and \( \lambda_{\text{max}} \) per ensemble are given in Table V. In the simulation, we explicitly check that the eigenvalues of \( \hat{Q}_h^2 \) remain within these bounds.

The factors in the approximation can be grouped

\[ R(\hat{Q}_h^2) = A \hat{r}_i(\hat{Q}_h^2) \cdot \hat{r}_j(\hat{Q}_h^2) \cdot \ldots, \]

where

\[ r_{n_i}(\hat{Q}_h^2) = \prod_{i=n_f}^{n_k} \hat{Q}_h^2 + a_{2i-1} = \text{rat}(n_f, n_k). \]

We perform partial fraction expansions of the terms

\[ r_{n_i}(\hat{Q}_h^2) = 1 + \sum_{i=n_f}^{n_k} \frac{q_i}{\hat{Q}_h^2 + \mu_i^2}, \]

such that the necessary matrix inverse can be calculated efficiently using a multishift solver. The coefficients \( q_i \) are given by

\[ q_i = (a_{2i-1} - a_{2i}) \prod_{m=n_f, m \neq i}^{n_k} \frac{a_{2m-1} - a_{2i}}{a_{2m} - a_{2i}}, \quad i = n_f, \ldots, n_k. \]

We can further define \( \mu_i = \sqrt{a_{2i}} \) and \( \nu_i = \sqrt{a_{2i-1}} \) and express \( q_i \) as

\[ q_i = (\nu_i^2 - \mu_i^2) \prod_{m=n_f, m \neq i}^{n_k} \frac{\nu_m^2 - \mu_i^2}{\mu_m^2 - \mu_i^2}, \quad i = n_f, \ldots, n_k. \]

At the beginning of each trajectory, pseudofermion fields are generated as follows: again a random spinor field \( R \) is sampled from a Gaussian distribution. Now, we need to compute pseudofermion fields \( \phi \) from

\[ R^2 R = \phi^\dagger R \phi \]

and, therefore, we need operators \( C^\dagger \) and \( C \) with the property

\[ R^{-1} = C^\dagger \cdot C, \quad \Rightarrow \phi = C \cdot R. \]

\( C \) is given by (inspired by twisted mass)

\[ C = \prod_{i=1}^{N} \hat{Q}_h + i\mu_i, \]

which can again be written as a partial fraction

\[ C = 1 + i \sum_{i=1}^{N} \frac{r_i}{\hat{Q}_h + i\nu_i}, \]

with

\[ r_i = (\mu_i - \nu_i) \prod_{m=1, m \neq i}^{N} \frac{\mu_m - \nu_i}{\nu_m - \nu_i}, \quad i = 1, \ldots, N. \]

The rational approximation \( R \) can be applied to a vector using a multimass solver and the partial fraction representation. The same works for \( C \); after solving \( N \) equations simultaneously for \( (\hat{Q}_h^2 + \nu_i^2)^{-1} \), \( i = 1, \ldots, N \), we have to multiply every term with \( (\hat{Q}_h - i\nu_i) \). The Hermitian conjugate of \( C \) is given by

\[ C^\dagger = 1 - i \sum_{i=1}^{N} \frac{r_i}{\hat{Q}_h - i\nu_i}, \]

using \( \hat{Q}_h^\dagger = \hat{Q}_h \). For the acceptance step only the application of \( R \) is needed.

**Rational approximation correction factor** [ratcor\( (n) \)]

The rational approximation \( R \) only has a finite precision. This finite precision can be accounted for during the acceptance step in the hybrid Monte Carlo (HMC) by estimating [45] \( 1 - |\hat{Q}_h^2|R \), if the rational approximation is precise enough. This can be achieved by including a monomial \( \text{det}(\hat{Q}_h^2) \) in the simulation, for which one needs an operator \( B \)

\[ B \cdot B^\dagger = |\hat{Q}_h^2| R. \]

Following Ref. [45], \( B \) can be written as

\[ B = (1 + Z)^{1/4} = \sum_{i=0}^{m} c_i Z^i \]

\[ = 1 + \frac{1}{4} Z - \frac{3}{32} Z^2 + \frac{7}{128} Z^3 + \ldots. \]
with $Z = \hat{Q}_k^2 R^2 - 1$. The series converges rapidly and can, thus, be truncated after a few terms, $m + 1$. The convergence can actually be controlled during the simulation and the truncation does not need to be fixed. We choose to sum the series until the contribution of the given term to the acceptance Hamiltonian is below the residual precision squared, $r^2$, that we employ for the solution of the linear systems involved in the approximation of $R$ in the acceptance step, such that $|c_n z_m \phi|^2 < r^2$, which we typically choose to be at the limit of double precision arithmetic.

For this monomial the pseudofermion field is computed from

$$\phi = B \cdot R,$$

where $R$ is again a Gaussian random vector, see above.

**b. Simulation parameters**

In Table VI we list monomials and parameters used per ensemble. The monomials are grouped in various timescales where the one with the highest id is the outermost timescale (with the fewest integration steps) into which the other time scales are nested. For the various timescales two integrator types are used, either the second order minimal norm (2MN) integrator or its extension with a force gradient (2MNFG), making the latter a fourth-order integrator with multiple timescales, experience suggests that further deviations from this optimal value improve the acceptance rate, such that we often use schemes with increasing values of $\lambda$ from the innermost to the outermost time scale, as shown in Table VI.

The simulations presented in this study have been generated using the tmLQCD [48–50] software suite, which provides all the necessary components to perform $N_f = 2 + 1 + 1$ simulations of twisted mass clover fermions, including implementations of the polynomial and rational HMC algorithms for the nondegenerate determinant. To enable multigrid solvers to be used in simulations [51], tmLQCD provides an interface to DDaAMG [52], a multigrid solver library optimized for twisted mass (clover) fermions [53]. The force calculation of some monomials in the light quark sector is accelerated by a 3-level multigrid approach. Moreover, we extended the DDaAMG method for the mass nondegenerate twisted mass operator. The multigrid solver used in the rational approximation [54] is particularly helpful for the lowest terms of the rational approximation, as well as for the rational approximation corrections in the acceptance step, where it yields a speed up of two over the standard multimass shifted conjugate gradient solver on traditional distributed-memory machines based on Intel Skylake or AMD EPYC architectures.

On the other hand, especially on machines based on Intel’s Knight’s Landing (KNL) architecture, only the most poorly conditioned monomials benefit from the usage of DDaAMG, to the point where (on KNL) the inversion of the nondegenerate operator does not benefit at all. To improve efficiency, tmLQCD also provides an interface to the QPhiX [55] lattice QCD library, which we have refactored and extended [56] to support twisted mass clover fermions, including the nondegenerate doublet. For solves related to the degenerate determinant and determinant ratios, this allows us to efficiently and flexibly combine mixed-precision CG and SIMD vector lengths of 8 or 16 as required by AVX512. On KNL, single-precision QPhiX kernels are up to a factor of 5 more efficient than their tmLQCD-native equivalents. Also in the multimass shifted conjugate gradient solves in the nondegenerate sector, the double-precision kernels in QPhiX are up to a factor of 2 more efficient than the tmLQCD-native equivalents on KNL. Combined, these efficiency improvements lead to overall speedup factors of 2–3 in the HMC on this architecture with smaller overall gains on Skylake and EPYC.
APPENDIX B: EXTRACTION OF $aM_{\pi}$ AND $af_{\pi}$ USING THE ODE PROCEDURE

The spectral decomposition of the pion correlation function (14) can be investigated adopting the ODE procedure of Ref. [22]. This method is able to extract exponential signals from the temporal dependence of a lattice correlator without any a priori knowledge of the multiplicity of each signal and it does not require any starting input for the masses and the amplitudes of the signals.

