F_K/F_π from Möbius domain-wall fermions solved on gradient-flowed HISQ ensembles

Nolan Miller,[1] Henry Monge-Camacho,[1] Chia Cheng Chang (張家丞),[2,3,4] Ben Hörz,[3] Enrico Rinaldi,[5,2] Dean Howarth,[6,3] Evan Berkowitz,[7,8] David A. Brantley,[6] Arjun Singh Gambhir,[9,3] Christopher Körber,[4,3] Christopher J. Monahan,[10,11] M. A. Clark,[12] Bálint Joó,[13] Thorsten Kurth,[12] Amy Nicholson,[1] Kostas Orginos,[10,11] Pavlos Vranas,[6,3] and André Walker-Loud[3,6,4]

[1]*Department of Physics and Astronomy, University of North Carolina, Chapel Hill, North Carolina 27516-3255, USA*
[2]*Interdisciplinary Theoretical and Mathematical Sciences Program (iTHEMS), RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan*
[3]*Nuclear Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[4]*Department of Physics, University of California, Berkeley, California 94720, USA*
[5]*Arithmer Inc., R&D Headquarters, Minato, Tokyo 106-6040, Japan*
[6]*Physics Division, Lawrence Livermore National Laboratory, Livermore, California 94550, USA*
[7]*Department of Physics, University of Maryland, College Park, Maryland 20742, USA*
[8]*Institut für Kernphysik and Institute for Advanced Simulation, Forschungszentrum Jülich, 54245 Jülich Germany*
[9]*Design Physics Division, Lawrence Livermore National Laboratory, Livermore, California 94550, USA*
[10]*Department of Physics, The College of William & Mary, Williamsburg, Virginia 23187, USA*
[11]*Theory Center, Thomas Jefferson National Accelerator Facility, Newport News, Virginia 23606, USA*
[12]*NVIDIA Corporation, 2701 San Tomas Expressway, Santa Clara, California 95050, USA*
[13]*Scientific Computing Group, Thomas Jefferson National Accelerator Facility, Newport News, Virginia 23606, USA*

(Received 16 May 2020; accepted 17 July 2020; published 25 August 2020)

We report the results of a lattice quantum chromodynamics calculation of F_K/F_π using Möbius domain-wall fermions computed on gradient-flowed $N_f = 2+1+1$ highly improved staggered quark (HISQ) ensembles. The calculation is performed with five values of the pion mass ranging from $130 \lesssim m_\pi \lesssim 400$ MeV, four lattice spacings of $a \sim 0.15$, 0.12, 0.09 and 0.06 fm and multiple values of the lattice volume. The interpolation/extrapolation to the physical pion and kaon mass point, the continuum, and infinite volume limits are performed with a variety of different extrapolation functions utilizing both the relevant mixed-action effective field theory expressions as well as discretization-enhanced continuum chiral perturbation theory formulas. We find that the $a \sim 0.06$ fm ensemble is helpful, but not necessary to achieve a subpercent determination of F_K/F_π. We also include an estimate of the strong isospin breaking corrections and arrive at a final result of $F_{\hat{K}^+}/F_{\hat{\pi}^+} = 1.1942(45)$ with all sources of statistical and systematic uncertainty included. This is consistent with the Flavour Lattice Averaging Group average value, providing an important benchmark for our lattice action. Combining our result with experimental measurements of the pion and kaon leptonic decays leads to a determination of $|V_{us}|/|V_{ud}| = 0.2311(10)$.

DOI: 10.1103/PhysRevD.102.034507

I. INTRODUCTION

Leptonic decays of the charged pions and kaons provide a means for probing flavor-changing interactions of the Standard Model (SM). In particular, the SM predicts that the Cabibbo-Kobayashi-Maskawa (CKM) matrix is unitary, providing strict constraints on various sums of the matrix elements. Thus, a violation of these constraints is indicative of new, beyond the SM physics. There is a substantial flavor physics program dedicated to searching indirectly for potential violations.

CKM matrix elements may be determined through a combination of experimental leptonic decay widths and theoretical determinations of the meson decay constants. For example, the ratio of the kaon and pion decay constants, F_K, F_π, respectively, may be related to the ratio

of light and strange CKM matrix elements $|V_{us}|$, $|V_{ud}|$ via [1,2],

$$\frac{\Gamma(K \to l\bar{\nu}_l)}{\Gamma(\pi \to l\bar{\nu}_l)} = \frac{|V_{us}|^2}{|V_{ud}|^2}\frac{F_K^2}{F_\pi^2}\frac{m_K}{m_\pi}\frac{\left(1-\frac{m_l^2}{m_K^2}\right)^2}{\left(1-\frac{m_l^2}{m_\pi^2}\right)^2} \times [1+\delta_{\mathrm{EM}}+\delta_{SU(2)}]. \tag{1.1}$$

In this expression, $l = e, \mu$, the one-loop radiative quantum electrodynamics (QED) correction is δ_{EM} [3,4] and $\delta_{SU(2)}$ is the strong isospin breaking correction that relates F_K^2/F_π^2 in the isospin limit to $F_{K^+}^2/F_{\pi^+}^2$ that includes $m_d - m_u$ corrections [5]

$$\frac{F_{\hat{K}^+}^2}{F_{\hat{\pi}^+}^2} = \frac{F_K^2}{F_\pi^2}[1+\delta_{SU(2)}].$$

Using lattice quantum chromodynamics (QCD) calculations of the ratio of decay constants in the above expression yields one of the most precise determinations of $|V_{us}|/|V_{ud}|$ [6]. Combining the results obtained through lattice QCD with independent determinations of the CKM matrix elements, such as semileptonic meson decays, provides a means for testing the unitarity of the CKM matrix and obtaining signals of new physics.

F_K/F_π is a so-called *gold-plated* quantity [7] for calculating within lattice QCD. This dimensionless ratio skirts the issue of determining a physical scale for the lattices, and gives precise results due to the correlated statistical fluctuations between numerator and denominator, as well as the lack of signal-to-noise issues associated with calculations involving, for instance, nucleons. Lattice QCD calculations of F_K/F_π are now a mature endeavor, with state-of-the-art calculations determining this quantity consistently with subpercent precision. The most recent review by the Flavour Lattice Averaging Group (FLAG), which performs global averages of quantities that have been calculated and extrapolated to the physical point by multiple groups, quotes a value of

$$\frac{F_{\hat{K}^+}}{F_{\pi^+}} = 1.1932(19) \tag{1.2}$$

for $N_f = 2+1+1$ dynamical quark flavors, including strong-isospin breaking corrections [8].

This average includes calculations derived from two different lattice actions, one [9] with twisted-mass fermions [10,11] and the other two [12,13] with the highly improved staggered quark (HISQ) action [14]. The results obtained using the HISQ action are approximately seven times more precise than those from twisted mass and so the universality of the continuum limit for F_K/F_π from $N_f = 2+1+1$ results has not been tested with precision yet: in the continuum limit, all lattice actions should reduce to a single universal limit, that of SM QCD, provided all systematics are properly accounted for. Thus, in addition to lending more confidence to its global average, the calculation of a gold-plated quantity also allows for precise testing of new lattice actions, and the demonstration of control over systematic uncertainties for a given action. FLAG also reports averages for $N_F = 2+1$, $F_{K^\pm}/F_{\pi^\pm} = 1.1917(37)$ from Refs. [15–20] and for $N_f = 2$, $F_{K^\pm}/F_{\pi^\pm} = 1.1205(18)$ from Refs. [21], though we restrict our direct comparisons to the $N_f = 2+1+1$ results just for simplicity.

In this work, we report a new determination of F_K/F_π calculated with Möbius domain-wall fermions computed on gradient-flowed $N_f = 2+1+1$ HISQ ensembles [22]. Our final result in the isospin symmetric limit, Sec. IV D, including a breakdown in terms of statistical (s), pion and kaon mass extrapolation (χ), continuum limit (a), infinite volume limit (V), physical point (phys) and model selection (M) uncertainties, is

$$\frac{F_K}{F_\pi} = 1.1964(32)^s(12)^\chi(20)^a(01)^V(15)^{\mathrm{phys}}(12)^M = 1.1964(44). \tag{1.3}$$

With our estimated strong isospin breaking corrections, Sec. IV E, our result including $m_d - m_u$ effects is

$$\frac{F_{\hat{K}^+}}{F_{\hat{\pi}^+}} = 1.1942(44)(07)^{\mathrm{iso}} = 1.1942(45), \tag{1.4}$$

where the first uncertainty in the first line is the combination of those in Eq. (1.3).

In the following sections we will discuss details of our lattice calculation, including a brief synopsis of the action and ensembles used, as well as our strategy for extracting the relevant quantities from correlation functions. We will then detail our procedure for extrapolating to the physical point via combined continuum, infinite volume, and physical pion and kaon mass limits and the resulting uncertainty breakdown. We discuss the impact of the $a \sim 0.06$ fm ensemble on our analysis, the convergence of the $SU(3)$-flavor chiral expansion, and the estimate of the strong isospin breaking corrections. We conclude with an estimate of the impact our result has on improving the extraction of $|V_{us}|/|V_{ud}|$ and an outlook.

II. DETAILS OF THE LATTICE CALCULATION

A. MDWF on gradient-flowed HISQ

There are many choices for discretizing QCD, with each choice being commonly referred to as a lattice action. These actions correspond to different UV theories that share a common low-energy theory, QCD. Sufficiently close to the continuum limit, the discrete lattice actions can

be expanded as a series of local operators known as the Symanzik expansion [23,24], the low-energy effective field theory (EFT) for the discrete lattice action. The Symanzik EFT contains a series of operators having higher dimension than those in QCD, multiplied by appropriate powers of the lattice spacing, a. For all lattice actions, the only operators of mass-dimension ≤ 4 are those of QCD, such that the explicit effects from the various discretizations are encoded only in higher-dimensional operators which are all irrelevant in the renormalization sense. There is a universality of the continuum limit, $a \to 0$, in that all lattice actions, if calculated using sufficiently small lattice spacing, will recover the target theory of QCD, provided there are no surprises from nonperturbative effects.

Performing lattice QCD (LQCD) calculations with different actions is therefore valuable to test this universality, to help ensure a given action is not accidentally in a different phase of QCD, and to protect against unknown systematic uncertainties arising from a particular calculation with a particular action. In this work, we use a mixed-action [25] in which the discretization scheme for the valence quarks is the Möbius domain-wall fermion (MDWF) action [26–28] while the discretization scheme for the sea quarks is the highly improved staggered quark action [14]. Before solving the MDWF propagators, we apply a gradient-flow [29–31] smoothing algorithm [32,33] to the gluons to dampen UV fluctuations, which also significantly improves the chiral symmetry properties of the MDWF action [22] (for example, the residual chiral symmetry breaking scale of domain-wall fermions m^{res} is held to less than 10% of m_l for reasonable values of L_5 and M_5, see Table I). Our motivation to perform this calculation is to improve our understanding of F_K/F_π and to test the MDWF on gradient-flowed HISQ action we have used to compute the $\pi^- \to \pi^+$ neutrinoless double beta decay matrix elements arising from prospective higher-dimension lepton-number-violating physics [34], and the axial coupling of the nucleon g_A [35,36]. As there is an otherwise straightforward path to determining g_A to sub-percent precision with pre-exascale computing such as Summit at Oak Ridge Leadership Computing Facility (OLCF) and Lassen at Lawrence Livermore National Laboratory (LLNL) [37], it is important to ensure this action is consistent with known results at this level of precision.

There are several motivations for choosing this mixed-action (MA) scheme [25,41]. The MILC Collaboration provides their gauge configurations to any interested party and we have made heavy use of them. They have generated the configurations covering a large parameter space allowing one to fully control the physical pion mass, infinite volume and continuum limit extrapolations [42,43]. The good chiral symmetry properties of the domain wall (DW) action [44–46] significantly suppress sources of chiral symmetry breaking from any sea-quark action, motivating the use of this mixed-action setup. While this action is not unitary at finite lattice spacing, we have tuned the valence quark masses such that the valence pion mass

TABLE I. Input parameters for our lattice action. The abbreviated ensemble name [38] indicates the approximate lattice spacing in fm and pion mass in MeV. The S, L, XL which come after an ensemble name denote a relatively small, large and extra-large volume with respect to $m_\pi L = 4$.

Ensemble	β	N_{cfg}	Volume	am_l	am_s	am_c	L_5/a	aM_5	b_5, c_5	am_l^{val}	$am_l^{\text{res}} \times 10^4$	am_s^{val}	$am_s^{\text{res}} \times 10^4$	σ	N	N_{src}
a15m400[a]	5.80	1000	$16^3 \times 48$	0.0217	0.065	0.838	12	1.3	1.50, 0.50	0.0278	9.365(87)	0.0902	6.937(63)	3.0	30	8
a15m350[a]	5.80	1000	$16^3 \times 48$	0.0166	0.065	0.838	12	1.3	1.50, 0.50	0.0206	9.416(90)	0.0902	6.688(62)	3.0	30	16
a15m310	5.80	1000	$16^3 \times 48$	0.013	0.065	0.838	12	1.3	1.50, 0.50	0.0158	9.563(67)	0.0902	6.640(44)	4.2	45	24
a15m220	5.80	1000	$24^3 \times 48$	0.0064	0.064	0.828	16	1.3	1.75, 0.75	0.00712	5.736(38)	0.0902	3.890(25)	4.5	60	16
a15m135XL[a]	5.80	1000	$48^3 \times 64$	0.002426	0.06730	0.8447	24	1.3	2.25, 1.25	0.00237	2.706(08)	0.0945	1.860(09)	3.0	30	32
a12m400[a]	6.00	1000	$24^3 \times 64$	0.0170	0.0509	0.635	8	1.2	1.25, 0.25	0.0219	7.337(50)	0.0693	5.129(35)	3.0	30	8
a12m350[a]	6.00	1000	$24^3 \times 64$	0.0130	0.0509	0.635	8	1.2	1.25, 0.25	0.0166	7.579(52)	0.0693	5.062(34)	3.0	30	8
a12m310	6.00	1053	$24^3 \times 64$	0.0102	0.0509	0.635	8	1.2	1.25, 0.25	0.0126	7.702(52)	0.0693	4.950(35)	3.0	30	8
a12m220S	6.00	1000	$24^4 \times 64$	0.00507	0.0507	0.628	12	1.2	1.50, 0.50	0.00600	3.990(42)	0.0693	2.390(24)	6.0	90	4
a12m220	6.00	1000	$32^3 \times 64$	0.00507	0.0507	0.628	12	1.2	1.50, 0.50	0.00600	4.050(20)	0.0693	2.364(15)	6.0	90	4
a12m220L	6.00	1000	$40^3 \times 64$	0.00507	0.0507	0.628	12	1.2	1.50, 0.50	0.00600	4.040(26)	0.0693	2.361(19)	6.0	90	4
a12m130	6.00	1000	$48^3 \times 64$	0.00184	0.0507	0.628	20	1.2	2.00, 1.00	0.00195	1.642(09)	0.0693	0.945(08)	3.0	30	32
a09m400[a]	6.30	1201	$32^3 \times 64$	0.0124	0.037	0.44	6	1.1	1.25, 0.25	0.0160	2.532(23)	0.0491	1.957(17)	3.5	45	8
a09m350[a]	6.30	1201	$32^3 \times 64$	0.00945	0.037	0.44	6	1.1	1.25, 0.25	0.0121	2.560(24)	0.0491	1.899(16)	3.5	45	8
a09m310	6.30	780	$32^3 \times 96$	0.0074	0.037	0.44	6	1.1	1.25, 0.25	0.00951	2.694(26)	0.0491	1.912(15)	6.7	167	8
a09m220	6.30	1001	$48^3 \times 96$	0.00363	0.0363	0.43	8	1.1	1.25, 0.25	0.00449	1.659(13)	0.0491	0.834(07)	8.0	150	6
a09m135[a]	6.30	1010	$64^3 \times 96$	0.001326	0.03636	0.4313	12	1.1	1.50, 0.50	0.00152	0.938(06)	0.04735	0.418(04)	3.5	45	16
a06m310[a]	6.72	1000	$72^3 \times 96$	0.0048	0.024	0.286	6	1.0	1.25, 0.25	0.00617	0.225(03)	0.0309	0.165(02)	3.5	45	8

[a]Additional ensembles generated by CalLat using the MILC code. The m350 and m400 ensembles were made on the Vulcan supercomputer at LLNL while the a15m135XL, a09m135, and a06m310L ensembles were made on the Sierra and Lassen supercomputers at LLNL and the Summit supercomputer at OLCF using QUDA[39,40]. These configurations are available to any interested party upon request, and will be available for easy anonymous downloading—hopefully soon.

matches the taste-5 HISQ pion mass within a few percent, so as the continuum limit is taken, we recover a unitary theory.