The ODE approach is sensitive to the noise of the correlator, so that pure oscillating signals (conjugate pairs of imaginary masses) may typically appear in the ODE spectral decomposition. Therefore, we adopt an improved version of the ODE procedure, in which a subsequent $\chi^2$-minimization procedure is applied to the non-noisy part of the ODE spectral decomposition [22]. In this way the accuracy of the physical (i.e., non-noisy) part of the ODE spectral decomposition is improved.

The time intervals $[t_{\text{min}},t_{\text{max}}]$ adopted for the analysis and the extracted values of the pion mass and decay constant in lattice units are collected in Table VII.

Within the ODE procedure we searched for eight exponential signals in the time intervals of Table VII and in all cases at least two physical (non-noisy) exponential signals were found. Then, a $\chi^2$-minimization procedure was applied using the physical ODE solution as the starting point. The minimized values of the $\chi^2$ variable turned out to be always less than 1.

The extracted values as well as their statistical errors of the ground-state mass and decay constant, collected in Table VII, are nicely consistent with the corresponding ones obtained by the direct single exponential fit (15) shown in Table II.

Using the above pion data for $f_{\pi}$ the NLO analysis of Sec. IV A, including the discretization term proportional to $a^2 M_{\pi}^2$, yields for the GF scale $w_0$ the value in agreement with the result (52). Analogously, the use of the data for $X_{\pi}$ and of the NLO fit (56) with $A_2^s = F_{\text{FVE}} = 0$ (see Sec. IV B) leads to

$$w_0 = 0.1740(16) \text{ fm} \quad (B1)$$

in agreement with the corresponding result shown in the second row of Table III.

APPENDIX C: MAXIMAL TWIST CORRECTIONS FOR MASSES AND DECAY CONSTANTS

We follow the general approach of Ref. [30] to the mixed action formulation of twisted mass lattice QCD, which ensures an unitary continuum limit (provided sea and valence renormalized quark masses are matched). Here however we allow for small deviations (due e.g., to numerical errors) from the maximal twist case, i.e., for $m_0 \neq m_{cr}$.

1. $N_f = 2 + 1 + 1$ isosymmetric QCD with twisted clover Wilson quarks

The lattice action can be conveniently written in terms of gauge, sea quark and valence quark plus valence ghost fields. If the sea quarks are arranged in two-flavor fields $\chi_f$ and $\chi_h$ and the valence quarks are described by one-flavor fields $\chi_f$, with $f = u, d, \ldots$, then we have

$$S = S_g[U] + S_{\text{tm}}[\chi_f, \tilde{\chi}_f, U; \mu_f, 0; m_0]$$
$$+ S_{\text{sh}}[\chi_h, \tilde{\chi}_h, U; \mu_h, \mu_0; m_0]$$
$$+ S_{\text{val}}[\{\chi_f, \tilde{\chi}_f\}, U; \{\mu_f\}, m_0], \quad (C1)$$

with the valence sector given by

| Ensemble | $\beta$ | $V/\alpha^4$ | $[t_{\text{min}}/a, t_{\text{max}}/a]$ | $aM_{\pi}$ | $af_{\pi}$ |
|----------|---------|-------------|----------------------------------|-------------|-----------|
| cA211.53.24 | 1.726 | $24^3 \times 48$ | [5, 24] | 0.16621(40) | 0.07106(36) |
| cA211.40.24 | 242 | $24^3 \times 48$ | [5, 24] | 0.14473(76) | 0.06809(30) |
| cA211.30.32 | 322 | $32^3 \times 64$ | [6, 32] | 0.12523(18) | 0.06675(15) |
| cA211.12.48 | 483 | $48^3 \times 96$ | [6, 48] | 0.08000(28) | 0.06139(34) |
| cB211.25.24 | 1.778 | $24^3 \times 48$ | [6, 24] | 0.10750(189) | 0.05351(48) |
| cB211.25.32 | 322 | $32^3 \times 64$ | [6, 32] | 0.10454(43) | 0.05656(37) |
| cB211.25.48 | 483 | $48^3 \times 96$ | [6, 48] | 0.10454(13) | 0.05727(11) |
| cB211.14.64 | 643 | $64^3 \times 128$ | [7, 64] | 0.07845(8) | 0.05476(12) |
| cB211.072.64 | 643 | $64^3 \times 128$ | [7, 64] | 0.05659(8) | 0.05266(15) |
| cC211.06.80 | 1.836 | 803 | [7, 80] | 0.04721(7) | 0.04504(10) |
where ellipses stand for possible replica \((\chi'_j)\) of the valence quarks with \(\mu'_f = -\mu_f\) and the ghost terms exactly cancel the valence fermion contributions to the effective action. Here we find it convenient to express all fermion fields in the canonical quark basis for untwisted Wilson fermions and denote by \(D_{Wclo\nu}\) the well-known clever improved (gauge covariant) Dirac matrix: \(D_{Wclo\nu} = D_{Wclo\nu}[U] = \gamma \cdot \nabla[U] - (a/2)[\nabla^* \cdot \nabla][U] + i(c/4)\sigma \cdot F_{clo\nu}[U].\)

We start by discussing the light valence quark sector, the extension to heavier flavors is straightforward. Following Refs. [27,31], we define (as customary) the twist angle \(\omega_\ell\) in terms of the bare mass parameters of the light valence quark \((u, d)\) doublet \(X_\ell = (\chi_u^0, \chi_d^0)^T\), viz.

\[
\sin \omega_\ell = \frac{\mu_\ell}{\sqrt{Z_A m_{PCAC}^2 + \mu_\ell^2}}, \\
\cos \omega_\ell = \frac{Z_A m_{PCAC}}{\sqrt{Z_A m_{PCAC}^2 + \mu_\ell^2}},
\]

(C3)

with \(Z_A\) the renormalization constant of \(\bar{X}_\ell \gamma_\mu \gamma_5 (\tau^{1,2,3}/2) X_\ell\), which, being independent of quark mass parameters, is defined in the chiral limit \(\mu_f \to 0, m_0 \to m_{cr}\). Maximal twist corresponds to \(|\omega_\ell| = \pi/2\), i.e., to angle \(\theta_\ell = \pi/2 - \omega_\ell\) equal to zero or \(\pi\). We thus have

\[
\cos \theta_\ell = \sin \omega_\ell = \frac{1}{\sqrt{1 + (Z_A m_{PCAC}^2)/\mu_\ell^2}}, \\
\sin \theta_\ell = \cos \omega_\ell = \frac{1}{\sqrt{1 + \mu_\ell^2/(Z_A m_{PCAC}^2)}}.
\]

(C4)

Here \(\mu_\ell\) is the bare twisted mass parameters for the \((u, d)\) doublet and \(m_{PCAC}\) denotes the untwisted bare quark mass of the \((u, d)\) doublet as obtained from the nonsinglet Ward-Takahashi Identity (WTIs)—hence a function of \(m_0\) plus the other bare parameters. We recall that \(m_{PCAC} \propto m_0 - m_{cr}\). The renormalized twisted and untwisted quark mass parameters that appear in the chiral WTIs read (up to discretization effects)

\[
\mu_\ell^R = \mu_\ell \frac{Z_p}{Z_p}, \quad m^R = Z_A m_{PCAC} \frac{1}{Z_p} = (m_0 - m_{cr}) \frac{1}{Z_{S^0}}, \quad (C5)
\]

where \(Z_p\) and \(Z_{S^0}\) are the renormalization constants of the pseudoscalar nonsinglet and the scalar singlet densities (in the canonical basis for untwisted Wilson quarks).

Defining the twist angle and hence formulating the maximal twist condition in terms of \(m_{PCAC}\), as measured on the ensembles with \(2 + 1 + 1\) dynamical flavors, effectively takes care of (compensates for) all the UV cutoff effects related to the breaking of chiral symmetry, including those coming from the \(2 + 1 + 1\) sea quark flavors.

2. Pion mass and decay constant

We argue here that in the case of small enough numerical deviations from maximal twist the lattice charged pion quantities

\[
M_\pi|_{L} = \frac{|2\mu_\ell |\pi^1(0)|P_1^1(0)/(M^2_\pi \cos \theta_\ell)|}{|L|}, \quad (C6)
\]

with \(P_1^1 = \bar{X}_\ell \gamma_5 (\tau^1/2) X_\ell\) and \(X_\ell = (\chi_u^0, \chi_d^0)^T\), approach \(M_\pi\) and \(f_\pi\) as \(a \to 0\) with lattice artifacts having numerically small, and (we shall see) within errors immaterial, differences as compared to the \(O(a^2)\) cutoff effects occurring at maximal twist. Of course these values of \(M_\pi\) and \(f_\pi\) correspond to the light quark renormalized mass \(M_\pi^R = \sqrt{(m^R)^2 + (\mu_\ell^R)^2}\).

The numerical information on \(M_\pi\) and \(f_\pi\) comes from the simple correlator \(C_{PP}(x_0) = a^3 \sum (P_1^1(x)P_1^1(0))\) (and \(C_{PP}(x_0))\). The large-\(x_0\) behavior of \(C_{PP}(x_0)\) determines \(M_\pi\) and an exact lattice WTI relates the operator \(P_1^1\) to the four-divergence of a conserved lattice (backward one-point split) current, which we denote by \(\bar{V}^1_{\chi_\mu}\), viz.

\[
\partial_\mu \bar{V}^1_{\chi_\mu}(x) = 2 \mu_\ell |P_1^1(x) = 2 \mu_\ell^R |P_R^1(x), \quad (C7)
\]

implying that the pion-to-vacuum matrix element of \(\bar{V}^1_{\chi_\mu}\) gives information on \(f_\pi\) (barring the case of \(\cos \theta_\ell = 0\)). In Eq. (C7) \(P_R^1 = Z_p P^1\) and the equalities hold at operator level for finite lattice spacing \((a > 0)\). Hence the lhs of Eq. (C7) is a renormalized operator and information on the approach of its matrix elements to the continuum limit can be obtained by studying the behavior as \(a \to 0\) of the corresponding matrix elements of \(2 \mu_\ell^R |P_R^1\).

Taking the matrix element of Eq. (C7) between the vacuum and a one-\(\pi^1\) state of zero three-momentum and noting that in the continuum limit

\[
\bar{V}^1_{\chi_\mu} \rightarrow (\bar{X}_\ell \gamma_5 (\tau^1/2) X_\ell) R \quad \Rightarrow \quad \sin \theta_\ell \left( \psi \gamma_\mu \frac{e^2}{2} \psi \right)^R + \cos \theta_\ell \left( \psi \gamma_\mu \gamma_5 \frac{e^1}{2} \psi \right)^R, \quad (C8)
\]

where \(\psi = (u, d)^T\) obeys the (continuum) equation of motion (e.o.m.) \((\gamma \cdot D + M^R_\pi)\psi = 0\), for \(a > 0\) one obtains
$2[\mu_r\langle\pi^1(0)|P^1(0)\rangle]_L$

$$= \left[ \cos \theta_r M^2_{\alpha} f_{\pi} + \sin \theta_r \left( \langle\pi^1(0)|\partial_0 \left( \frac{1}{2} P \cdot \psi \right) \rangle \right) \right]_L,$$

$$= \cos \theta_r M^2_{\alpha} f_{\pi} + O(a). \quad (C9)$$

This relation implies that as $a \rightarrow 0$ the ratio

$$\left[2\mu_r\langle\pi^1(0)|P^1(0)\rangle/(M^2_{\alpha} \cos \theta_r)\right]_L \rightarrow f_{\pi}, \quad (C10)$$

generic $\theta_r \neq \pm \pi/2$. Hence at $a > 0$ the ratio (C10) represents a bona fide lattice estimator of $f_{\pi}$, while its discretization errors depend on the lattice artifacts in $M^2_{\alpha}$, $\theta_r$ (or equivalently $m^R$, $\mu^R$) and the renormalized quantity $2\mu_r G_{\pi} = 2\mu_r \langle\pi^1(0)|P^1(0)\rangle$.

We are interested here in situations where $Z_A m_{PCAC}/\mu_r < 1$ but not negligibly small, say slightly above 0.1. This situation indeed occurs in our gauge configuration ensemble cA211.12.48, at $a \sim 0.095$ fm and $a\mu_r = 0.0012$, where we find $Z_A m_{PCAC}/\mu_r \sim 0.15$. In this case an analysis à la Symanzik of $M^2_{\alpha}$, $m^R$ and $2\mu_r G_{\pi}$ shows (see below) that the change in the lattice artifacts of our lattice estimator of $f_{\pi}$, with respect to those purely $O(a^{2n})$ (with $n$ integer) that occur at maximal twist, is smaller than 0.001$f_{\pi}$, therefore numerically immaterial within statistical errors that are typically of order 0.005$f_{\pi}$.

**a. The change in the lattice artifacts for $M^2_{\alpha}$ and $f_{\pi}$**

If $|a m_{PCAC}|$ is nonzero, though definitely smaller than $|a\mu_r|$, then the same holds for $|a(m_0 - m_{cr})|$ and one expects that the appropriate lattice estimators of $M^2_{\alpha}$, $f_{\pi}$ and any other physical quantity will be altered already at $O(a)$ as compared to their counterparts at maximal twist. Correcting analytically the lattice estimators for the deviation from maximal twist at order $a^0$ is hence not enough and one must also check that the out-of-maximal-twist modifications in order $a$ and order $a^2$ lattice artifacts are numerically negligible within statistical errors. Otherwise, analyzing data corrected for deviations from maximal twist on some gauge ensembles together with data evaluated at maximal twist on other gauge ensembles might lead to a systematic bias in the continuum extrapolation, where one typically assumes uniform $O(a^2)$ artifacts—as expected if all data are obtained at maximal twist.