EFT can be used to understand the salient features of such Mixed-action LQCD (MALQCD) calculations. Chiral perturbation theory (χPT) [47–49] can be extended to incorporate discretization effects into the analytic formula describing the quark-mass dependence of various hadronic quantities [50]. The MA EFT [51] for DW valence fermions on dynamical rooted staggered fermions is well developed [52–59]. The use of valence fermions which respect chiral symmetry leads to a universal form of the MA EFT extrapolation formulas at next-to-leading order (NLO) in the joint quark mass and lattice spacing expansions [55,58], which follows from the suppression of chiral symmetry breaking discretization effects.

B. Correlation function construction and analysis

The correlation function construction and analysis follows closely the strategy of Refs. [22] and [35,36]. Here we summarize the relevant details for this work.

The pseudoscalar decay constants F can be obtained from standard two-point correlation functions by making use of the 5D Ward-Takahashi identity [60,61]

$$F^{q_1 q_2} = z_{0p}^{q_1 q_2} \frac{m^{q_1} + m_{\text{res}}^{q_1} + m^{q_2} + m_{\text{res}}^{q_2}}{(E_0^{q_1 q_2})^{3/2}}, \tag{2.1}$$

where q_1 and q_2 denote the quark content of the meson with lattice input masses m_{q_1} and m_{q_2} respectively. The point-sink ground-state overlap-factor z_{0p} and ground-state energy E_0 are extracted from a two-point correlation function analysis with the model

$$C_{(ss)ps}^{q_1 q_2}(t) = \sum_n z_{n(s)p}^{q_1 q_2} z_{ns}^{q_1 q_2 \dagger} (e^{-E_n^{q_1 q_2} t} + e^{-E_n^{q_1 q_2}(T-t)}), \tag{2.2}$$

where n encompasses in general an infinite tower of states, t is the source-sink time separation, T is the temporal box size and we have both smeared (s) and point (p) correlation functions which both come from smeared sources. From Ref. [22], we show that gradient-flow smearing leads to the suppression of the domain-wall fermion oscillating mode (which also decouples as $M_5 \to 1$, at least in free-field [62]), and therefore this mode is not included in the correlator fit model. Finally, the residual chiral symmetry breaking m_{res} is calculated by the ratio of two-point correlation functions evaluated at the midpoint of the fifth dimension $L_5/2$ and bounded on the domain wall [28]

$$m_{\text{res}}(t) = \frac{\sum_{\mathbf{x}} \langle \pi(t, \mathbf{x}, L_5/2) \pi(0, \mathbf{0}, 0) \rangle}{\sum_{\mathbf{x}} \langle \pi(t, \mathbf{x}, 0) \pi(0, \mathbf{0}, 0) \rangle}, \tag{2.3}$$

where $\pi(t, \mathbf{x}, s) \equiv \bar{q}(t, \mathbf{x}, w) \gamma_5 q(t, \mathbf{x}, w)$ is the pseudoscalar interpolating operator at time t, space $\mathbf{x}$ and fifth dimension s. We extract m_{res} by fitting Eq. (2.3) to a constant.

1. Analysis strategy

For all two-point correlation function parameters (MDWF and mixed MDWF-HISQ), we infer posterior parameter distributions in a Bayesian framework using a 4-state model which simultaneously describes the smeared- and point-sink two-point correlation functions (the source is always smeared). The joint posterior distribution is approximated by a multivariate normal distribution (we later refer to this procedure as *fitting*). The two-point correlation functions are folded in time to double the statistics. The analysis of the pion, kaon, $\bar{s}\gamma_5 s$, and mixed MDWF-HISQ mesons are performed independently, with correlations accounted for under bootstrap resampling.

We analyze data of source-sink time separations between 0.72 and 3.6 fm for all 0.09 and 0.12 fm lattice spacing two-point correlation functions, and 0.75 to 3.6 fm for all 0.15 fm lattice spacing two-point correlation functions.

We choose normally distributed priors for the ground-state energy and all overlap factors, and log-normal distributions for excited-state energy priors. The ground-state energy and overlap factors are motivated by the plateau values of the effective masses and scaled correlation function, and a prior width of 10% of the central value. The excited-state energy splittings are set to the value of two pion masses with a width allowing for fluctuations down to one pion mass within one standard deviation. The excited-state overlap factors are set to zero, with a width set to the mean value of the ground-state overlap factor.

Additionally, we fit a constant to the correlation functions in Eq. (2.3). For the 0.09 to 0.12 fm ensembles, we analyze source-sink separations that are greater than 0.72 fm. For the 0.12 fm ensemble, the minimum source-sink separation is 0.75 fm. The prior distribution for the residual chiral symmetry breaking parameter is set to the observed value per ensemble, with a width that is 100% of the central value. The uncertainty is propagated with bootstrap resampling.

We emphasize that all input fit parameters (i.e., number of states, fit region, priors) are chosen to have the same values in physical units for all observables, to the extent that a discretized lattice allows. Additionally, we note that the extracted ground-state observables from these correlation functions are insensitive to variations around the chosen set of input fit parameters. Figure 1 shows the stability of the determination of E_0 for the pion and kaon on the a12m130 ensemble versus t_{min} and the number of states.

III. EXTRAPOLATION FUNCTIONS

We now turn to the extrapolation/interpolation to the physical point. We have three ensembles at the physical pion mass with relatively high statistics and precise

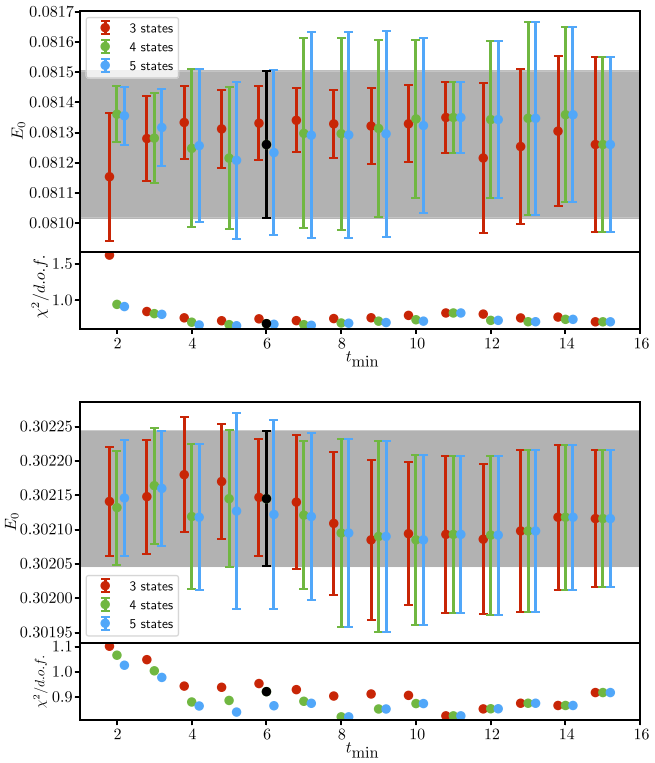

FIG. 1. Stability of the ground-state mass determination of the pion (top) and kaon (bottom) on the a12m130 ensemble. The x axis is the value of $t_{\min}$ used in the analysis and the resulting E_0 for a given $t_{\min}$ and number of states in the analysis is plotted. The 68% confidence interval of the chosen fit (black) is plotted as a horizontal band to guide the eye.

determinations of F_K/F_π (a15m135XL, a12m130, and a09m135, see Table II) such that the physical quark mass extrapolation is an interpolation. Nevertheless, we explore how the ensembles with heavier pion masses impact the physical point prediction and we use our dataset to explore uncertainty arising in the $SU(3)$-flavor chiral expansion.

We begin by assuming a canonical power-counting scheme for our MA LQCD action [52] in which $\mathrm{O}(m_\pi^2) \sim \mathrm{O}(m_K^2) \sim \mathrm{O}(a^2\Lambda_{\mathrm{QCD}}^4)$ are all treated as small scales. For the quark mass expansion, the dimensionless small parameters $(m_P/4\pi F)^2$ naturally emerge from χPT where $P \in \{\pi, K, \eta\}$. For the discretization corrections, while $a\Lambda_{\mathrm{QCD}}^2$ is often used to estimate the relative size of corrections compared to typical hadronic mass scales, it is a bit unnatural to use this in a low-energy EFT as Λ_{QCD} is a QCD scale that does not emerge in χPT.

We chose to use another hadronic scale to form a dimensionless parameter with a, that being the gradient flow scale $w_0 \sim 0.17$ fm [63]. This quantity is easy to compute, has mild quark mass dependence, and the value is roughly $w_0^{-1} \simeq 4\pi F_\pi$. We then define the dimensionless small parameters for controlling the expansion to be

$$\epsilon_P^2 = \left(\frac{m_P}{\Lambda_\chi}\right)^2 = \left(\frac{m_P}{4\pi F}\right)^2, \qquad \epsilon_a^2 = \left(\frac{a}{2w_0}\right)^2. \tag{3.1}$$

We leave F ambiguous, as we will explore taking $F = F_\pi$, $F = F_K$ and $F^2 = F_\pi F_K$ in our definition of Λ_χ. This particular choice of ϵ_a is chosen such that the range of

TABLE II. Extracted masses and decay constants from correlation functions. An HDF5 file is provided with this publication which includes the resulting bootstrap samples which can be used to construct the correlated uncertainties. The small parameters in this table are defined as $\epsilon_{\pi,K} = m_{\pi,K}/(4\pi F_\pi)$, $\epsilon_a = a/(2w_0)$. The normalization of ϵ_a is chosen such that as a small parameter, it spans the range of $\epsilon_\pi^2 \lesssim \epsilon_a^2 \lesssim \epsilon_K^2$.

Ensemble	am_π	am_K	ϵ_π^2	ϵ_K^2	$m_\pi L$	ϵ_a^2	α_S	aF_π	aF_K	F_K/F_π
a15m400	0.30281(31)	0.42723(27)	0.09216(33)	0.18344(62)	4.85	0.19378(13)	0.58801	0.07938(12)	0.08504(09)	1.0713(09)
a15m350	0.26473(30)	0.41369(28)	0.07505(28)	0.18326(60)	4.24	0.19378(13)	0.58801	0.07690(11)	0.08370(09)	1.0884(09)
a15m310	0.23601(29)	0.40457(25)	0.06223(17)	0.18285(48)	3.78	0.19378(13)	0.58801	0.07529(09)	0.08293(09)	1.1015(13)
a15m220	0.16533(19)	0.38690(21)	0.03269(11)	0.17901(48)	3.97	0.19378(13)	0.58801	0.07277(08)	0.08196(10)	1.1263(15)
a15m135XL	0.10293(07)	0.38755(14)	0.01319(05)	0.18704(59)	4.94	0.19378(13)	0.58801	0.07131(11)	0.08276(10)	1.1606(18)
a12m400	0.24347(16)	0.34341(14)	0.08889(30)	0.17685(63)	5.84	0.12376(18)	0.53796	0.06498(11)	0.06979(07)	1.0739(17)
a12m350	0.21397(20)	0.33306(16)	0.07307(37)	0.17704(83)	5.14	0.12376(18)	0.53796	0.06299(14)	0.06851(07)	1.0876(27)
a12m310	0.18870(17)	0.32414(21)	0.05984(25)	0.17657(69)	4.53	0.12376(18)	0.53796	0.06138(11)	0.06773(10)	1.1033(21)
a12m220S	0.13557(32)	0.31043(22)	0.03384(19)	0.1774(10)	3.25	0.12376(18)	0.53796	0.05865(16)	0.06673(11)	1.1378(27)
a12m220L	0.13402(15)	0.31021(19)	0.03289(15)	0.17621(79)	5.36	0.12376(18)	0.53796	0.05881(13)	0.06631(17)	1.1276(29)
a12m220	0.13428(17)	0.31001(17)	0.03314(15)	0.17666(81)	4.30	0.12376(18)	0.53796	0.05870(13)	0.06636(11)	1.1306(22)
a12m130	0.08126(16)	0.30215(11)	0.01287(08)	0.17788(71)	3.90	0.12376(18)	0.53796	0.05701(11)	0.06624(08)	1.1619(21)
a09m400	0.18116(15)	0.25523(13)	0.08883(32)	0.17633(59)	5.80	0.06515(08)	0.43356	0.04837(08)	0.05229(07)	1.0810(09)
a09m350	0.15785(20)	0.24696(12)	0.07256(32)	0.17761(68)	5.05	0.06515(08)	0.43356	0.04663(08)	0.05127(07)	1.0994(10)
a09m310	0.14072(12)	0.24106(14)	0.06051(22)	0.17757(59)	4.50	0.06515(08)	0.43356	0.04552(07)	0.05053(08)	1.1101(16)
a09m220	0.09790(06)	0.22870(09)	0.03307(14)	0.18045(70)	4.70	0.06515(08)	0.43356	0.04284(08)	0.04899(07)	1.1434(18)
a09m135	0.05946(06)	0.21850(08)	0.01346(08)	0.18175(91)	3.81	0.06515(08)	0.43356	0.04079(10)	0.04804(06)	1.1778(22)
a06m310L	0.09456(06)	0.16205(07)	0.06141(35)	0.1803(10)	6.81	0.02726(03)	0.29985	0.03037(08)	0.03403(07)	1.1205(17)

values of this small parameter roughly corresponds to $\epsilon_\pi^2 \lesssim \epsilon_a^2 \lesssim \epsilon_K^2$ as the lattice spacing is varied, similar to the variation of ϵ_π^2 itself over the range of pion masses used, see Table II. As we will discuss in Sec. IV, this choice of ϵ_a seems natural as determined by the size of the discretization low-energy constants (LECs) which are found in the analysis. Note, this differs from the choice used in our analysis of g_A [35,36].

With this power-counting scheme, the different orders in the expansion are defined to be

$$\begin{aligned}
&\text{NLO}: \mathrm{O}(\epsilon_P^2) \sim \mathrm{O}(\epsilon_a^2), \\
&\text{N}^2\text{LO}: \mathrm{O}(\epsilon_P^4) \sim \mathrm{O}(\epsilon_P^2\epsilon_a^2) \sim \mathrm{O}(\epsilon_a^4), \\
&\text{N}^3\text{LO}: \mathrm{O}(\epsilon_P^6) \sim \mathrm{O}(\epsilon_P^4\epsilon_a^2) \sim \mathrm{O}(\epsilon_P^2\epsilon_a^4) \sim \mathrm{O}(\epsilon_a^6).
\end{aligned} \tag{3.2}$$

Even at finite lattice spacing, $F_K = F_\pi$ in the $SU(3)$ flavor symmetry limit, also known as the $SU(3)$ vector limit $SU(3)_V$, and so there cannot be a pure $\mathrm{O}(\epsilon_a^2)$ correction as it must accompany terms which vanish in the $SU(3)_V$ limit, such as $\epsilon_K^2 - \epsilon_\pi^2$. Therefore, at NLO, there cannot be any counterterms proportional to ϵ_a^2 and the only discretization effects that can appear at NLO come through modification of the various meson masses that appear in the MA EFT.