**b. Structure of the Symanzik effective Lagrangian**

Let us analyze à la Symanzik the $N_f = 2 + 1 + 1$ lattice QCD theory (C1) out-of-maximal-twist and focus here on the light valence sector. We assume the reader is familiar with the basic literature on this topic such as [27,57,58] and references therein. The Symanzik local effective Lagrangian to be used in our analysis of $O(a)$ artifacts then reads

$$L_{\text{Syma}} = L_4 + a L_5 + a^2 L_6 + O(a^3),$$

$$L_4 = \frac{1}{4}(F \cdot F) + \bar{\psi}_{\ell}(\gamma \cdot D + m^R + i\gamma^3 \mu_{\ell}^R)\psi_{\ell}$$

$$+ \bar{\psi}_h(\gamma \cdot D + m^R + i\gamma^3 \mu_{h}^R + \epsilon_h)\psi_{h}$$

$$+ \bar{\psi}_{\ell}^{\omega}(\gamma \cdot D + m^R + i\gamma^3 \mu_{\ell}^R)\psi_{\ell}^{\omega} + \ldots \quad (C11)$$

where $X_{\ell}^{\omega} = (\chi_{\omega}, \chi_{\ell})^T$ describes the valence light quarks in the same basis as in Eq. (C2) while ellipses stands for $d \leq 4$ terms involving heavier valence quarks and ghost terms. Upon taking the continuum limit in isosymmetric QCD with $2 + 1 + 1$ dynamical flavors, we must have coinciding sea and valence renormalized masses for each flavor and keep constant as $a \rightarrow 0$ the renormalized parameters $g_{\mu}^2$, $\mu_{\ell}^R$, $\mu_{h}^R$, $\epsilon_{h}$.

The local effective Lagrangian terms $L_n$ are suitable linear combinations of the $d = n > 4$ operator terms allowed by the symmetries of the lattice theory (C1). In particular it turns out that

$$L_5 = \ldots + (c - c_{\text{SW}}) \frac{1}{4} X_{\ell}^{\omega} \cdot \xi F X_{\ell}^{\omega} - b g m^2 (F \cdot F - (b_m m^2 + b_m^2) X_{\ell}^{\omega} X_{\ell}^{\omega}$$

$$- b_b m \bar{\psi}_{\ell}^{\omega} i\gamma^3 \psi_{\ell}^{\omega} + \ldots \quad (C12)$$

where $m = Z_A m_{PCAC} \propto m_{0} - m_{cr}$, while ellipses stand here for terms involving only heavier valence quark operators as well as sea quark and ghost terms (all of them are omitted since they are immaterial for this section). With $c - c_{\text{SW}}$ we indicate the difference between the 1-loop tadpole improved estimate (c) employed in our simulation and the exact value ($c_{\text{SW}}$) of the coefficient of the clover term.3

Concerning $L_6$, it will be enough to focus on its $m$-dependent sector and to note the structure

$$L_6(m; \mu_{\ell}, \mu_{h}, \epsilon_{h}, \mu_{s}, \ldots) = L_6(0; \mu_{\ell}, \mu_{h}, \epsilon_{h}, \mu_{s}, \ldots) + mO_5 + m^2O_4. \quad (C13)$$

where $O_5$ ($O_4$) is a linear combination of the $m$-independent terms allowed in $L_5$ ($L_4$). Among the latter, since we employ in our correlators flavor diagonal OS valence quark fields $\chi_{\ell}^{\omega}$ with twisted mass $\mu_{\ell} > 0$, or $X_{\ell}^{\omega}$ fields with twisted mass $\mu_{\ell} = -\mu_{\ell}$, for the purposes of this section only the terms bilinear in the $X_{\ell}^{\omega}$ and $\bar{\psi}_{\ell}^{\omega}$, or $X_{\ell}^{\omega}$ and

3From our experience in simulations with two dynamical flavors in a lattice setup where $c_{\text{SW}}$ is known, we expect that $|c - c_{\text{SW}}| < 0.15$, which suppresses $O(a)$ lattice artifact by nearly one order of magnitude and, provided $|a m_{\text{PCAC}}| < 1$ is small enough, makes undesired $O(a)$ numerically negligible in most observables.
\(\tilde{\mathcal{X}}_{\sigma}^{\text{val}}\), fields are relevant, which we may call \(mO_{\chi\rho}^{\text{val}}\) and \(m^2O_{\chi}^{\text{val}}\).

We note that exact or spurious lattice symmetries rule out\(^4\) the \(L_5\) terms of the form

\[
\mu_f i \tilde{F}, \quad \mu_f \tilde{X}_f \mathbb{t}^{0, 1, 2, 3} 3 \gamma \cdot DX_f, \quad \mu_f \tilde{X}_f \mathbb{t}^{0, 1, 2, 3} 3 \gamma \cdot DX_f,
\]
as well as the analogous \(L_6\) terms of the form

\[
m_f i \tilde{F}, \quad m_f \tilde{X}_f \mathbb{t}^{0, 1, 2, 3} 3 \gamma \cdot DX_f, \quad m_f \tilde{X}_f \mathbb{t}^{0, 1, 2, 3} 3 \gamma \cdot DX_f.
\]

\(\text{c. The discretization effects on } M_\sigma | L\)

In the case of \(|am| \sim 0.0002 < |\mu_\sigma|\), for the quantity \(M_2^\sigma | L\), the \(O(a)\) deviation from its continuum limit value \(M_2^\sigma = 2 B^R \sqrt{(m^R)^2 + (\mu_\sigma^R)^2 + O((m^R)^2 + (\mu_\sigma^R)^2)} \sim 2 B^R \mu_\sigma^R\) is given by

\[
\delta_1 M_2^\sigma = \delta_{1A} M_2^\sigma + \delta_{1B} M_2^\sigma,
\]

where, writing operators in terms of the physical basis fermion doublet fields

\[
\psi_\sigma = e^{i \alpha 4 \gamma_5 \tau^3} X_\sigma, \quad \tilde{\psi}_\sigma = \tilde{X}_\sigma e^{i \alpha 4 \gamma_5 \tau^3/2}
\]

and exploiting parity and isospin symmetries of continuum \(N_f = 2 + 1 + 1 \text{ QCD}\), one has

\[
\delta_{1A} M_2^\sigma = \frac{a \sin \theta_\sigma (\pi^{1, 2}(0)) \left| (c - c_{\text{SW}}) \tilde{\psi}_\sigma \right| \psi_\sigma \frac{i}{4} \sigma \cdot F \tilde{\psi}_\sigma 
\]

\[
- \left( b_m m^2 + \tilde{b}_m \mu_\sigma^R \right) \tilde{\psi}_\sigma \psi_\sigma \left| \pi^{1, 2}(0) \right|
\]

\[
\lesssim 0.003 \frac{\alpha_s(\Lambda_{\text{QCD}})^2}{4 \pi} \Lambda_{\text{QCD}}^2 \sim 0.001 M_2^\sigma \]

and (approximating \(\cos \theta_\sigma\) with 1, since \(\sin \theta_\sigma \approx 0.2\))

\[
\delta_{1B} M_2^\sigma = a m \left( \pi^{1, 2}(0) \right) \left( b_{\mu} \tilde{\psi}_\sigma \psi_\sigma + b_\sigma \left| F \cdot F \right| \pi^{1, 2}(0) \right)
\]

\[
\lesssim 2 B^R \mu_\sigma^R 0.0005.
\]

The numerical estimate in Eq. (15) results from \(|c - c_{\text{SW}}| \lesssim 0.15, \quad \alpha_s(\Lambda_{\text{QCD}}) \approx 1, \quad a \Lambda_{\text{QCD}} \approx 0.1\) and \(\Lambda_{\text{QCD}}/\Lambda_\sigma \approx 2\), while \(b_m = -1/2 + O(g_0^2)\) and (making the choice advocated in Ref. [58]) \(\tilde{b}_m = -1/2\).

The numerical estimate in Eq. (16) follows from

\[
\left| \pi^{1, 2}(0) \right| \left| \tilde{\psi}_\sigma \psi_\sigma \left| \pi^{1, 2}(0) \right| \right| \sim 2 B^R \mu_\sigma^R \] and \(|b_{\mu}| = O(g_0^2) < 1\). In fact soft pion theorems (i.e.,

\(\text{d. The discretization effects on } f_{\pi} | L\)

Based on Eq. (C10) the lattice artifacts on \(f_\pi\) can be estimated in terms of the cutoff effects in \(\langle \mu_\pi / \cos \theta_\pi | \pi^{1}(0) | P(0) | L \rangle\) and in \(M_2^\pi | L\). We discussed above the lattice artifacts of \(M_2^\pi | L\). Concerning \(\langle \mu_\pi / \cos \theta_\pi | \pi^{1}(0) | P(0) | L \rangle\), we can see it as the product of the renormalized quantities \(Z_{\text{p}}^\pi \langle \mu_\pi / \cos \theta_\pi | M_2^\pi \rangle\) and \(Z_{\text{p}} \langle \pi^{1}(0) | P(0) | L \rangle\rangle_{\text{p}} = G_{\text{p}}^\pi\), and then discuss separately the lattice artifacts in each of these two factors.\(^5\)

As for \(Z_{\text{p}}^\pi \langle \mu_\pi / \cos \theta_\pi \rangle = \sqrt{(m^R)^2 + (\mu_\sigma^R)^2}\), the form of the \(m\)-dependent terms in \(L_\sigma\) i.e., those with coefficients \(b_m\) and \(b_\sigma\) (which are all in modulus \(\lesssim 1\)), implies that the \(O(a)\) corrections to \(m^R\) and \(\mu_\sigma^R\), and hence to \(M_2^\sigma\) are only of relative size \(|am| \approx 0.0002\). Even smaller are the corrections to \(M_2^\sigma\) of order \(a^2 m\).

Let us then consider the out-of-maximal-twist induced cutoff artifacts in the matrix element \(Z_{\text{p}} \langle \pi^{1}(0) | P(0) | L \rangle_{\text{p}} = G_{\text{p}}^\pi\). They clearly arise from the lattice two-point correlator \(C_4^\pi(x_0)\). At order \(a\), since for the operator \(P^1 = \tilde{\psi}_\chi \chi \pi^{1}(0) / 2\psi_\sigma\) it is known that \(c_P = 0, \quad \tilde{b}_P = 0\) and \(b_P = 1 + O(g_0^2)\), implying \(|P_{\text{p}} | am| \approx 0.0002\), the numerically dominant lattice artifacts stem from the term \(a \int d^4 y \langle \pi^{1}(0) | P(0) | L \rangle_{\text{p}}(y) \rangle_{\text{p}} \langle \text{X}_{\pi}(x) \rangle_{\text{p}}(x) \rangle_{\text{p}}\). Inserting intermediate states and considering the possible \(\gamma_0\) orderings one checks that, owing to the structure of \(L_\sigma\) and \(|am| \approx 0.0002\), the numerically leading cutoff effects in \(G_{\pi}\) are proportional to

\[
(c - c_{\text{SW}}) \sin \theta_\pi a \left[ \langle \pi^{1} | \tilde{\psi}_\sigma \left| \psi_\sigma \right| \pi^{1} \rangle \right] G_{\pi}
\]

\[
+ \sum_n |n| \langle \psi_\sigma^{\text{val}} \left| \psi_\sigma \right| n | P^{1} | n \rangle \]

and hence of relative order \(\sin \theta_\pi a (c - c_{\text{SW}}) |4 \pi|\), which, if \(|am| \approx 0.0002\) and \(|c - c_{\text{SW}}| < 0.15\), noting that \(a \Lambda_{\text{QCD}} a / (4 \pi) < 0.1\), turns out to be

\(\text{\footnote{To this goal it is enough to consider charge combination,}
\(P \times (\mu_\rho e_h \times -\mu_\rho e_h)\) and \((X_h \times i \tau^3 \tilde{X}_h) \times (X_h \times i \tau^3 \tilde{X}_h) \times (\mu_\rho e_h \times -\mu_\rho e_h)\) invariances, with \(P\) meaning parity transformation on gauge fields combined with \(X_h(\tilde{X}) \to X_h(\tilde{X}) \times (X_h(\tilde{X}) \to X_h(\tilde{X})) \times \left( X_h(\tilde{X}) \to X_h(\tilde{X}) \right)\), and \(X_0 = (x_0, -\vec{x})\).}

\(\text{\footnote{Of course we do not worry about possible cutoff effects on}
\(Z_{\text{p}}\), which cancel in the product.}\)
$\lesssim 0.0002$. Compared to this the $O(a^2 m)$ artifacts in $G_{a1}$ are expected to be smaller by at least one order of magnitude.

Hence our lattice estimator of $f_\pi$ is estimated to be affected by the small deviation from maximal twist observed on the ensemble ca211.A12.48, where $|am| \sim 0.0002$, only at a level of $\lesssim 0.0004 f_\pi$, which is negligible as compared to current statistical errors.

e. Analysis of $M_\pi^2$ and $f_\pi$ in terms of the renormalized light quark mass

The discussion above is valid also in case the observables $M_\pi^2$ and $f_\pi$ are analyzed as functions of the renormalized light quark mass

$$M_\pi^R = \sqrt{(m^R)^2 + (\mu^R)^2} = Z_\pi^{-1}(\mu/\cos \theta_\pi).$$

(C17)

As already noted in Sec. C 2d, for the ensemble ca211.12.48 the observed small deviation from maximal twist leads to an undesired $O(am)$ relative change in $M_\pi^R$ that is only of order $|am| \approx 0.0002$ and thus fully negligible in comparison to other errors.

A rather obvious but practically important and general caveat follows in the case one insists in analyzing the lattice terms linear in $am$ is not much difference between a PS meson made out of two light valence quarks (pion, with $|\mu_s| = |\mu_d|$) and a PS meson made out of a light (mass $\mu_\pi$) and a heavy (mass $\mu_\pi$) valence quark. This is a consequence of the fact that the extraction of the PS meson mass and decay constant relies on general positivity properties of two-point correlation functions (close enough to the continuum limit) and on exact chiral WTI, which hold valid irrespectively of the value of valence quark masses. One might thus deal in a fully analogous way with PS mesons made out of two nonlight valence quarks. For definiteness, however, we shall here focus on the case of the kaon, i.e., $\mu_s = \mu_s \gg \mu_\pi$ and $\mu_d = \mu_\pi > 0$.

It turns out that in the case of small enough numerical deviations from maximal twist, say $0.1 a \mu_s < |am| \ll a \mu_\pi < a \mu_s$, the lattice charged kaon quantities

$$M_K |_{L.} \quad [((\mu_\pi + \mu_\pi^0)(K(0))^{P_{s,d}}/0)/(M_K^R \cos((\theta_\pi + \theta_\pi^0)/2))],$$

(C19)

with $P_{s,d} = \mathcal{X}_{s,d} T (x^i/2) \mathcal{X}_{s,d}$ and $\mathcal{X}_{s,d} = (\mathcal{X}_s, \mathcal{X}_d)^T$, approach $M_K$ and $f_K$ as $a \rightarrow 0$ with lattice artifacts having numerically small, and within errors immaterial, differences as compared to the $O(a^2)$ cutoff effects occurring at maximal twist. These values of $M_K$ and $f_K$ correspond to the (light) quark renormalized mass $M_\pi^R = \sqrt{(m^R)^2 + (\mu^R)^2}$ and to the strange ($s$) quark renormalized mass $M_s^R = \sqrt{(m^R)^2 + (\mu_s^R)^2}$.