We find that the precision of our results requires including terms higher than NLO, and we have to work at a hybrid N^{3}LO order to obtain a good description of our data. Therefore, we will begin with a discussion of the full N^{2}LO χPT theory expression for F_K/F_π in the continuum limit [64–67].

A. N^{2}LO χPT

The analytic expression for F_K/F_π up to N^{2}LO is [67]

$$\begin{aligned}
\frac{F_K}{F_\pi} = 1 &+ \frac{5}{8}\ell_\pi - \frac{1}{4}\ell_k - \frac{3}{8}\ell_\eta + 4\bar{L}_5(\epsilon_K^2 - \epsilon_\pi^2) \\
&+ \epsilon_K^4 F_F\left(\frac{m_\pi^2}{m_K^2}\right) + \hat{K}_1^r\lambda_\pi^2 + \hat{K}_2^r\lambda_\pi\lambda_K \\
&+ \hat{K}_3^r\lambda_\pi\lambda_\eta + \hat{K}_4^r\lambda_K^2 + \hat{K}_5^r\lambda_K\lambda_\eta + \hat{K}_6^r\lambda_\eta^2 \\
&+ \hat{C}_1^r\lambda_\pi + \hat{C}_2^r\lambda_K + \hat{C}_3^r\lambda_\eta + \hat{C}_4^r.
\end{aligned} \tag{3.3}$$

The first line is the LO (1) plus NLO terms, while the next three lines are the N^{2}LO terms. Several nonunique choices were made to arrive at this formula. Prior to discussing these choices, we first define the parameters appearing in Eq. (3.3). First, the small parameters were all defined as

$$\epsilon_P^2 = \left(\frac{m_P}{4\pi F_\pi(m_P)}\right)^2, \tag{3.4}$$

where $F_\pi(m_P)$ is the "on-shell" pion decay constant at the masses m_P. The quantities ℓ_P are defined as

$$\ell_P = \epsilon_P^2 \ln\left(\frac{m_P^2}{\mu^2}\right), \tag{3.5}$$

where μ is a renormalization scale. The coefficient $\bar{L}_5 = (4\pi)^2 L_5^r(\mu)$ is one of the regulated Gasser-Leutwyler LECs [68] which has a renormalization scale dependence that exactly cancels against the dependence arising from the logarithms appearing at the same order. In the following, we define all of the Gasser-Leutwyler LECs with the extra $(4\pi)^2$ for convenience:

$$\bar{L}_i \equiv (4\pi)^2 L_i^r(\mu). \tag{3.6}$$

The η mass has been defined through the Gell-Mann–Okubo (GMO) relation

$$m_\eta^2 \equiv \frac{4}{3}m_K^2 - \frac{1}{3}m_\pi^2, \tag{3.7}$$

with the corrections to this relation being propagated into Eq. (3.3) for consistency at N^{2}LO. The logs are

$$\lambda_P \equiv \ln\left(\frac{m_P^2}{\mu^2}\right). \tag{3.8}$$

The $\ln^2$ terms are encapsulated in the $F_F(x)$ function, defined in Eqs. (8-17) of Ref. [67],[1] and the $\hat{K}_i^r\lambda_P\lambda_{P'}$ terms whose coefficients are given by[2]

$$\begin{aligned}
\hat{K}_1^r &= \frac{11}{24}\epsilon_\pi^2\epsilon_K^2 - \frac{131}{192}\epsilon_\pi^4, & \hat{K}_2^r &= -\frac{41}{96}\epsilon_\pi^2\epsilon_K^2 - \frac{3}{32}\epsilon_\pi^4, \\
\hat{K}_3^r &= \frac{13}{24}\epsilon_\pi^2\epsilon_K^2 + \frac{59}{96}\epsilon_\pi^4, & \hat{K}_4^r &= \frac{17}{36}\epsilon_K^4 + \frac{7}{144}\epsilon_\pi^2\epsilon_K^2, \\
\hat{K}_5^r &= -\frac{163}{144}\epsilon_K^4 - \frac{67}{288}\epsilon_\pi^2\epsilon_K^2 + \frac{3}{32}\epsilon_\pi^4, \\
\hat{K}_6^r &= \frac{241}{288}\epsilon_K^4 - \frac{13}{72}\epsilon_\pi^2\epsilon_K^2 - \frac{61}{192}\epsilon_\pi^4.
\end{aligned} \tag{3.9}$$

The single log coefficients $\hat{C}_{1-3}^r$ are combinations of the NLO Gasser-Leutwyler coefficients

$$\hat{C}_i^r = c_i^{\pi\pi}\epsilon_\pi^4 + c_i^{K\pi}\epsilon_K^2\epsilon_\pi^2 + c_i^{KK}\epsilon_K^4, \tag{3.10}$$

where

[1]They also provide an approximate formula which is easy to implement, but our numerical results are sufficiently precise to require the exact expression. To implement this function in our analysis, we have modified an interface C++ file provided by Bijnens to CHIRON [69], the package for two-loop χPT functions. We have provided a PYTHON interface as well so that the function can be called from our main analysis code, which is provided with this article.

[2]We correct a typographical error in the K_6^r term presented in Ref. [67]: a simple power-counting reveals the $\xi_K^2 = \epsilon_K^4$ accompanying this term should not be there.

$$
\begin{aligned}
c_1^{\pi\pi} &= -\frac{113}{72} - 2(2\bar{L}_1 + 5\bar{L}_2) - \frac{13}{2}\bar{L}_3 + \frac{21}{2}\bar{L}_5, \\
c_1^{K\pi} &= -\frac{7}{9} - \frac{11}{2}\bar{L}_5, \\
c_1^{KK} &= c_2^{\pi\pi} = 0, \\
c_2^{K\pi} &= \frac{209}{144} + 3\bar{L}_5, \\
c_2^{KK} &= \frac{53}{96} + 2(2\bar{L}_1 + 5\bar{L}_2) + 5\bar{L}_3 - 5\bar{L}_5, \\
c_3^{\pi\pi} &= \frac{19}{288} + \frac{1}{6}\bar{L}_3 + \frac{11}{6}\bar{L}_5 - 8(2\bar{L}_7 + \bar{L}_8), \\
c_3^{K\pi} &= -\frac{4}{9} - \frac{4}{3}\bar{L}_3 - \frac{25}{6}\bar{L}_5 + 16(2\bar{L}_7 + \bar{L}_8), \\
c_3^{KK} &= \frac{13}{18} + \frac{8}{3}\bar{L}_3 - \frac{2}{3}\bar{L}_5 - 8(2\bar{L}_7 + \bar{L}_8).
\end{aligned} \tag{3.11}
$$

Finally, $\hat{C}_4^r$ is a combination of these L_i^r coefficients as well as counterterms appearing at N^{2}LO. At N^{2}LO, only two counterterm structures can appear due to the $SU(3)_V$ constraints:

$$
\hat{C}_4^r = (\epsilon_K^2 - \epsilon_\pi^2)[(A_K^4 + L_K^4)\epsilon_K^2 + (A_\pi^4 + L_\pi^4)\epsilon_\pi^2] \tag{3.12}
$$

which are linear combinations of the N^{2}LO counterterms

$$
\begin{aligned}
A_K^4 &= 16(4\pi)^4(C_{14}^r + C_{15}^r), \\
A_\pi^4 &= 8(4\pi)^4(C_{15}^r + 2C_{17}^r),
\end{aligned} \tag{3.13}
$$

and contributions from the Gasser-Leutwyler LECs (Eq. 7 of Ref. [67])

$$
\begin{aligned}
L_K^4 &= 8\bar{L}_5(8(\bar{L}_4 - 2\bar{L}_6) + 3\bar{L}_5 - 8\bar{L}_8) \\
&\quad - 2\bar{L}_1 - \bar{L}_2 - \frac{1}{18}\bar{L}_3 + \frac{4}{3}\bar{L}_5 - 8(2\bar{L}_7 + \bar{L}_8), \\
L_\pi^4 &= 8\bar{L}_5(4(\bar{L}_4 - 2\bar{L}_6) + 5\bar{L}_5 - 8\bar{L}_8) \\
&\quad - 2\bar{L}_1 - \bar{L}_2 - \frac{5}{18}\bar{L}_3 - \frac{4}{3}\bar{L}_5 + 8(2\bar{L}_7 + \bar{L}_8).
\end{aligned} \tag{3.14}
$$

There were several nonunique choices that went into the determination of Eq. (3.3). When working with the full N^{2}LO χPT expression, the different choices one can make result in different N^{3}LO or higher corrections and exploring these different choices in the analysis will expose sensitivity to higher-order contributions that are not explicitly included. The first choice we discuss is the Taylor expansion of the ratio of F_K/F_π

$$
\begin{aligned}
\frac{F_K}{F_\pi} &= \frac{1 + \delta F_K^{\mathrm{NLO}} + \delta F_K^{\mathrm{N^2LO}} + \cdots}{1 + \delta F_\pi^{\mathrm{NLO}} + \delta F_\pi^{\mathrm{N^2LO}} + \cdots} \\
&= 1 + \delta F_{K-\pi}^{\mathrm{NLO}} + \delta F_{K-\pi}^{\mathrm{N^2LO}} \\
&\quad + (\delta F_\pi^{\mathrm{NLO}})^2 - \delta F_\pi^{\mathrm{NLO}}\delta F_K^{\mathrm{NLO}} + \cdots,
\end{aligned} \tag{3.15}
$$

where the $\cdots$ represent higher order terms in the expansion and $\delta F_{K-\pi}^{\mathrm{N^2LO}} = \delta F_K^{\mathrm{N^2LO}} - \delta F_\pi^{\mathrm{N^2LO}}$. Equation (3.3) has been derived from this standard Taylor-expanded form with the choices mentioned above: the use of the on-shell renormalized value of $F \to F_\pi$ and the definition of the η mass through the GMO relation. The NLO expressions are the standard ones [68],

$$
\begin{aligned}
\delta F_K^{\mathrm{NLO}} &= -\frac{3}{8}\ell_\pi - \frac{3}{4}\ell_K - \frac{3}{8}\ell_\eta + 4\bar{L}_5\epsilon_K^2 \\
&\quad + 4\bar{L}_4(\epsilon_\pi^2 + 2\epsilon_K^2), \\
\delta F_\pi^{\mathrm{NLO}} &= -\ell_\pi - \frac{1}{2}\ell_K + 4\bar{L}_5\epsilon_\pi^2 + 4\bar{L}_4(\epsilon_\pi^2 + 2\epsilon_K^2), \\
\delta F_{K-\pi}^{\mathrm{NLO}} &= \frac{5}{8}\ell_\pi - \frac{1}{4}\ell_K - \frac{3}{8}\ell_\eta + 4\bar{L}_5(\epsilon_K^2 - \epsilon_\pi^2).
\end{aligned} \tag{3.16}
$$

The $\delta F_P^{\mathrm{N^2LO}}$ terms have been determined in Ref. [64] and cast into analytic forms in Refs. [65,66]. The NLO terms are of O(20%) and so Taylor expanding this ratio leads to sizable corrections from the $(\delta F_\pi^{\mathrm{NLO}})^2 - \delta F_\pi^{\mathrm{NLO}}\delta F_K^{\mathrm{NLO}}$ contributions. Utilizing the full ratio expression could in principle lead to a noticeable difference in the analysis (a different determination of the values of the LECs for example). Rather than implementing the full $\delta F_P^{\mathrm{N^2LO}}$ expressions for kaon and pion, we explore this convergence by instead just resumming the NLO terms which will dominate the potential differences in higher-order corrections. A consistent expression at N^{2}LO is

$$
\frac{F_K}{F_\pi}[(3.3)] = \frac{1 + \delta F_K^{\mathrm{NLO}}}{1 + \delta F_\pi^{\mathrm{NLO}}} + \delta F^{\mathrm{N^2LO}} + \delta_{\mathrm{ratio}}^{\mathrm{N^2LO}} + \cdots, \tag{3.17}
$$

where $\delta F^{\mathrm{N^2LO}}$ is the full N^{2}LO expression in Eq. (3.15)

$$
\delta F^{\mathrm{N^2LO}} = \delta F_{K-\pi}^{\mathrm{N^2LO}} + (\delta F_\pi^{\mathrm{NLO}})^2 - \delta F_\pi^{\mathrm{NLO}}\delta F_K^{\mathrm{NLO}}, \tag{3.18}
$$

and the ratio correction is given by

$$
\delta_{\mathrm{ratio}}^{\mathrm{N^2LO}} = \delta F_\pi^{\mathrm{NLO}}\delta F_K^{\mathrm{NLO}} - (\delta F_\pi^{\mathrm{NLO}})^2. \tag{3.19}
$$

Another choice we explore is the use of $F \to F_\pi$ in the definition of the small parameters. Such a choice is very convenient as it allows one to express the small parameters entirely in terms of observables one can determine in the lattice calculation (unlike the bare parameters, such as χPT's F_0 and Bm_q, which must be determined through extrapolation analysis). Equally valid, one could have

chosen $F \to F_K$ or $F^2 \to F_\pi F_K$. Each choice induces explicit corrections one must account for at N^{2}LO to have a consistent expression at this order. The NLO corrections in Eq. (3.3) are proportional to

$$\begin{aligned}\epsilon_P^2 &= \frac{m_P^2}{(4\pi F_\pi)^2} \\ &= \frac{m_P^2}{(4\pi)^2 F_\pi F_K}\frac{F_K}{F_\pi} = \frac{m_P^2}{(4\pi)^2 F_\pi F_K}(1+\delta F_{K-\pi}^{\rm NLO}) \\ &= \frac{m_P^2}{(4\pi F_K)^2}\frac{F_K^2}{F_\pi^2} = \frac{m_P^2}{(4\pi F_K)^2}(1+2\delta F_{K-\pi}^{\rm NLO}),\end{aligned} \tag{3.20}$$

plus higher-order corrections.

Related to this choice, Eq. (3.3) is implicitly defined at the standard renormalization scale [67]

$$\mu_0^\rho = m_\rho = 770\ \text{MeV}. \tag{3.21}$$

While F_K/F_π of course does not depend upon this choice, the numerical values of the LECs do. Further, a scale setting would be required to utilize this or any fixed value of μ. Instead, as was first advocated in Ref. [70] to the best of our knowledge, it is more convenient to set the renormalization scale on each ensemble with a lattice quantity. For example, Ref. [70] used $\mu = f_\pi^{\rm latt} = \sqrt{2}F_\pi^{\rm latt}$ where $F_\pi^{\rm latt}$ is the lattice-determined value of the pion decay constant on a given ensemble. The advantage of this choice is that the entire extrapolation can be expressed in terms of ratios of lattice quantities such that a scale setting is not required to perform the extrapolation to the physical point.