In full analogy to the definitions adopted for light valence quarks (see Sec. C1) we define $\mu_s^R = \mu_s / Z_p$, where $Z_p$ is (can be conveniently taken as) the same mass-independent renormalization constant of the PS non-singlet quark bilinear operator as above, and

$$\cos \theta_s = \sin \omega_s = \frac{1}{\sqrt{1 + (Z_s^2 m^2_{\text{PCAC}})/\mu_s^R}},$$

$$\sin \theta_s = \cos \omega_s = \frac{1}{\sqrt{1 + \mu_s^2 / (Z_s^2 m^2_{\text{PCAC}})}}.$$  

(C20)

a. On the lattice estimators (C19) of $M_K$ and $f_K$ at $O(a^0)$

The numerical information on $M_K$ and $f_K$ comes in fact from the simple correlator

$$C_{K,K}^{x_0}(x_0) = a^3 \sum_{x_0} \langle P_{s,d} (x) P_{d,s} (0) \rangle.$$  

The large-$x_0$ behavior of $C_{K,K}^{x_0}(x_0)$ determines $M_K$ as the kaon mass that, owing to renormalizability (and unitarity in the continuum limit) of the lattice theory of Sec. C 1, corresponds to the light and strange quark masses $M_\pi^R$ and $M_s^R$. An exact lattice WTI relates the operator $P_{s,d}^{\pi_\pi}$ to the four-divergence of a conserved lattice (backward one-point split) current, which we denote by $\tilde{\mathcal{V}}_{s,d}^{\pi_\pi}$, viz.

$$\partial_\mu \tilde{\mathcal{V}}_{s,d}^{\pi_\pi}(x) = (\mu_\pi + \mu_\pi)(P_{s,d}^{\pi_\pi}(x) = (\mu_\pi^R + \mu_\pi^R)P_{s,d}^{\pi_\pi}(x),$$

(C21)

implying that the kaon-to-vacuum matrix element of $\tilde{\mathcal{V}}_{s,d}^{\pi_\pi}$ gives information on $f_K$, barring the case of $\cos((\theta_\pi + \theta_\pi)/2) = 0$. In Eq. (C21) $P_{s,d}^{\pi_\pi} = Z_p P_{s,d}^{\pi_\pi}$ and
the equalities hold at operator level for \( a > 0 \). Hence the lhs of Eq. (C21) is a renormalized operator and information on the approach of its matrix elements to the continuum limit can be obtained by studying the behavior as \( a \to 0 \) of the corresponding matrix elements of \( (\mu_\epsilon^R + \mu_\epsilon^R) P_{\epsilon,s}^{\text{sd}} \).

Taking the matrix element of Eq. (C21) between the vacuum and a one-\( K \) state of zero three-momentum and noting that in the continuum limit

\[
\sections{a}{14} \epsilon, \theta \to 0 \quad \left( \bar{X}_{\epsilon,s} \gamma_{\mu} (r^2/2) X_{\epsilon,s} \right)^R
= \sin \left( \frac{\theta_\epsilon + \theta_s}{2} \psi_{\epsilon,d,\mu} \frac{r^2}{2} \psi_{s,d} \right)^R
+ \cos \left( \frac{\theta_\epsilon + \theta_s}{2} \right) \psi_{\epsilon,d,\mu} \psi_{s,d}^R,
\]

(C22)

where \( \psi_{\epsilon,d} = (s, d)^T \) obeys the (continuum) e.o.m.

\[
(\gamma \cdot D + \frac{M^2 + M_\pi^2}{2} + \frac{M^2 - M_\pi^2}{2} r^2) \psi_{\epsilon,d} = 0, \quad \text{for} \quad a > 0
\]

one obtains

\[
\left[ (\mu_\epsilon + \mu_s) \left( K(0) \right) |P_{\epsilon,s,d}^\text{sd}\right]_L = \left[ \cos((\theta_\epsilon + \theta_s)/2) M_K^2 f_K \right]_L + O(a).
\]

(C23)

This relation implies that as \( a \to 0 \) the ratio

\[
\left[ (\mu_\epsilon + \mu_s) \left( K(0) \right) |P_{\epsilon,s,d}^\text{sd}\right]_L / \left( M_K^2 \cos((\theta_\epsilon + \theta_s)/2) \right)_L \to f_K,
\]

(C24)

for generic \( \theta_\epsilon + \theta_s \neq \pm \pi \). Hence at \( a > 0 \) the ratio (C24) represents a bona fide lattice estimator of \( f_K \), while its discretization errors depend on the lattice artifacts in \( M_K^2, \theta_\epsilon, \theta_s \) (or equivalently \( m_K, \mu_s^R, \mu_\epsilon^R \)) and the renormalized quantity \( (\mu_\epsilon + \mu_s) G_K = (\mu_\epsilon + \mu_s) \left( K(0) \right) |P_{\epsilon,s,d}^\text{sd}\right] \).

b. Theoretical control and numerical size of \( O(a) \) and \( O(a^2) \) artifacts

Concerning the impact of a nonzero \( am \) value on the lattice artifact in the kaon sector it is important to note that, as phenomenology dictates \( \mu_\epsilon \approx 0.037 \mu_s \), we have in general

\[
\theta_s \ll \theta_\epsilon, \quad \cos((\theta_\epsilon + \theta_s)/2) \approx \cos(\theta_\epsilon/2) [1 - O(\theta_s^2)] - 0(\theta_s \theta_\epsilon),
\]

(C25)

In particular, for the gauge ensemble cA211.12.48, where \( \am = Z_A \mp N \mu_s / \mu_\epsilon \approx -0.15 \), we find

\[
\theta_s \approx -0.15, \quad \theta_\epsilon \approx -0.006, \quad \cos((\theta_\epsilon + \theta_s)/2) \approx \cos(\theta_\epsilon/2) - O(0.0009).
\]

(C26)

For the discussion of \( O(a) \) lattice artifacts in \( M_K^2 \) and \( f_K \) one should consider of course the occurrence of flavor diagonal terms involving the valence quark fields \( \chi_s, \chi_s \) both in \( L_5 \) for Eq. (C12), where they take the form

\[
L_5 \supset (c - c_{SW}) \frac{C_{27}}{4} \bar{X}_s \gamma_\tau \mu \gamma_\tau \mu \sigma \cdot F_{\epsilon,s}^\text{val} - (b m^2 + \bar{b} m^2) \bar{X}_s \gamma_\tau \mu \sigma \cdot \chi_s^\text{val}
- b m_\pi^2 \bar{X}_s \gamma_\tau \mu \psi_{\epsilon,s}^\text{val} \chi_s^\text{val},
\]

(C27)

and in \( L_6 \), for which the structure in Eq. (C13) remains valid. Looking back to Eq. (C12) for the light valence quark and gluonic terms in \( L_5 \), one finds that the discussion for the kaon case closely follows the one for the pion (see Sec. C2a), provided one replaces the valence quark field pair \( X_\epsilon^\text{val} = (\chi_\mu, \chi_d)^T \), used for the latter, with the valence quark field pair \( X_\epsilon^\text{val} = (\chi_s, \chi_d)^T \) relevant for the kaon, as well as \( \theta_\epsilon \) with \( (\theta_\epsilon + \theta_s)/2 \). This implies that for \( M_K^2 \) and \( f_K \) the numerically dominant changes in the lattice artifacts as compared to the maximal twist case, which for \( M_K^2 \) and \( f_K \) were proportional to \( \sin \theta_\epsilon \), turn out to be proportional to

\[
\sin((\theta_\epsilon + \theta_s)/2) \approx \sin(\theta_\epsilon/2) - O(0.006)
\]

(C28)

thereby getting reduced by a factor of about two with respect to the pion case.

In conclusion, if \( |am| \sim 0.0002 \) (as it happens for our ensemble cA211.12.48), we estimate a lattice artifact modification, with respect to the case of maximal twist, that does not exceed 0.0005 \( M_K^2 \) for \( M_K^2 \) and 0.0002 \( f_K \) for \( f_K \) and is hence safely negligible as compared to our current statistical errors.

From the arguments above it should also be clear that the same quantitative estimates hold also for the lattice artifact changes induced by \( am \neq 0 \) in the mass and the decay constant of heavy-light PS mesons with a charm or even heavier non-light valence quark having a mass \( \mu_s \gg \mu_s \). In fact the heavier the valence quark, the smaller \( |m_\mu| \approx |m/s| \) and the more numerically irrelevant the effect of a small nonzero value of \( am \).

APPENDIX D: DETERMINATION OF THE GF SCALES \( \sqrt{t_0} \) AND \( t_0/w_0 \)

In this appendix we describe the calculations of the relative GF scales \( w_0/a, \sqrt{t_0}/a \) and \( t_0/(w_0 a) \) at the physical pion point and we summarize the determination of the absolute scales \( \sqrt{t_0} \) and \( t_0/w_0 \) using the SU(2) ChPT analysis of the data for \( X_\epsilon \) carried out in Sec. IV B in the case of the GF scale \( w_0 \).
1. Determination of the relative GF scales

In this section we provide the details of the determinations of the GF scales at the physical point. Our analysis is based on the values of the gradient flow scales in Table VIII as obtained on the ensembles in Table I. We calculate the scales following the definitions in [12,13,59] using the standard Wilson action for the gradient flow evolution and the symmetrized discretization of the action density. The errors are calculated by taking into account the statistical errors. The details of the extrapolations are summarized in Table IX. In order to illustrate the extrapolations and compare them at different lattice spacings, in Fig. 10 we show the scales normalized by their values at the physical point as a function of Δ2.

In order to use the scales in the analysis of the light meson sector, we need the values at the physical pion-mass point. To achieve this, we perform an extrapolation for the two sets of ensembles cA211 and cB211 to the physical point in terms of Δ2 ≡ (Mπ/fπ)2 − (Mπ/fπ)2phys, such that the physical point is reached when this quantity is zero, while for the ensemble cC211.06.80 we directly use the value at the physical point as given in Table VIII. We note that for this ensemble Δ2 = 0.025 such that the potential corrections would be tiny and in fact smaller than the statistical errors. The details of the extrapolations are summarized in Table IX. In order to illustrate the extrapolations and compare them at different lattice spacings, in Fig. 10 we show the scales normalized by their values at the physical point as a function of Δ2.

From the plots and the data in the table it is obvious that the quark-mass dependence of the scale (t0/w0)/a is very small, i.e., the corrections are less than 0.5% at our largest pion mass ensemble cA211.53.24, to be compared to 2.5% for √t0/a and 4.5% for w0/a. We note that the quark-mass dependence exhibits clearly visible lattice artifacts, but the dependence seems to become weaker towards the continuum limit. One peculiar feature is the fact that for the scale (t0/w0)/a the slope of the quark-mass dependence changes sign when going from the coarser to the finer lattice spacing. This certainly warrants further investigation, once more data is available, however, one should keep in mind that for this quantity the slope is consistent with zero within less than 3σ.

In order to examine the lattice artifacts of the GF scales further, we now turn to the dimensionless ratio s0/w0. In Fig. 11 we show the ratio at the physical point as a function of

| Ensemble     | Nmeas | s0/a | w0/a | (t0/w0)/a | s0/w0 | t0/τint | w0/τint | s0/w0 | w0/τint |
|--------------|-------|------|------|-----------|-------|---------|---------|-------|---------|
| cA211.53.24  | 1122  | 1.5306(21) | 1.7597(43) | 1.33139(89) | 0.86982(100) | 23(6) | 25(7) | 7(1) | 18(4) |
| cA211.40.24  | 1219  | 1.5384(18) | 1.7766(33) | 1.33213(96) | 0.86592(64) | 20(5) | 18(4) | 7(1) | 9(2) |
| cA211.30.32  | 2559  | 1.5460(9)  | 1.7928(17) | 1.33314(47) | 0.86233(32) | 22(5) | 21(4) | 9(1) | 10(2) |
| cA211.12.48  | 326   | 1.5614(22) | 1.8249(33) | 1.33590(155) | 0.85559(29) | 69(30) | 63(27) | 57(25) | 16(5) |
| cB211.25.24  | 1145  | 1.7937(22) | 2.0992(46) | 1.53260(108) | 0.85455(77) | 21(5) | 25(6) | 5(1) | 12(2) |
| cB211.25.32  | 990   | 1.7922(19) | 2.0991(47) | 1.53018(72) | 0.85380(91) | 35(10) | 45(14) | 6(1) | 28(7) |
| cB211.25.48  | 1175  | 1.7915(8)  | 2.0972(19) | 1.52966(41) | 0.85384(38) | 28(8) | 31(9) | 9(2) | 20(5) |
| cB211.14.64  | 619   | 1.7992(5)  | 2.1175(11) | 1.52875(23) | 0.84968(23) | 30(8) | 32(9) | 8(1) | 23(6) |
| cB211.072.64 | 191   | 1.8028(8)  | 2.1272(19) | 1.52784(42) | 0.84750(41) | 45(18) | 52(22) | 16(5) | 41(16) |
| cC211.06.80  | 785   | 2.1094(8)  | 2.5045(17) | 1.77670(37) | 0.84226(27) | 46(17) | 42(16) | 14(3) | 26(8) |
the lattice spacing expressed in units of the three GF scales $s_0^\text{phys}$, $w_0^\text{phys}$ and $t_0^\text{phys}/w_0^\text{phys}$ at the physical point. Note that for the lattice spacing cC211 this corresponds to the value obtained on the ensemble cC211.06.680. The data show a precise $O(a^2)$ scaling towards the continuum and allow continuum extrapolations in terms of $a^2$. The continuum extrapolations using in turn $a^2/(t_0/w_0)^2$, $a^2/t_0$ and $a^2/w_0^2$ yield $(s_0/w_0)^{\text{phys}} = 0.8285(13)$, $0.8291(13)$ and $0.8298(12)$, respectively, with $\chi^2/\text{d.o.f.} = 0.20$, 0.12 and 0.06. The values in the continuum are perfectly consistent with each other and averaging them in the usual way gives $(s_0/w_0)^{\text{phys}} = 0.8291(13)(5)14$, where the second error reflects the spread of the results while the error in the square bracket is the combined one. These results provide a non-trivial check of the expected scaling behavior with quantities determined with an accuracy of between 1–2 permille for the