At NLO in the expansion, one is free to make this choice as the corrections appear at N^{2}LO. In the present work, we must account for these corrections for a consistent expression at this order, which is still defined at a fixed renormalization scale. To understand these corrections, we take as our fixed scale

$$\mu_0 = 4\pi F_0, \tag{3.22}$$

where F_0 is the decay constant in the $SU(3)$ chiral limit. Define $\mu_\pi = 4\pi F_\pi$ and consider the NLO expression

$$\begin{aligned}\frac{F_K}{F_\pi} &= 1 + \frac{5}{8}\ell_\pi^{\mu_0} - \frac{1}{4}\ell_K^{\mu_0} - \frac{3}{8}\ell_\eta^{\mu_0} + 4(\epsilon_K^2-\epsilon_\pi^2)\bar{L}_5(\mu_0) \\ &= +\frac{5}{8}\epsilon_\pi^2 \ln\left(\epsilon_\pi^2\frac{\mu_\pi^2}{\mu_0^2}\right) - \frac{1}{4}\epsilon_K^2\ln\left(\epsilon_K^2\frac{\mu_\pi^2}{\mu_0^2}\right) \\ &\quad - \frac{3}{8}\epsilon_\eta^2\ln\left(\epsilon_\eta^2\frac{\mu_\pi^2}{\mu_0^2}\right) + 4(\epsilon_K^2-\epsilon_\pi^2)\bar{L}_5(\mu_0) \\ &= 1 + \frac{5}{8}\ell_\pi^{\mu_\pi} - \frac{1}{4}\ell_K^{\mu_\pi} - \frac{3}{8}\ell_\eta^{\mu_\pi} + 4(\epsilon_K^2-\epsilon_\pi^2)\bar{L}_5(\mu_0) \\ &\quad + \ln\left(\frac{\mu_\pi^2}{\mu_0^2}\right)\left[\frac{5}{8}\epsilon_\pi^2 - \frac{1}{4}\epsilon_K^2 - \frac{3}{8}\epsilon_\eta^2\right],\end{aligned} \tag{3.23}$$

where we have introduced the notation

$$\ell_P^\mu = \epsilon_P^2 \ln\left(\frac{\epsilon_P^2}{\mu^2}\right). \tag{3.24}$$

If we chose the renormalization scale μ_π and add the second term of the last equality, then this expression is equivalent to working with the scale μ_0 through N^{2}LO. The convenience of this choice becomes clear as μ_π/μ_0 has a familiar expansion

$$\frac{\mu_\pi}{\mu_0} = 1 + \delta F_\pi^{\rm NLO} + \cdots. \tag{3.25}$$

Using the GMO relation Eq. (3.7) and expanding $\ln(1+x)$ for small x, this expression becomes

$$\begin{aligned}\frac{F_K}{F_\pi} &= 1 + \frac{5}{8}\ell_\pi^{\mu_\pi} - \frac{1}{4}\ell_K^{\mu_\pi} - \frac{3}{8}\ell_\eta^{\mu_\pi} + 4(\epsilon_K^2-\epsilon_\pi^2)\bar{L}_5(\mu_0) \\ &\quad - \frac{3}{2}(\epsilon_K^2-\epsilon_\pi^2)\delta F_\pi^{\rm NLO}.\end{aligned} \tag{3.26}$$

Similar expressions can be derived for the choices $\mu_{\pi K} = 4\pi F_{\pi K}$ (where $F_{\pi K} = \sqrt{F_\pi F_K}$) and $\mu_K = 4\pi F_K$ which are made more convenient if one also makes the replacements $F_\pi^2 \to \{F_\pi F_K, F_K^2\}$ in the definition of the small parameters plus the corresponding N^{2}LO corrections that accompany these choices.

If we temporarily expose the implicit dependence of the expression for F_K/F_π on the choices of F and μ, such that Eq. (3.3) is defined as

$$\frac{F_K}{F_\pi}[(3.3)] = \frac{F_K}{F_\pi}(F_\pi, \mu_0^\rho), \tag{3.27}$$

then the following expressions are all equivalent at N^{2}LO:

$$\begin{aligned}\frac{F_K}{F_\pi}(F_\pi,\mu_0) &= \frac{F_K}{F_\pi}(F_K,\mu_K) + \delta_{F_K}^{\rm N^2LO} \\ &= \frac{F_K}{F_\pi}(F_{\pi K},\mu_{\pi K}) + \delta_{F_{\pi K}}^{\rm N^2LO} \\ &= \frac{F_K}{F_\pi}(F_\pi,\mu_\pi) + \delta_{F_\pi}^{\rm N^2LO},\end{aligned} \tag{3.28}$$

where

$$\begin{aligned}\delta_{F_K}^{\rm N^2LO} &= -\frac{3}{2}(\epsilon_K^2-\epsilon_\pi^2)\delta F_K^{\rm NLO} + 2(\delta F_{K-\pi}^{\rm NLO})^2, \\ \delta_{F_{\pi K}}^{\rm N^2LO} &= -\frac{3}{4}(\epsilon_K^2-\epsilon_\pi^2)(\delta F_K^{\rm NLO} + \delta F_\pi^{\rm NLO}) + (\delta F_{K-\pi}^{\rm NLO})^2, \\ \delta_{F_\pi}^{\rm N^2LO} &= -\frac{3}{2}(\epsilon_K^2-\epsilon_\pi^2)\delta F_\pi^{\rm NLO},\end{aligned} \tag{3.29}$$

and the LECs in these expressions are related to those at the standard scale by evolving them from $\mu_0^\rho \to \mu_0$ with their

known scale dependence [68]. Implicit in these expressions is the normalization of the small parameters

$$\epsilon_P^2 = \begin{cases} \frac{m_P^2}{(4\pi F_\pi)^2}, & \text{for } F \to F_\pi \\ \frac{m_P^2}{(4\pi)^2 F_\pi F_K}, & \text{for } F \to \sqrt{F_\pi F_K}. \\ \frac{m_P^2}{(4\pi F_K)^2}, & \text{for } F \to F_K \end{cases} \tag{3.30}$$

We have described several choices one can make in parametrizing the χPT formula for F_K/F_π. The key point is that if the underlying chiral expansion is well behaved, the formulas resulting from each choice are all equivalent through N^{2}LO in the $SU(3)$ chiral expansion, with differences only appearing at N^{3}LO and beyond. Therefore, by studying the variance in the extrapolated answer upon these choices, one is assessing some of the uncertainty arising from the truncation of the chiral extrapolation formula.

B. Discretization corrections

We now turn to the discretization corrections. We explore two parametrizations for incorporating the corrections arising at finite lattice spacing. The simplest approach is to use the continuum extrapolation formula and enhance it by adding contributions from all allowed powers of ϵ_P^2 and ϵ_a^2 to a given order in the expansion. This is very similar to including only the contributions from local counterterms that appear at the given order. At N^{2}LO, the set of discretization corrections is given by[3]

$$\delta_a^{\mathrm{N^2LO}} = A_s^4 \epsilon_a^2 (\epsilon_K^2 - \epsilon_\pi^2) + A_{\alpha_s}^4 \alpha_S \epsilon_a^2 (\epsilon_K^2 - \epsilon_\pi^2), \tag{3.32}$$

where A_s^4 and $A_{\alpha_s}^4$ are the LECs and α_S is the running QCD coupling that emerges in the Symanzik expansion of the lattice expansion through loop corrections. Each contribution at this order must vanish in the $SU(3)_V$ limit because the discretization corrections are flavor blind and so we have the limiting constraint

$$\lim_{m_l \to m_s} \frac{F_K}{F_\pi} = 1, \tag{3.33}$$

at any lattice spacing.

From a purist EFT perspective, we should instead utilize the MA EFT expression. Unfortunately, the MA EFT expression is only known at NLO [52] and our results require higher orders to provide good fits. Nevertheless, we can explore the utility of the MA EFT by replacing the NLO χPT expression with the NLO MA EFT expression and using the continuum expression enhanced with the local discretization corrections at higher orders, Eq. (3.32).

Using the parametrization of the hairpin contributions from Ref. [55], the NLO MA EFT expressions are

$$\begin{aligned} \delta F_\pi^{\mathrm{MA}} &= -\ell_{ju} - \frac{\ell_{ru}}{2} + 4\bar{L}_5 \epsilon_\pi^2 + 4\bar{L}_4(\epsilon_\pi^2 + 2\epsilon_K^2) + \epsilon_a^2 \bar{L}_a, \\ \delta F_K^{\mathrm{MA}} &= -\frac{\ell_{ju}}{2} + \frac{\ell_\pi}{8} - \frac{\ell_{ru}}{4} - \frac{\ell_{js}}{2} - \frac{\ell_{rs}}{4} + \frac{\ell_{ss}}{4} - \frac{3\ell_X}{8} \\ &\quad + 4\bar{L}_5 \epsilon_K^2 + 4\bar{L}_4(\epsilon_\pi^2 + 2\epsilon_K^2) + \epsilon_a^2 \bar{L}_a \\ &\quad - \frac{\delta_{ju}^2}{8}(d\ell_\pi - 2\mathcal{K}_{\pi X}) - \frac{\delta_{ju}^4}{24}\mathcal{K}_{\pi X}^{(2,1)} \\ &\quad + \frac{\delta_{rs}^2}{4}\left(\mathcal{K}_{ssX} - \frac{2}{3}(\epsilon_K^2 - \epsilon_\pi^2)\mathcal{K}_{ssX}^{(2,1)}\right) \\ &\quad + \frac{\delta_{ju}^2 \delta_{rs}^2}{12}(\mathcal{K}_{ssX}^{(2,1)} - 2\mathcal{K}_{\pi ssX}). \end{aligned} \tag{3.34}$$

In these expressions, we use the partially quenched flavor notation [74] in which

$$\begin{aligned} \pi &: \text{ valence-valence pion} \\ K &: \text{ valence-valence kaon} \\ u &: \text{ valence light flavor quark} \\ j &: \text{ sea light flavor quark} \\ s &: \text{ valence strange flavor quark} \\ r &: \text{ sea strange flavor quark} \\ X &: \text{ sea-sea eta meson} \end{aligned}, \tag{3.35}$$

and so, for example

$$\ell_{ju} = \frac{m_{ju}^2}{(4\pi F_\pi)^2} \ln\left(\frac{m_{ju}^2}{(4\pi F_\pi)^2}\right), \tag{3.36}$$

where m_{ju} is the mass of a mixed valence-sea pion. The partial quenching parameters δ_{ju^2} and δ_{rs}^2 provide a measure of the unitarity violation in the theory. For our MDWF on HISQ action, at LO in MA EFT, they are given by the

[3] One can use the renormalization-group to resum corrections from radiative gluons that modify the leading asymptotic scaling behavior [71,72]. For actions without dimension-5 operators in the Symanzik EFT, these resummed scaling violations are known to be proportional to

$$\delta_a^{\mathrm{Symanzik}} = c_2^{\mathrm{O}} a^2 \alpha_S^{n+\hat{\gamma}_1}, \tag{3.31}$$

where c_2^{O} is an LEC for operator O and whose value depends upon the lattice action. The power is $n = 0$ for unimproved actions (such as our MDWF valence action), $n = 1$ for tree-level improved actions (such as the HISQ action) and $n = 2$ for one-loop improved actions. The anomalous dimension $\hat{\gamma}_1$ can be determined in the asymptotic scaling regime which has been recently done for Yang-Mills and Wilson fermion actions with $\hat{\gamma}_1^{\mathrm{YM}} = 7/11$ [73]. This anomalous dimension is not known for our action. In principle, one could perform a fit where instead of treating the a^2 and $\alpha_S a^2$ terms with different LECs, one could combine them as in Eq. (3.31) and try and fit both c_2^{O} and $\hat{\gamma}_1$. We leave this to future studies and in this work, we use Eq. (3.32).

splitting in the quark masses plus a discretization correction arising from the taste-identity splitting

$$\delta^2_{ju} = \frac{2B_0(m_j - m_u) + a^2\Delta_{\mathrm{I}}}{(4\pi F_\pi)^2},$$
$$\delta^2_{rs} = \frac{2B_0(m_r - m_s) + a^2\Delta_{\mathrm{I}}}{(4\pi F_\pi)^2}. \tag{3.37}$$

For the tuning we have done, setting the valence-valence pion mass equal to the taste-5 sea-sea pion mass, these parameters are given just by the discretization terms as $m_u = m_j$ and $m_s = m_r$ within 1%–2%. The sea-sea eta mass in this tuning is given at LO in MA EFT as

$$m^2_X = \frac{4}{3}m^2_K - \frac{1}{3}m^2_\pi + a^2\Delta_{\mathrm{I}}. \tag{3.38}$$

These parameters, and the corresponding meson masses are provided in Table III. The expressions for $d\ell_\pi$, $\mathcal{K}_{\phi_1\phi_2}$, $\mathcal{K}^{(2,1)}_{\phi_1\phi_2}$ and $\mathcal{K}_{\phi_1\phi_2\phi_3}$ are provided in the Appendix B.

At NLO in the MA EFT, the LECs which contribute to δF_K and δF_π are the same as in the continuum, L_4 and L_5, plus a discretization LEC which we have denoted $\bar{L}_a$. Just like the L_4 contribution, the contribution from $\bar{L}_a$ exactly cancels in $\delta F_K - \delta F_\pi$. At N^{2}LO, beyond the continuum counterterm contributions, Eq. (3.12), there are the two additional LECs contributions, Eq. (3.32).

C. Finite volume corrections

We now discuss the corrections arising from the finite spatial volume. The leading finite volume (FV) corrections arise from the tadpole integrals which arise at NLO in both the χPT and MA expressions. The well-known modification to the integral can be expressed as [75–77]

$$\ell_P^{\mu_\pi,\mathrm{FV}} = \ell_P^{\mu_\pi} + 4\epsilon_P^2 \sum_{|\mathbf{n}|\neq 0} \frac{c_n}{m_P L|\mathbf{n}|} K_1(m_P L|\mathbf{n}|), \tag{3.39}$$

where the sum runs over all nonzero integer three-vectors. Each value of $|\mathbf{n}|$ can be thought of as a winding of the meson P around the finite universe. The c_n are multiplicity factors counting all the ways to form a vector of length $|\mathbf{n}|$ from triplets of integers, see Table IV for the first few. $K_1(x)$ is a modified Bessel function of the second kind. In the asymptotically large volume limit, the finite volume correction to these integrals is

$$\begin{aligned}\delta^{\mathrm{FV}}\ell_P &\equiv \ell_P^{\mathrm{FV}} - \ell_P \\ &= \epsilon_P^2 2\sqrt{2\pi}\frac{e^{-m_P L}}{(m_P L)^{3/2}} \\ &\quad + \epsilon_P^2 \times \mathrm{O}\left(\frac{e^{-m_P L\sqrt{2}}}{(m_P L\sqrt{2})^{3/2}}, \frac{e^{-m_P L}}{(m_P L)^{5/2}}\right).\end{aligned} \tag{3.40}$$

TABLE III. Extracted masses of the mixed MDWF-HISQ mesons. We use the notation from Ref. [74] in which m_π and m_K denote the masses of the valence pion and kaon and j and r denote the light and strange flavors of the sea quarks while u and s denote the light and strange flavors of the valence quarks. Since we have tuned the valence MDWF pion and $\bar{s}s$ mesons to have the same mass as the HISQ sea pion and $\bar{s}s$ mesons within a few percent, the quantities $m^2_{ju} - m^2_\pi$ and other splittings provide an estimate of the additive mixed-meson mass splitting due to discretization effects, $a^2\Delta_{\mathrm{Mix}}$ [52] and additional additive corrections [59]. At LO in MA EFT, these splittings are predicted to be quark mass independent, which we find to be approximately true, with a notable decrease in the splitting as the valence-quark mass is increased as first observed in Ref. [56] as well as a milder decrease as the seq-quark mass is increased.

Ensemble	am_{ju}	am_{js}	am_{ru}	am_{rs}	am_{ss}	$w_0^2\Delta^2_{\mathrm{Mix,ju}}$	$w_0^2\Delta^2_{\mathrm{Mix,js}}$	$w_0^2\Delta^2_{\mathrm{Mix,ru}}$	$w_0^2\Delta^2_{\mathrm{Mix,rs}}$	$w_0^2a^2\Delta_{\mathrm{I}}$
a15m400	0.3597(17)	0.4586(24)	0.4717(19)	0.5537(11)	0.5219(02)	0.0486(15)	0.0359(28)	0.0516(23)	0.0440(16)	0.112(14)
a15m350	0.3308(23)	0.4463(14)	0.4598(16)	0.5526(10)	0.5201(02)	0.0508(20)	0.0362(17)	0.0519(19)	0.0451(15)	0.112(14)
a15m310	0.3060(17)	0.4345(16)	0.4508(14)	0.5490(12)	0.5188(02)	0.0489(13)	0.0324(18)	0.0511(17)	0.0416(16)	0.112(14)
a15m220	0.2564(27)	0.4115(17)	0.4320(29)	0.5420(08)	0.5150(01)	0.0495(18)	0.0253(19)	0.0476(33)	0.0368(11)	0.112(14)
a15m135XL	0.232(13)	0.4058(56)	0.4337(84)	0.5560(31)	0.5257(02)	0.0559(75)	0.0187(59)	0.0489(94)	0.0423(45)	0.112(14)
a12m400	0.2678(06)	0.3560(08)	0.3624(07)	0.4333(06)	0.4207(01)	0.0251(07)	0.0177(12)	0.0271(10)	0.0217(11)	0.063(05)
a12m350	0.2303(08)	0.3446(07)	0.3454(10)	0.4322(05)	0.4197(01)	0.0147(07)	0.0158(10)	0.0168(15)	0.0214(09)	0.063(05)
a12m310	0.2189(09)	0.3344(10)	0.3439(09)	0.4305(05)	0.4180(02)	0.0248(08)	0.0136(14)	0.0266(13)	0.0213(09)	0.063(05)
a12m220S	0.1774(14)	0.3187(12)	0.3323(17)	0.4286(10)	0.4158(02)	0.0264(10)	0.0105(16)	0.0283(24)	0.0219(18)	0.063(05)
a12m220L	0.1774(14)	0.3187(12)	0.3323(17)	0.4286(10)	0.4156(02)	0.0273(10)	0.0107(16)	0.0286(23)	0.0222(18)	0.063(05)
a12m220	0.1774(14)	0.3187(12)	0.3323(17)	0.4286(10)	0.4154(01)	0.0272(10)	0.0110(16)	0.0289(23)	0.0225(18)	0.063(05)
a12m130	0.1491(20)	0.3080(15)	0.3240(26)	0.4271(08)	0.4141(01)	0.0316(12)	0.0073(19)	0.0276(34)	0.0220(14)	0.063(05)
a09m400	0.1878(05)	0.2581(06)	0.2607(06)	0.3162(05)	0.3133(01)	0.0094(07)	0.0056(12)	0.0109(11)	0.0071(12)	0.020(02)
a09m350	0.1654(06)	0.2498(05)	0.2526(06)	0.3159(04)	0.3124(01)	0.0093(07)	0.0054(10)	0.0108(12)	0.0083(11)	0.020(02)
a09m310	0.1485(06)	0.2428(05)	0.2472(10)	0.3150(04)	0.3117(01)	0.0086(07)	0.0032(10)	0.0114(20)	0.0080(09)	0.020(02)
a09m220	0.1090(09)	0.2303(06)	0.2334(07)	0.3115(03)	0.3094(01)	0.0088(07)	0.0028(10)	0.0083(12)	0.0051(08)	0.020(02)
a09m135	0.0786(15)	0.2187(11)	0.2270(15)	0.3079(05)	0.3027(07)	0.0102(09)	0.0004(19)	0.0146(26)	0.0123(19)	0.020(02)
a06m310L	0.0957(08)	0.1619(11)	0.1619(12)	0.2103(10)	0.2098(01)	0.0020(14)	−0.0004(34)	−0.0004(34)	0.0020(40)	0.004(00)

TABLE IV. Multiplicity factors for the finite volume corrections of the first 10 vector lengths, $|\mathbf{n}|$.

$\lvert\mathbf{n}\rvert$	1	$\sqrt{2}$	$\sqrt{3}$	$\sqrt{4}$	$\sqrt{5}$	$\sqrt{6}$	$\sqrt{7}$	$\sqrt{8}$	$\sqrt{9}$	$\sqrt{10}$
c_n	6	12	8	6	24	24	0	12	30	24

The full finite volume corrections to the continuum formula are also known at N^{2}LO [78] as well as in the partially quenched χPT [79]. In this work, we restrict the corrections to those arising from the NLO corrections as our results are not sensitive to higher-order FV corrections. This is because, with the ensembles used in this work, all ensembles except a12m220S satisfy $m_\pi L \gtrsim 4$ (see Table II). MILC generated three volumes for this a12m220 ensemble series to study FV corrections. Figure 2 shows a comparison of the results from the a12m220L, a12m220, and a12m220S along with the predicted volume corrections arising from NLO in χPT. The uncertainty band arises from an N^{3}LO fit using the full N^{2}LO continuum χPT formula enhanced with discretization LECs and N^{3}LO corrections arising from continuum and finite lattice spacing corrections. Even with one of the most precise fits, we see that the numerical results are consistent with the predicted NLO FV corrections.

D. N^{3}LO corrections

The numerical dataset in this work requires us to add N^{3}LO corrections to obtain a good fit quality. At this order, we only consider local counterterm contributions, of which there are three new continuumlike corrections and three discretization corrections. A nonunique, but complete parametrization is

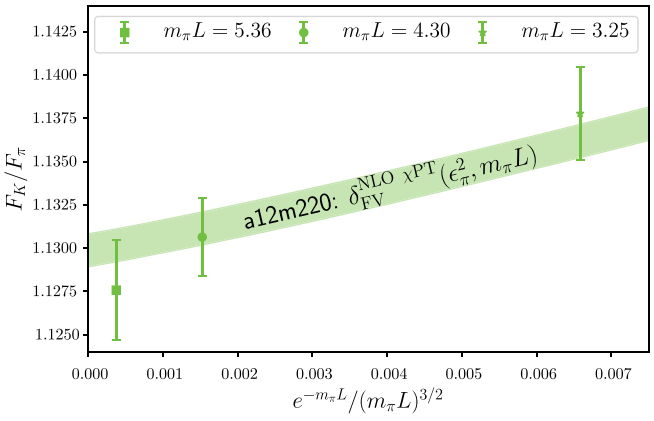

FIG. 2. We compare the finite volume results on a12m220L, a12m220 and a12m220S to the predicted finite volume corrections from NLO χPT. The uncertainty band is from the full N^{3}LO χPT extrapolation, plotted with fixed mesons masses (ϵ_P^2) and fixed lattice spacing (ϵ_a^2), determined from the a12m220L ensemble. At the one-sigma level, our data are consistent with the leading FV corrections.

$$\delta^{\mathrm{N^3LO}} = (\epsilon_K^2 - \epsilon_\pi^2)\{\epsilon_a^4 A_s^6 + \epsilon_a^2(A_{s,K}^6\epsilon_K^2 + A_{s,\pi}^6\epsilon_\pi^2) + A_{K\pi}^6\epsilon_K^2\epsilon_\pi^2 + (\epsilon_K^2 - \epsilon_\pi^2)(A_K^6\epsilon_K^2 + A_\pi^6\epsilon_\pi^2)\}. \quad (3.41)$$

In principle, we could also add counterterms proportional to higher powers of α_S but with four lattice spacings, we would not be able to resolve the difference between the complete set of operators including all possible additional α_S corrections. The set of operators we do include is sufficient to parametrize the approach to the continuum limit.

IV. EXTRAPOLATION DETAILS AND UNCERTAINTY ANALYSIS

We now carry out the extrapolation/interpolation to the physical point, which we perform in a Bayesian framework. To obtain a good fit, we must work to N^{3}LO in the mixed chiral and continuum expansion. The results from the a06m310L ensemble drive this need, in particular, for higher-order discretization corrections to parametrize the results from all the ensembles. We will explore the impact of the a06m310L ensemble in more detail in this section. First, we discuss the values of the priors we set and the definition of the physical point.

A. Prior widths for LECs

The number of additional LECs we need to determine at each order in the expansion is

order	N_{L_i}	N_χ	N_a
NLO	1	0	0
N^{2}LO	7	2	2
N^{3}LO	0	3	3
Total	8	5	5

N_{L_i} is the number of Gasser-Leutwyler coefficients, N_χ the number of counterterms associated with the continuum χPT expansion and N_a is the number of counterterms associated with the discretization corrections. In total, there are 18 unknown LECs. While we utilize 18 ensembles in this analysis, the span of parameter space is not sufficient to constrain all the LECs without prior knowledge. In particular, the introduction of all 8 L_i coefficients requires prior widths informed from phenomenology.

In the literature, the L_i are typically quoted at the renormalization scale $\mu_\rho = 770$ MeV while in our work, we use the scale $\mu_{F_0} = 4\pi F_0$. We can use the BE14 values of the L_i LECs from Ref. [80] and the known scale dependence [68] to convert them from μ_ρ to μ_{F_0}:

$$L_i^r(\mu_2) = L_i^r(\mu_1) - \frac{\Gamma_i}{(4\pi)^2}\ln\left(\frac{\mu_2}{\mu_1}\right), \quad (4.1)$$

with the values of Γ_i listed in Table V for convenience. We use $F_0 = 80$ MeV, which is the value adopted by

TABLE V. Γ_i coefficients that appear in the scale dependence of the $L_i(\mu)$. We evolve the $L_i(\mu)$ from the typical scale $\mu = 770$ MeV, Eq. (3.21) to $\mu_0 = 4\pi F_0$, beginning with the BE14 estimates from the review [80] (Table 3), using their known scale dependence [68], Eq. (4.1). We assign the following slightly more conservative uncertainty as a prior width in the minimization: If a value of L_i is less than 0.5×10^{-3}, we assign it a 100% uncertainty at the scale $\mu = 770$ MeV; if the value is larger than 0.5×10^{-3}, we assign it the larger of 0.5 or 1/3 of the mean value.

L_i	L_1	L_2	L_3	L_4	L_5	L_6	L_7	L_8
Γ_i	3/32	3/16	0	1/8	3/8	11/144	0	5/48
$10^3 L_i(m_\rho)$	0.53(50)	0.81(50)	−3.1(1.0)	0.30(30)	1.01(50)	0.14(14)	−0.34(34)	0.47(47)
$10^3 L_i(\mu_0)$	0.37(50)	0.49(50)	−3.1(1.0)	0.09(30)	0.38(50)	0.01(14)	−0.34(34)	0.29(47)

FLAG [8]. We set the central value of all the L_i with this procedure and the widths are set as described in Table V.

Next, we must determine priors for the N^{2}LO and N^{3}LO local counterterm coefficients, $A^n_{K,\pi,s}$. We set the central value of all these priors to 0 and then perform a simple grid search varying the widths to find preferred values of the width, as measured by the Bayes factor. Our goal is not to optimize the width of each prior individually for each model used in the fit, but rather find a set of prior widths that is close to optimal for all models. To this end, we vary the width of the χPT LECs together at each order (N^{2}LO, N^{3}LO) and the discretization LECs together at each order (N^{2}LO, N^{3}LO) for a four-parameter search. We apply a very crude grid where the values of the widths are taken to be 2, 5, or 10.

We find taking the width of all these $A^n_{K,\pi,s}$ LECs equal to 2 results in good fits with near-optimal values. This provides evidence the normalization of small parameters we have chosen for ϵ_P^2 and ϵ_a^2, Eq. (3.1), is "natural" and supports the power-counting we have assumed, Eq. (3.2). The N^{2}LO LECs mostly favor a width of 2 while the N^{3}LO discretization LECs prefer 5 and the N^{3}LO χPT LECs vary from model to model with 5 a reasonable value for all. As a result of this search, we pick as our priors

$$\tilde{A}^4_{K,\pi} = 0 \pm 2, \qquad \tilde{A}^4_{s} = 0 \pm 2,$$
$$\tilde{A}^6_{K,\pi} = 0 \pm 5, \qquad \tilde{A}^6_{s} = 0 \pm 5. \tag{4.2}$$

B. Physical point

As our calculation is performed with isospin symmetric configurations and valence quarks, we must define a physical point to quote our final result. We adopt the definition of the physical point from FLAG. FLAG[2017] [81] defines the isospin symmetric pion and kaon masses to be [Eq. (16)]

$$\bar{M}_\pi = 134.8(3)\ \text{MeV},$$
$$\bar{M}_K = 494.2(3)\ \text{MeV}. \tag{4.3}$$

The values of F_{π^+} and F_{K^+} are taken from the $N_f = 2+1$ results from FLAG[2020] [8] (we divide the values by $\sqrt{2}$ to convert to the normalization used in this work)

$$F^{\text{phys}}_{\pi^+} = 92.07(57)\ \text{MeV},$$
$$F^{\text{phys}}_{K^+} = 110.10(49)\ \text{MeV}. \tag{4.4}$$

The isospin symmetric physical point is then defined by extrapolating our results to the values (for the choice $F \to F_\pi$)

$$(\epsilon_\pi^{\text{phys}})^2 = \left(\frac{\bar{M}_\pi}{4\pi F^{\text{phys}}_{\pi^+}}\right)^2,$$
$$(\epsilon_K^{\text{phys}})^2 = \left(\frac{\bar{M}_K}{4\pi F^{\text{phys}}_{\pi^+}}\right)^2. \tag{4.5}$$

C. Model averaging procedure

Our model average is performed under a Bayesian framework following the procedure described in [36,82]. Suppose we are interested in estimating the posterior distribution of $Y = F_K/F_\pi$, i.e., $P(Y|D)$ given our data D. To that end, we must marginalize over the different models M_k.

$$P(Y|D) = \sum_k P(Y|M_k, D) P(M_k|D). \tag{4.6}$$

Here $P(Y|M_k, D)$ is the distribution of Y for a given model M_k and dataset D, while $P(M_k|D)$ is the posterior distribution of M_k given D. The latter can be written, per Bayes's theorem, as

$$P(M_k|D) = \frac{P(D|M_k)P(M_k)}{\sum_l P(D|M_l)P(M_l)}. \tag{4.7}$$

We can be more explicit with what the latter is in the context of our fits. First, mind that we are *a priori* agnostic in our choice of M_k. We thus take the distribution $P(M_k)$ to be uniform over the different models. We calculate $P(D|M_k)$ by marginalizing over the parameters (LECs) in our fits:

$$P(D|M_k) = \int \prod_j \mathrm{d}\theta_j^{(k)} P(D|\theta_j^{(k)}, M_k) P(\theta_j^{(k)}|M_k). \tag{4.8}$$

After marginalization, $P(D|M_k)$ is just a number. Specifically, it is the Bayes factor of M_k: $P(D|M_k) = \exp(\texttt{log GBF})_{M_k}$, where logGBF is the log of the Bayes factor as reported by LSQFIT [83]. Thus

$$P(M_k|D) = \frac{\exp(\texttt{log GBF})_{M_k}}{\sum_{l=1}^{K}\exp(\texttt{log GBF})_{M_l}} \tag{4.9}$$

with K the number of models included in our average. We emphasize that this model selection criterion not only rates the quality of the description of data but also penalizes parameters which do not improve this description. This helps rule out models which overparametrize data.