| Table IX. Results for the extrapolations of the GF scales to the physical point in terms of $\Delta^2 = (M_\pi/f_\pi)^2 - (M_\pi/f_\pi)^{\text{phys}}$, i.e., $X(M_\pi)/a = (X/a)^{\text{phys}} + c \cdot \Delta^2$. Note that for the ensembles cA211 we have $N_{\text{d.o.f.}} = 2$, while for cB211 $N_{\text{d.o.f.}} = 1$. |
|-----------------------------------------------|
| $(s_0/a)_{\text{phys}}$ | $c$ | $\chi^2/\text{d.o.f.}$ | $(w_0/a)_{\text{phys}}$ | $c$ | $\chi^2/\text{d.o.f.}$ |
| cA211 | 1.5660(22) | −0.0082(8) | 0.02 | 1.8352(35) | −0.0174(13) | 0.00 |
| cB211 | 1.80396(68) | −0.0053(5) | 2.14 | 2.1299(16) | −0.0136(12) | 1.84 |

| $(t_0/w_0/a)_{\text{phys}}$ | $c$ | $\chi^2/\text{d.o.f.}$ | $(s_0/w_0)_{\text{phys}}$ | $c$ | $\chi^2/\text{d.o.f.}$ |
| cA211 | 1.3359(12) | −0.0011(4) | 0.15 | 0.8531(10) | 0.0038(4) | 0.05 |
| cB211 | 1.52789(33) | 0.0008(3) | 0.18 | 0.84697(37) | 0.0030(3) | 1.26 |

FIG. 10. Light quark-mass dependence and extrapolations to the physical point of the GF scales $s_0 \equiv \sqrt{t_0}$, $w_0$, $t_0/w_0$, normalized by their values at the physical point for ease of comparing the results at different lattice spacings.
scales and subpermille for the ratio, and they nicely confirm the automatic $O(a)$ improvement in place for TM Wilson fermions at maximal twist.

Nevertheless, we may attempt to fit our data with the lattice spacing, it is not clear how solid this conclusion is. Information on the quark-mass dependence at the finest lattice spacing far away from the physical point, and hence no lattice are not included in the lattice artifacts as described by the global fit.

The global fit suggests that

$$s_0/w_0 = s_0^\text{phys}/w_0^\text{phys} + a_1 \cdot \left( s_0^\text{cont}/w_0^\text{cont} \right)^2,$$

which includes a light quark-mass dependence proportional to $\Delta^2$ described by $B_0$, and lattice artifacts proportional to $a^2/w_0^2$ described by $A_1$ and $B_1$. The latter coefficient describes the lattice artifacts on the quark-mass dependence. The global fit suggests that $B_0$, describing the quark-mass dependence in the continuum, is well consistent with zero, i.e., $B_0 = 0.001(7)$. That is, in the continuum the ratio $\sqrt{t_0}/w_0$ appears to have no dependence on the pion mass at all and the observed quark-mass dependence at finite lattice spacings is apparently just a lattice artifact. However, given the fact that we do not have data for the ratio at the lattice spacing $aC211$ away from the physical point, and hence no information on the quark-mass dependence at the finest lattice spacing, it is not clear how solid this conclusion is. Nevertheless, we may attempt to fit our data with $B_0 = 0$ fixed, and in Fig. 12 we show the results for this global fit. The colored lines with error bands show the light quark-mass dependence of the ratio and the extrapolations to the physical point for each lattice spacing, while the black line with the error band at the bottom shows the fit result in the continuum. The data points on this line represent our data corrected by the lattice artifacts as described by the global fit.

For the ratio at the physical point and in the continuum the fit yields

$$s_0/w_0 = s_0^\text{phys}/w_0^\text{phys} + a_1 \cdot \left( s_0^\text{cont}/w_0^\text{cont} \right)^2,$$

with $\chi^2/\text{d.o.f.} = 0.42$, $N_{\text{d.o.f.}} = 5$ and $A_1 = 0.0806(30)$, $B_1 = 0.0129(5)$.

We note that the ratio is determined with a precision in the subpermille region, i.e., better than 0.8 permille. As such, it provides an interesting consistency crosscheck on any other, independent determination of the scales, e.g., through hadronic quantities.

### 2. Determination of the GF scales $\sqrt{t_0}$ and $t_0/w_0$

The SU(2) ChPT analysis of the data for $X_{\pi}$, carried out in Sec. IV B adopting the GF scale $w_0$, can be repeated in the case of the scales $\sqrt{t_0}$ and $t_0/w_0$. The values of the relative GF scales $w_0/\alpha$, $\sqrt{t_0}/\alpha$ and $t_0/(w_0\alpha)$ have been calculated at the physical pion point in the previous Sec. D 1 and, for sake of clarity, we recollect them in Table X.

| $\beta$ | $w_0/\alpha$ | $\sqrt{t_0}/\alpha$ | $t_0/(w_0\alpha)$ |
|-------|-------------|---------------------|------------------|
| 1.726 | 1.8352 (35) | 1.5660 (22) | 1.3359 (12) |
| 1.778 | 2.1299 (16) | 1.80396 (68) | 1.52789 (33) |
| 1.836 | 2.5045 (17) | 2.1094 (8) | 1.77670 (37) |
Adopting fitting functions similar to the ones given by Eq. (56) used in the case of the scale \(w_0\), we obtain

\[
w_0 = 0.17383(57)_{\text{stat+fit}}^{(26)}_{\text{syst}}(63) \text{ fm}, \tag{D3}
\]

\[
\sqrt{t_0} = 0.14436(54)_{\text{stat+fit}}^{(30)}_{\text{syst}}(61) \text{ fm}, \tag{D4}
\]

\[
t_0/w_0 = 0.11969(52)_{\text{stat+fit}}^{(33)}_{\text{syst}}(62) \text{ fm}. \tag{D5}
\]

The quality of the fitting procedure in the case of the GF scales \(\sqrt{t_0}\) and \(t_0/w_0\) is illustrated in Fig. 13 and it should be compared with the one shown in Fig. 6 in the case of the GF scale \(w_0\).

Some values obtained for the continuum-limit fitting parameters \(f\) and \(\hat{\rho}_{\text{phys}}\) and for the discretization parameters \(D'_0\) and \(D'_t\) are collected in Table XI. It can clearly be seen that the pion mass dependence of \(X_\pi\) in the continuum limit is stable against the choice of the specific GF scale, while the values of the discretization parameters \(D'_0\) and \(D'_t\) depend on the above choice. The discretization effects on \(X_\pi\) appear to be smaller in the case of the GF scale \(t_0/w_0\).

Finally, the values of the lattice spacing \(a\) corresponding to the three GF scales (D3)–(D5) and to the relative scales given in Table X are shown in Table XII. The three determinations of \(a\) differ by \(\mathcal{O}(a^2)\) effects, as shown in Fig. 14. In particular, we get \(a(\sqrt{t_0})/a(w_0) \approx 1 - 0.09(2)a^2(w_0)/w_0^2\) and \(a(t_0/w_0)/a(w_0) \approx 1 - 0.18(2)a^2(w_0)/w_0^2\).
[57] M. Luscher, S. Sint, R. Sommer, and P. Weisz, Nucl. Phys. B478, 365 (1996).
[58] R. Frezzotti et al. (ALPHA Collaboration), J. High Energy Phys. 07 (2001) 048.
[59] M. Luscher and P. Weisz, J. High Energy Phys. 02 (2011) 051.
[60] A. Deuzeman and U. Wenger, Proc. Sci., LATTICE2012 (2012) 162.