Now we can estimate the expectation value and variance of Y:

$$\mathrm{E}[Y] = \sum_k \mathrm{E}[Y|M_k]P(M_k|D), \tag{4.10}$$

$$\begin{aligned}\mathrm{Var}[Y] = &\left[\sum_k \mathrm{Var}[Y|M_k]P(M_k|D)\right] \\ &+ \left[\left(\sum_k \mathrm{E}^2[Y|M_k]P(M_k|D)\right) - \mathrm{E}^2[Y|D]\right].\end{aligned} \tag{4.11}$$

The variance $\mathrm{Var}[Y]$ results from the total law of variance; the first term in brackets is known as the *expected value of the process variance* (which we refer to as the *model averaged variance*), while the latter is the *variance of the hypothetical means* (the root of which we refer to as the *model uncertainty*). After this work was completed, a similar, but more thorough discussion of Bayesian Model Averaging in the context of lattice QCD was presented [84].

D. Full analysis and uncertainty breakdown

In total, we consider 216 different models of extrapolation/interpolation to the physical point. The various choices for building a χPT or MA EFT model consist of

×2:	χPT or MA EFT at NLO
×3:	use $F^2 = \{F_\pi^2, F_\pi F_K, F_K^2\}$ in defining ϵ_P^2
×2:	fully expanded (3.15) or ratio (3.17) form
×2:	at N^{2}LO, use full χPT or just counterterms
×2:	include or not an α_S term at N^{2}LO
×2:	include or not the NLO FV corrections
×2:	include N^{3}LO counterterms or not
192:	total choices

We also consider pure Taylor expansion fits with only counterterms and no log corrections. For these fits, the set of models we explore is

×2:	work to N^{2}LO or N^{3}LO
×3:	use $F^2 = \{F_\pi^2, F_\pi F_K, F_K^2\}$ in defining ϵ_P^2
×2:	include or not an α_S term at N^{2}LO
×2:	include or not FV corrections
24:	total choices

Based upon the quality of fit (gauged by the Bayesian analog to the p-value, Q, or the reduced chi square, χ^2_ν) and/or the weight determined as discussed in the previous section, we can dramatically reduce the number of models used in the final averaging procedure. First, any model which does not include the FV correction from NLO is heavily penalized. This is not surprising given the observed volume dependence on the a12m220 ensembles, Fig. 2. However, even if we remove the a12m220S ensemble from the analysis, the Taylor-expanded fits have a relative weight of e^{-6} or less compared to those that have χPT form at NLO.

If we add FV corrections to the Taylor expansion fits (pure counterterm) and use all ensembles,

$$\frac{F_K}{F_\pi} = 1 + \bar{L}_5(\epsilon_K^2 - \epsilon_\pi^2)\left\{1 + t_{\mathrm{FV}}\sum_{|\mathbf{n}|\neq 0}\frac{c_n}{m_\pi L|\mathbf{n}|}K_1(m_\pi L|\mathbf{n}|)\right\} + \cdots \tag{4.12}$$

they still have weights which are $\sim e^{-8}$ over the normalized model distribution and also contribute negligibly to the model average.

We observe that the fits which use the MA EFT at NLO are also penalized with a relative weight of $\sim e^{-8}$, and fits which only work to N^{2}LO have unfavorable weights by $\sim e^{-5}$ (and are also accompanied by poor χ^2_ν values). Cutting all of these variations reduces our final set of models to be N^{3}LO χPT with the following variations:

×3:	use $F^2 = \{F_\pi^2, F_\pi F_K, F_K^2\}$ in defining ϵ_P^2
×2:	fully expanded (3.15) or ratio (3.15) form
×2:	at N^{2}LO, use full χPT or just counterterms
×2:	include or not an α_S term at N^{2}LO
24:	total choices which enter the model average

The final list of models, with their corresponding weights and resulting extrapolated values to the isospin symmetric physical point, is given in Table VII in Appendix A. Our final result in the isospin symmetric limit, defined as in Eq. (4.5) and analogously for other choices of F^2, including a breakdown in terms of statistical (s), pion mass extrapolation (χ), continuum limit (a), infinite volume limit (V), physical point (phys) and model selection (M) uncertainties, is as reported in Eq. (1.3)

$$\frac{F_K}{F_\pi} = 1.1964(32)^s(12)^\chi(20)^a(01)^V(15)^{\text{phys}}(12)^M = 1.1964(44).$$

The finite volume uncertainty is assessed by removing the a12m220S ensemble from the analysis, repeating the model averaging procedure and taking the difference. The final probability distribution broken down into the three choices of F^2 is shown in Fig. 3.

1. Impact of a06m310L ensemble

Next, we turn to understanding the impact of the a06m310L ensemble on our analysis. The biggest difference upon removing the a06m310L ensemble is that the data are not able to constrain the various terms contributing to the continuum extrapolation as well, particularly since there are up to three different types of scaling violations:

$$(\epsilon_K^2 - \epsilon_\pi^2) \times \{\epsilon_a^2, \alpha_S\epsilon_a^2, \epsilon_a^4\},$$

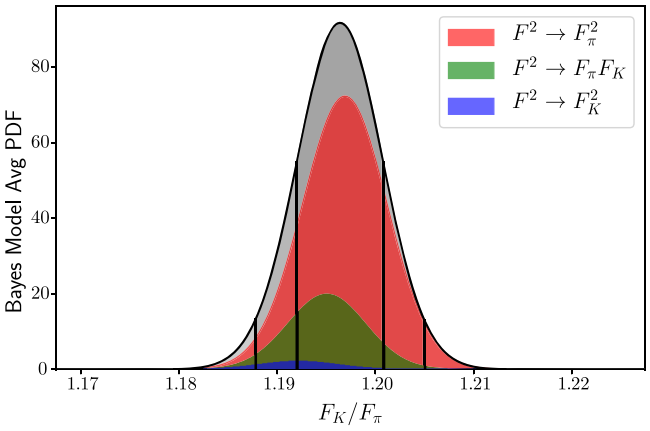

FIG. 3. Final probability distribution giving rise to Eq. (1.3), separated into the three choices of $F^2 = \{F_\pi^2, F_\pi F_K, F_K^2\}$ in the definition of the small parameters, Eq. (3.1). The parent "gray" distribution is the final PDF normalized to 1 when integrated.

and thus, the statistical uncertainty of the results grows as well as the model variance, with a total uncertainty growth from ∼0.0044 to ∼0.0057, and the mean of the extrapolated answer moves by approximately half a standard deviation. Furthermore, N^{2}LO fits become acceptable, though they are still grossly outweighed by the N^{3}LO fits. Including both effects, the final model average result shifts from

$$\frac{F_K}{F_\pi} = 1.1964(44) \to \left.\frac{F_K}{F_\pi}\right|_{\text{no a06}} = 1.1941(57). \quad (4.13)$$

In Fig. 4, we show the continuum extrapolation from three fits:

(i) Left: all ensembles, N^{3}LO χPT with only counterterms at N^{2}LO and N^{3}LO and $F = F_\pi$;
(ii) Middle: no a06m310L, N^{3}LO χPT with only counterterms at N^{2}LO and N^{3}LO and $F = F_\pi$;
(iii) Right: no a06m310L, N^{2}LO χPT with only counterterms at N^{2}LO and $F = F_\pi$.

As can be seen from the middle plot, the a15, a12 and a09 ensembles prefer contributions from both ϵ_a^2 and ϵ_a^4 contributions and are perfectly consistent with the result on the a06m310L ensemble. They are also consistent with an N^{2}LO fit (no ϵ_a^4 contributions) as can be seen in the right figure. However, the weight of the N^{3}LO fits is still significantly greater than the N^{2}LO fits even without the a06m310L data.

We conclude that the a06m310L ensemble is useful, but not necessary to obtain a subpercent determination of F_K/F_π with our lattice action. A more exhaustive comparison can be performed with the analysis notebook provided with this publication.

In Fig. 5, we show the stability of our final result for various choices discussed in this section.

2. Convergence of the chiral expansion

While the numerical analysis favors a fit function in which only counterterms are used at N^{2}LO and higher, it is interesting to study the convergence of the chiral expansion

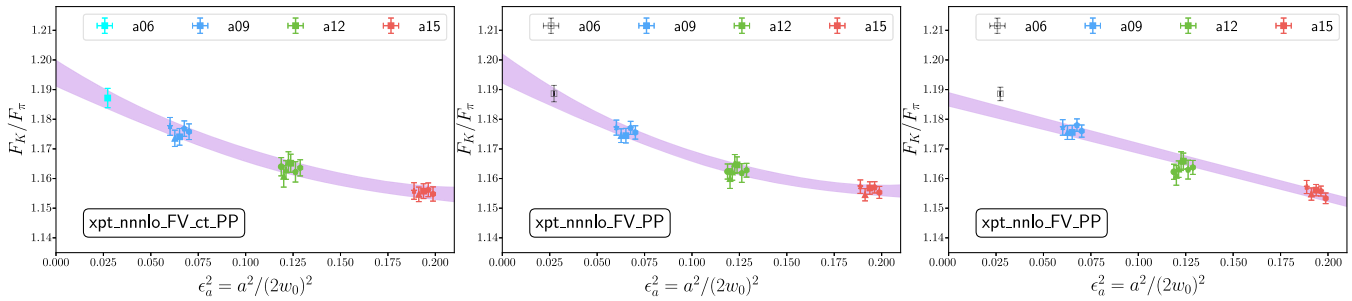

FIG. 4. Left: N^{3}LO fit to all ensembles. Middle: same fit to all ensembles excluding a06m310L. Right: representative N^{2}LO fit to all ensembles excluding a06m310L. In all plots, the results from each ensemble are shifted to the physical values of ϵ_π^2 and ϵ_K^2 and the infinite volume limit, with only the ϵ_a^2 dependence remaining. The labels are explained in Appendix A; the data points at each spacing are slightly offset horizontally for visual clarity.

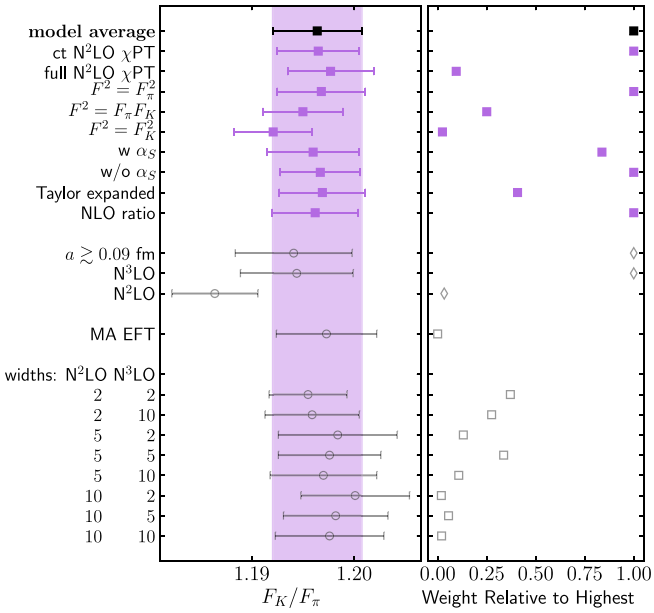

FIG. 5. Stability plot of final result compared to various model choices. The black square at the top is our final answer for the isospin symmetric determination of F_K/F_π. The vertical magenta band is the uncertainty of this fit to guide the eye. The solid magenta squares are various ways of decomposing the model selection that goes into the final average. The right panel shows the relative weight with respect to the maximum logGBF value, $\exp\{\log \mathrm{GBF}_i - \log \mathrm{GBF}_{\max}\}$. Below the set of models included in the average, we show sets of analyses that are not included in the average for comparison, which are indicated with open gray symbols. First, we show the impact of excluding the a06m310L ensemble. The logGBF cannot be directly compared between these fits and the main analysis as the number of data points are not the same, so the overall normalization is different; their logGBF are shown as open diamonds. The relative logGBF between the N³LO and N²LO analysis can be compared which indicates a large preference for the N³LO analysis. We also show the MA EFT analysis, which agrees well with the main analysis. Finally, we show the results if one were to change the widths of the N²LO and N³LO priors from those chosen in Eq. (4.2).

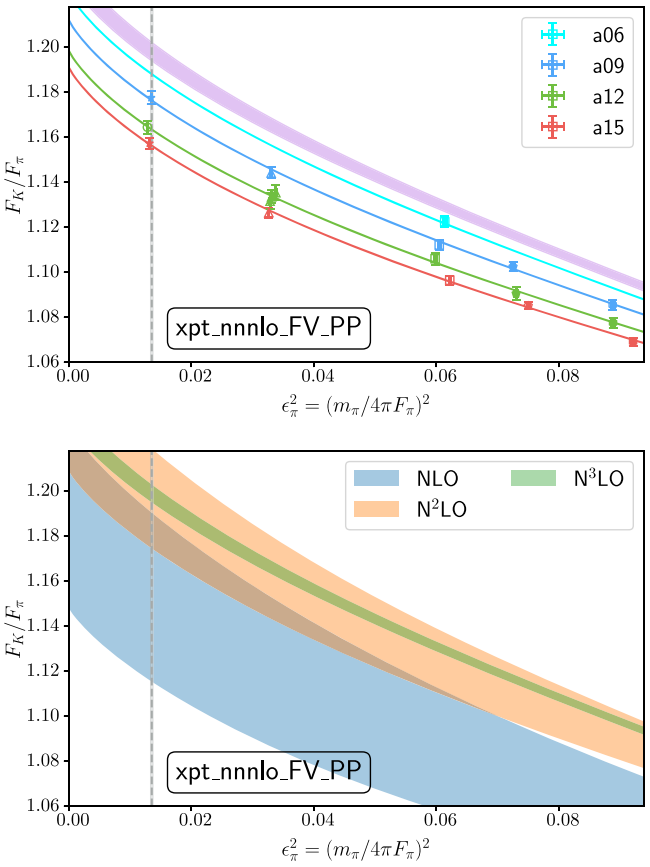

FIG. 6. Sample light quark mass dependence from a χPT fit with $F = F_\pi$ (N²LO χPT + N³LO counterterms). Top: curves are plotted with $\epsilon_K^2 = (\epsilon_K^{\mathrm{phys}})^2$, at fixed ϵ_a^2 for each lattice spacing, as a function of ϵ_π^2. The magenta band is the full uncertainty in the continuum, infinite volume limit. The data points have all been shifted from the values of $\epsilon_K^{\mathrm{latt}}$. to $\epsilon_K^{\mathrm{phys}}$ and to the infinite volume limit. Bottom: We show the convergence of the resulting fit as a function of ϵ_π^2. Each band corresponds to all contributions up to that order with the LECs determined from the full fit. The N³LO band corresponds to the continuum extrapolated band in the top figure.

by studying the fits which use the full χPT expression at N²LO.

In Fig. 6, we show the resulting light quark mass dependence using the N³LO extrapolation with the full N²LO χPT formula. After the analysis is performed, the results from each ensemble are shifted to the physical kaon mass point, leaving only dependence upon ϵ_π^2 and ϵ_a^2 as well as dependence upon the η mass defined by the GMO relation. The magenta band represents the full 68% confidence interval in the continuum, infinite volume limit. The different colored curves are the mean values as a function of ϵ_π^2 at the four different lattice spacings. We also show the convergence of this fit in the lower panel plot. From this convergence plot, one sees that roughly that at the physical pion mass (vertical gray line) the NLO contributions add a correction of ~0.16 compared to 1 at LO, the N²LO contributions add another ~0.04, and the N³LO corrections are not detectable by eye. The band at each order represents the sum of all terms up to that order determined from the full fit. The reduction in the uncertainty as the order is increased is due on large part to the induced correlation between the LECs at different orders through the fitting procedure.

In Fig. 3, we observe that the different choices of F are all consistent, indicating higher-order corrections (starting at N³LO in the noncounterterm contributions) are smaller than the uncertainty in our results. It is also interesting to note that choosing $F_{\pi K}$ or F_K is penalized by the analysis, indicating the numerical results prefer larger expansion parameters. In Table VI, we show the resulting χPT LECs determined in this analysis for the two choices $F = \{F_\pi, F_{K\pi}\}$, as well as whether the ratio form of the

TABLE VI. Resulting LECs from full N^{2}LO χPT analysis (also including N^{3}LO counterterms). For the Gasser-Leutwyler LECs L_i, we evolve them back to the standard scale $\mu = 770$ MeV, while for the other LECs, we leave them at the scale $\mu_0 = 4\pi F_0 \simeq 1005$ MeV.

		$F^2 = F_\pi^2$		$F^2 = F_\pi F_K$	
		Ratio		Ratio	
LEC $\mu = 770$	Prior	No	Yes	No	Yes
$10^3 L_1$	0.53(50)	0.47(49)	0.50(49)	0.45(49)	0.48(49)
$10^3 L_2$	0.81(50)	0.77(46)	0.84(46)	0.69(44)	0.77(45)
$10^3 L_3$	−3.1(1.0)	−3.02(85)	−2.84(86)	−3.26(81)	−3.05(82)
$10^3 L_4$	0.30(30)	0.24(29)	0.14(29)	0.24(29)	0.16(29)
$10^3 L_5$	1.01(50)	0.48(35)	0.52(34)	0.40(33)	0.47(34)
$10^3 L_6$	0.14(14)	0.14(14)	0.14(14)	0.14(14)	0.14(14)
$10^3 L_7$	−0.34(34)	−0.55(32)	−0.57(32)	−0.52(33)	−0.53(33)
$10^3 L_8$	0.47(47)	0.30(46)	0.28(46)	0.35(46)	0.32(46)
$\mu = \mu_0$					
A_K^4	0(2)	0.06(1.42)	0.09(1.41)	0.2(1.6)	0.2(1.5)
A_π^4	0(2)	2.5(1.2)	2.4(1.2)	2.0(1.3)	2.0(1.3)
$A_{K\pi}^6$	0(5)	2.8(4.7)	2.8(4.7)	1.9(4.7)	2.0(4.7)
A_K^6	0(5)	0.008(4.016)	0.3(4.0)	0.1(4.4)	0.2(4.4)
A_p^6	0(5)	2.6(4.0)	2.1(4.1)	2.4(4.4)	2.0(4.4)

fit is used, Eq. (3.17). For the Gasser-Leutwyler LECs, we evolve the values back from $\mu_0 \to \mu_\rho$ for a simpler comparison with the values quoted in literature. For most of the L_i, we observe the numerical results have very little influence on the parameters as they mostly return the prior value (also listed in the table for convenience). The only LECs influenced by the fit are L_5, L_7, and L_8 with L_5 getting pulled about one sigma away from the prior value and L_7 and L_8 only shifting by a third or half of the prior width. One interesting observation from our results is that our fit prefers a value of L_5 that is noticeably smaller than the value obtained by MILC [16] and HPQCD [12] and is also smaller than the BE14 result from Ref. [80], although the discrepancy is still less than 2 sigma. We also note that our value of L_5 is very compatible with that determined by RBC/UKQCD with domain-wall fermions and near-physical pion masses [18]. Those interested in exploring this in more detail can utilize our numerical results, and if desired, extrapolation code made available with this publication.

In Fig. 7, we show the impact of using the fully expanded expression, Eq. (3.15), versus the expression in which the NLO terms are kept in a ratio, Eq. (3.17). To simplify the comparison we restrict it to the choice $F = F_\pi$ and the full N^{2}LO χPT expression. We see that fits without the ratio form are preferred, but the central value of the final result depends minimally upon this choice.

In Fig. 8, we show that the results strongly favor the use of only counterterms at N^{2}LO as opposed to the full χPT fit function at that order. We focus on the choice $F = F_\pi$ to simplify the comparison.

Our results are not sufficient to understand why the fit favors only counterterms at N^{2}LO and higher. While the linear combination of LECs in Eq. (3.12) are redundant, the L_i LECs also appear in the single-log coefficients, Eqs. (3.10) and (3.11) in different linear combinations. Nevertheless, we double check that the fit is not penalized for the counterterm redundancy, Eq. (3.12). Using the priors for L_i from Table V, we find the contribution from the Gasser-Leutwyler LECs to these N^{2}LO counterterms, Eq. (3.14), are given by

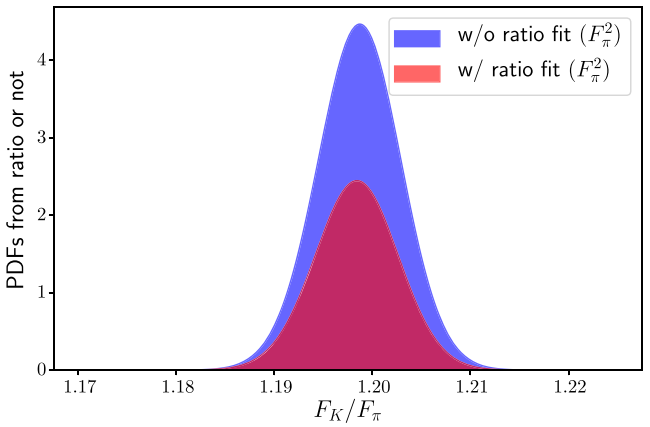

FIG. 7. Comparison of fits with the fully expanded Eq. (3.15) and ratio Eq. (3.17) expressions, all with the choice $F = F_\pi$. The PDFs are taken from the parent PDF, Fig. 3, without renormalizing such that height in this figure reflects the relative weight compared to the total PDF.

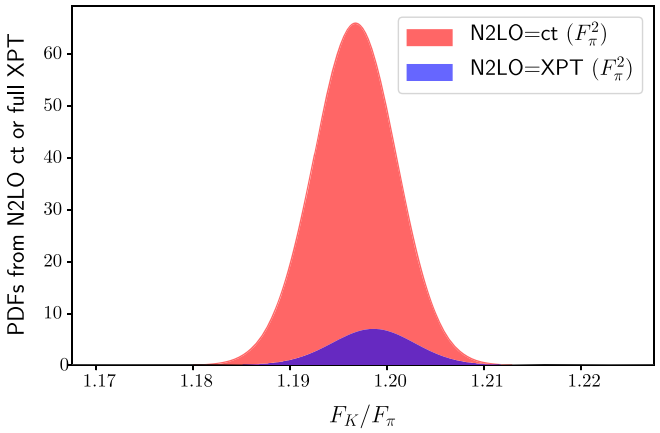

FIG. 8. Comparison of N^{3}LO χPT analysis with $F = F_\pi$ using the full N^{2}LO χPT expression (smaller histogram) versus only counterterms at N^{2}LO, Eqs. (3.12) and (3.32). As in Fig. 7, the PDFs are drawn from the parent PDF.

$$L_K^4 = 0.3(1.3), \qquad L_\pi^4 = -0.64(94). \tag{4.14}$$

As the A_P^4 terms are priored at 0(2), it is sufficient to rerun the analysis by simply setting $L_P^4 = 0$. We find this result marginally improves the Bayes factors but not statistically significantly, leaving us with the puzzle that the optimal fit is a hybrid NLO χPT plus counterterms (analytic terms) at higher orders. We note that it has been known for some time that using χPT at NLO plus purely analytic terms at NNLO and higher results in good quality extrapolation fits, at least in part because the NNLO chiral logarithms are relatively slowly varying for the range of pion masses for which the NNLO analytic terms are sizable enough to be important [2]. This is discussed in more detail in the review by Bernard [85]. The MILC Collaboration no longer reports analysis with just the analytic terms at NNLO [16] and so it is not clear if other groups observe the same preference for counterterms only at NNLO or not.

If the Taylor expansion fits (pure counterterm) were good and favored over the χPT fits, this could be a sign that the $SU(3)\chi$PT formula was failing to describe the lattice results. However, we have to include the NLO χPT expression, including its predicted (counterterm free) volume dependence to describe the numerical results. It would be nice to have the full N^{2}LO MA EFT expression to understand why the hybrid MA EFT fits are so relatively disfavored in the analysis. There may be compensating discretization effects that cancel against those at NLO to some degree that might allow the full N^{2}LO MA EFT to better describe the results. However, at two loops in χPT, the universality of MA EFT expressions [58] breaks down such that the MA EFT expression can no longer be "derived" from the corresponding PQχPT one (which is known for F_K/F_π at two loops [79,86–88]). It is therefore unlikely that the NNLO MA EFT expression specific to this MALQCD calculation will ever be derived, so this issue will most likely not be resolved with more clarity.

E. QCD isospin breaking corrections

Finally, we discuss the correction to our result to obtain a direct determination of F_{K^+}/F_{π^+} including strong isospin breaking corrections, but excluding QED corrections. This is the standard value quoted in the FLAG reviews [8,81]. Our calculations, like most, are performed in the isospin symmetric limit, and therefore, the strong isospin breaking correction must be estimated, rather than having a direct determination. The optimal approach is to incorporate both QED and QCD isospin breaking corrections into the calculations such that the separation is not necessary, as was done in Ref. [89] by incorporating both types of corrections through the perturbative modification of the path integral and correlation functions [90,91]. In this work, we have not performed these extensive computations and so we rely upon the $SU(3)\chi$PT prediction to estimate the correction due to strong isospin breaking. As we have observed in Sec. IV D, the $SU(3)$ chiral expansion behaves and converges nicely, so we expect this approximation to be reasonable.

The NLO corrections to F_K and F_π including the strong isospin breaking corrections are given by

$$\begin{aligned}
\delta F_{\pi^\pm}^{\mathrm{NLO}} &= -\frac{\ell_{\pi^0}}{2} - \frac{\ell_{\pi^\pm}}{2} - \frac{\ell_{K^0}}{4} - \frac{\ell_{K^\pm}}{4} \\
&\quad + 4\bar{L}_4(\epsilon_{\pi^\pm}^2 + \epsilon_{K^\pm}^2 + \epsilon_{K^0}^2) + 4\bar{L}_5\epsilon_{\pi^\pm}^2, \\
\delta F_{K^\pm}^{\mathrm{NLO}} &= -\frac{\ell_{\pi^0}}{8} - \frac{\ell_{\pi^\pm}}{4} - \frac{3\ell_\eta}{8} - \frac{\ell_{K^0}}{4} - \frac{\ell_{K^\pm}}{2} \\
&\quad + \frac{1}{4}(\epsilon_{K^0}^2 - \epsilon_{K^\pm}^2)\frac{\ell_\eta - \ell_{\pi^0}}{\epsilon_\eta^2 - \epsilon_{\pi^0}^2} \\
&\quad + 4\bar{L}_4(\epsilon_{\pi^\pm}^2 + \epsilon_{K^\pm}^2 + \epsilon_{K^0}^2) + 4\bar{L}_5\epsilon_{K^\pm}^2,
\end{aligned} \tag{4.15}$$

where we have kept explicit the contribution from each flavor of meson propagating in the loop. There are three points to note in these expressions:

(1) At NLO in the $SU(3)$ chiral expansion, there are no additional LECs that describe the isospin breaking corrections beyond those that contribute to the isospin symmetric limit. Therefore, one can make a parameter-free prediction of the isospin breaking corrections using lattice results from isospin symmetric calculations, with the only assumption being that $SU(3)\chi$PT converges for this observable.

(2) If we expand these corrections about the isospin limit, they agree with the known results [5], and $\delta F_{\pi^\pm}$ is free of isospin breaking corrections at this order.

(3) We have used the kaon mass splitting in place of the quark mass splitting, which is exact at LO in χPT $B(m_d - m_u) = (\hat{M}_{K^0}^2 - \hat{M}_{K^\pm}^2)$.

The estimated shift of our isospin-symmetric result to incorporate strong isospin breaking is then

$$\delta F^{\rm iso}_{K-\pi} \equiv \frac{F_{\hat{K}^+}}{F_{\hat{\pi}^+}} - \frac{F_K}{F_\pi} = -\frac{1}{4}(\ell_{\hat{K}^+} - \ell_{\bar{K}}) + 4\bar{L}_5(\epsilon^2_{\hat{K}^+} - \epsilon^2_{\bar{K}}) + \frac{1}{4}(\epsilon^2_{K^0} - \epsilon^2_{K^\pm})\frac{\ell_\eta - \ell_{\pi^0}}{\epsilon^2_\eta - \epsilon^2_{\pi^0}}. \quad (4.16)$$

Reference [5] suggested replacing $\bar{L}_5$ with the NLO expression equating it to the isospin symmetric F_K/F_π which yields

$$\delta F^{\rm iso'}_{K-\pi} = -\frac{1}{6}\frac{\epsilon^2_{K^0} - \epsilon^2_{K^\pm}}{\epsilon^2_\eta - \epsilon^2_{\pi^0}} \times \left[4\left(\frac{F_K}{F_\pi} - 1\right) + \epsilon^2_{\bar{\pi}} \ln\left(\frac{\epsilon^2_{\bar{K}}}{\epsilon^2_{\bar{\pi}}}\right) - \epsilon^2_{\bar{K}} + \epsilon^2_{\bar{\pi}}\right]. \quad (4.17)$$

In this expression, we have utilized the two relations

$$\ell_{\hat{K}^+} - \ell_{\bar{K}} = -\frac{2}{3}\frac{\epsilon^2_{K^0} - \epsilon^2_{K^\pm}}{\epsilon^2_\eta - \epsilon^2_{\pi^0}}(\epsilon^2_{\bar{K}} - \epsilon^2_{\bar{\pi}})(\ln \epsilon^2_{\bar{K}} + 1)$$
$$\epsilon^2_\eta - \epsilon^2_{\bar{\pi}} = \frac{4}{3}(\epsilon^2_{\bar{K}} - \epsilon^2_{\bar{\pi}}). \quad (4.18)$$

At this order, both Eqs. (4.16) and (4.17) are equivalent. However, they can result in shifts that differ by more than one standard deviation. Further, the direct estimate of the strong isospin breaking corrections [9] is larger in magnitude than either of them. Therefore, to estimate the strong isospin breaking corrections, we take the larger of the two corrections as the mean and the larger uncertainty of the two, and then add an additional 25% uncertainty for $SU(3)$ truncation errors. In Sec. IV D 2 we observe the N^{2}LO correction is ∼25% of the NLO correction (while NLO is ∼16% of LO).

In order to evaluate these expressions, we have to define the physical point with strong isospin breaking and without QED isospin breaking. We employ the values from FLAG [2017] [81] (except $\hat{M}_{\pi^0} = 134.6(3)$ MeV):

$$\hat{M}_{\pi^0} = \hat{M}_{\pi^+} = 134.8(3)\ \text{MeV},$$
$$\hat{M}_{K^0} = 497.2(4)\ \text{MeV},$$
$$\hat{M}_{K^+} = 491.2(5)\ \text{MeV}. \quad (4.19)$$

With this definition of the physical point, we find (under the same model average as Table VII)

$$\delta F^{\rm iso}_{K-\pi} = -0.00188(51),$$
$$\delta F^{\rm iso'}_{K-\pi} = -0.00215(24), \quad (4.20)$$

resulting in our estimated strong isospin breaking correction

TABLE VII. List of models used in final result, as described in the text.

Model	χ^2_ν	Q	logGBF	Weight	F_K/F_π
xpt-ratio_nnnlo_FV_ct_PP	0.847	0.645	77.728	0.273	1.1968(40)
xpt-ratio_nnnlo_FV_alphaS_ct_PP	0.843	0.650	77.551	0.229	1.1962(46)
xpt_nnnlo_FV_ct_PP	0.908	0.569	76.830	0.111	1.1974(40)
xpt_nnnlo_FV_alphaS_ct_PP	0.902	0.576	76.668	0.095	1.1966(46)
xpt-ratio_nnnlo_FV_ct_PK	1.014	0.439	76.343	0.068	1.1952(37)
xpt-ratio_nnnlo_FV_alphaS_ct_PK	1.006	0.449	76.234	0.061	1.1944(42)
xpt_nnnlo_FV_PP	0.949	0.517	75.371	0.026	1.1989(40)
xpt_nnnlo_FV_alphaS_PP	0.946	0.522	75.196	0.022	1.1983(46)
xpt_nnnlo_FV_ct_PK	1.135	0.309	75.084	0.019	1.1950(36)
xpt_nnnlo_FV_alphaS_ct_PK	1.123	0.321	75.007	0.018	1.1941(41)
xpt-ratio_nnnlo_FV_PP	1.014	0.439	74.765	0.014	1.1987(40)
xpt-ratio_nnnlo_FV_alphaS_PP	1.009	0.445	74.599	0.012	1.1980(46)
xpt_nnnlo_FV_PK	1.100	0.344	74.421	0.010	1.1969(37)
xpt_nnnlo_FV_alphaS_PK	1.093	0.352	74.306	0.009	1.1962(42)
xpt-ratio_nnnlo_FV_ct_KK	1.262	0.202	74.014	0.007	1.1920(36)
xpt-ratio_nnnlo_FV_alphaS_ct_KK	1.244	0.215	74.004	0.007	1.1912(39)
xpt-ratio_nnnlo_FV_PK	1.159	0.286	73.880	0.006	1.1967(37)
xpt-ratio_nnnlo_FV_alphaS_PK	1.150	0.295	73.780	0.005	1.1959(41)
xpt_nnnlo_FV_KK	1.288	0.184	72.757	0.002	1.1938(36)
xpt_nnnlo_FV_alphaS_KK	1.273	0.194	72.718	0.002	1.1930(40)
xpt-ratio_nnnlo_FV_KK	1.338	0.152	72.348	0.001	1.1938(36)
xpt-ratio_nnnlo_FV_alphaS_KK	1.322	0.162	72.323	0.001	1.1929(39)
xpt_nnnlo_FV_alphaS_ct_KK	1.536	0.068	71.459	0.001	1.1900(38)
xpt_nnnlo_FV_ct_KK	1.558	0.061	71.430	0.001	1.1909(35)
Bayes Model Average					1.1964(42)(12)

$$\frac{F_{\hat{K}^+}}{F_{\hat{\pi}^+}} - \frac{F_K}{F_\pi} = -0.00215(72) \tag{4.21}$$

and our final result as reported in Eq. (1.4)

$$\begin{aligned}\frac{F_{\hat{K}^+}}{F_{\hat{\pi}^+}} &= 1.1942(44)(07)^{\text{iso}} \\ &= 1.1942(45),\end{aligned}$$

where the first uncertainty in the first line is the combination of those in Eq. (1.3).

V. SUMMARY AND DISCUSSION

The ratio F_K/F_π may be used, in combination with experimental input for leptonic decay widths, to make a prediction for the ratio of CKM matrix elements, $|V_{us}|/|V_{ud}|$. Using the most recent data, Eq. (1.1) becomes [6]

$$\frac{|V_{us}|}{|V_{ud}|}\frac{F_{\hat{K}^+}}{F_{\hat{\pi}^+}} = 0.2760(4), \tag{5.1}$$

where strong isospin breaking effects must be included for direct comparison with experimental data. Combining this expression with our final result, we find

$$\frac{|V_{us}|}{|V_{ud}|} = 0.2311(10). \tag{5.2}$$

Utilizing the current global average, $|V_{ud}| = 0.97420(21)$, extracted from superallowed nuclear beta decays [6] results in

$$|V_{us}| = 0.2251(10). \tag{5.3}$$

Finally, we may use our results, combined with the value $|V_{ub}| = (3.94(36)) \times 10^{-3}$, as a test of unitarity for the CKM matrix, which states that $|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 1$. From our calculation we find

$$|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 0.99977(59). \tag{5.4}$$

Alternatively, rather than using the experimental determination of $|V_{ud}|$ as input for our test of unitarity, we may instead use the global lattice average for $|V_{us}| = 0.2231(7)$ [8], extracted via the quantity $f_+(0)$, the zero momentum transfer limit of a form factor relevant for the semileptonic decay $K^0 \to \pi^- l\nu$. This leads to

$$|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 0.9812(95), \tag{5.5}$$

leading to a roughly 2σ tension with unitarity. Our result, along with the reported experimental results for $|V_{ud}|$ and lattice results for $|V_{us}|$, are shown in Fig. 9. One could also combine our results with the more precise average in the FLAG review which would lead to a slight reduction their reported uncertainties, but we will leave that to the FLAG Collaboration in their next update.

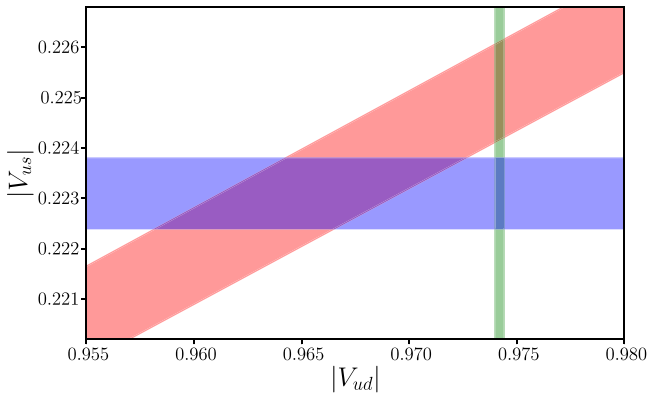

FIG. 9. Result for the ratio of CKM matrix elements, $|V_{us}|/|V_{ud}|$, extracted from the ratio F_K/F_π reported in this work (red band). The global lattice value for $|V_{us}|$ extracted from a semileptonic decay form factor, $f_+(0)$ [8], is shown as a horizontal blue band, while the global experimental average for $|V_{ud}|$ from nuclear beta decay [6] is given as a vertical green band. Note that the intersection between the red and green bands agrees well with the unitarity constraint for the CKM matrix, while the intersection between the red and blue bands shows $\sim 2\sigma$ tension.

Another motivation for this work was to precisely test (below 1%) whether the action we have used for our nucleon structure calculations [35–37] can be used to reproduce an accepted value from other lattice calculations that are known at the subpercent level. Our result provides the first subpercent cross-check of the universality of the continuum limit of this quantity with $N_f = 2 + 1 + 1$ dynamical flavors, albeit with the same sea-quark action as used by MILC/FNAL and HPQCD [12,13].

Critical in obtaining a subpercent determination of any quantity is control over the continuum extrapolation. This is relevant to our pursuit of a subpercent determination of g_A as another calculation, utilizing many of the same HISQ ensembles but with a different valence action (clover fermions), obtains a result that is in tension with our own [92,93]. While there has been speculation that this discrepancy is due to the continuum extrapolations [93], new work suggests the original work underestimated the systematic uncertainty in the correlation function analysis, and when accounted for, the tension between our results goes away [94].

In either case, to obtain a subpercent determination of g_A, which is relevant for trying to shed light on the neutron lifetime discrepancy [95], it is important to understand the scaling violations of our lattice action. While a smooth continuum extrapolation in one observable does not guarantee such a smooth extrapolation in another, it at least provides some reassurance of a well-behaved continuum extrapolation. Furthermore, the determination of F_K/F_π involves the same axial current that is relevant for the

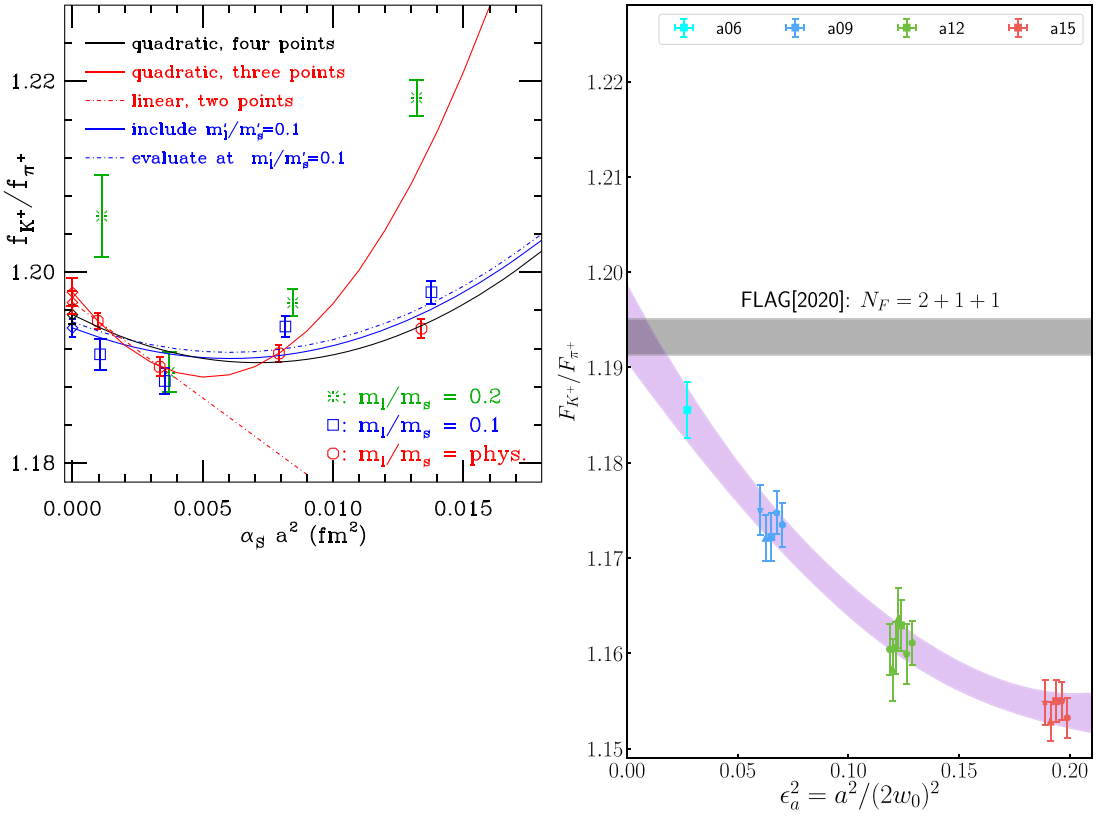

FIG. 10. Comparison of the continuum extrapolation from the MILC[2014] [96] result (left) with the continuum extrapolation in the present work with the MDWF on gradient-flowed HISQ action (right). In the right plot, we also include the FLAG[2020] average value from the $N_f = 2+1+1$ calculations [8]. While the a06m310L ensemble is not necessary for us to extrapolate to a consistent value as this FLAG average (see Fig. 4), the overall size of our discretization effects are larger. This is not necessarily surprising as the HISQ action used by MILC has perturbatively removed all $O(a^2)$ corrections such that the leading scaling violations begin at $O(\alpha_S a^2)$, as implied by the x axis of the left plot. See footnote 3 for a comment on the leading scaling violations [71–73].

computation of the nucleon matrix element used to compute g_A.

Figure 4 shows the continuum extrapolation of F_K/F_π from our analysis. The size of the discretization effects are noticeably larger than we observed in our calculation of g_A [36]. In Sec. IV D 1, we demonstrated that, while helpful, the a06m310L ensemble is not necessary to achieve a subpercent determination of F_K/F_π. This is in contrast to the determination by MILC which requires the $a \sim 0.06$ fm (or smaller) lattice spacings to control the continuum extrapolation (though we note, the HPQCD calculation [12], also performed on the HISQ ensembles, does not utilize the $a \sim 0.06$ fm ensembles but agrees with the MILC result). It should be noted, the MILC result does not rely on the heavier mass ensembles except to adjust for the slight mistuning of the input quark masses on their near-physical point ensembles. In Fig. 10, we compare our continuum extrapolation to that of MILC [96].

In Ref. [96], they also utilize the same four lattice spacings as in this work (they have subsequently improved their determination with an additional two finer lattice spacings [13].) A strong competition between the $O(a^2)$ and $O(a^4)$ corrections was observed in that work, such that the $a \sim .06$ fm ensemble is much more instrumental for a reliable continuum extrapolation than is the case in our setup. At the same time, the overall scale of their discretization effects is much smaller than we observe in the MDWF on gradient-flowed HISQ action for this quantity. This is not entirely surprising as the HISQ action has been tuned to perturbatively remove all $O(a^2)$ corrections such that the leading corrections formally begin as $O(\alpha_S a^2)$.

The analysis and supporting data for this article are openly available [97].

ACKNOWLEDGMENTS

We would like to thank V. Cirigliano, S. Simula, J. Simone, and T. Kaneko for helpful correspondence and discussions regarding the strong isospin breaking corrections to F_K/F_π. We would like to thank J. Bijnens for helpful correspondence on χPT and a C++ interface to CHIRON [69] that we used for the analysis presented in this work. We thank the MILC Collaboration for providing some of the HISQ configurations used in this work, and

A. Bazavov, C. Detar, and D. Toussaint for guidance on using their code to generate the new HISQ ensembles also used in this work. We would like to thank P. Lepage for enhancements to GVAR [98] and LSQFIT [83] that enable the pickling of LSQFIT.NONLINEAR_FIT objects. We also thank C. Bernard for useful correspondence concerning higher order extrapolation analysis and R. Sommer for comments on the leading asymptotic scaling violations. Computing time for this work was provided through the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program and the LLNL Multiprogrammatic and Institutional Computing program for Grand Challenge allocations on the LLNL supercomputers. This research utilized the NVIDIA GPU-accelerated Titan and Summit supercomputers at Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725 as well as the Surface, RZHasGPU, Pascal, Lassen, and Sierra supercomputers at Lawrence Livermore National Laboratory. The computations were performed utilizing LALIBE [99] which utilizes the CHROMA software suite [100] with QUDA solvers [39,40] and HDF5 [101] for I/O [102]. They were efficiently managed with METAQ [103,104] and status of tasks logged with EspressoDB [105]. The hybrid Monte Carlo was performed with the MILC Code [106], and for the ensembles new in this work, running on GPUs using QUDA. The final extrapolation analysis utilized GVAR v11.2 [98] and LSQFIT v11.5.1 [83] and CHIRON v0.54 [69]. This work was supported by the NVIDIA Corporation (M. A. C.), the Alexander von Humboldt Foundation through a Feodor Lynen Research Fellowship (C. K.), the DFG and the NSFC Sino-German CRC110 (E. B.), the RIKEN Special Postdoctoral Researcher Program (E. R.), the U.S. Department of Energy, Office of Science, Office of Nuclear Physics under Awards No. DE-AC02-05CH11231 (C. C. C., C. K., B. H., A. W. L.), No. DE-AC52-07NA27344 (D. A. B., D. H., A. S. G., P. V.), No. DE-FG02-93ER-40762 (E. B.), No. DE-AC05-06OR23177 (B. J., C. M., K. O.), No. DE-FG02-04ER41302 (K. O.); the Office of Advanced Scientific Computing (B. J.); the Nuclear Physics Double Beta Decay Topical Collaboration (D. A. B., H. M. C., A. N., A. W. L.); and the DOE Early Career Award Program (C. C. C., A. W. L.).

APPENDIX A: MODELS INCLUDED IN FINAL ANALYSIS

We list the models that have entered the final analysis as described in Sec. IV D and listed in Table VII. For example, the model

xpt-ratio_nnnlo_FV_alphaS_PP

indicates the model uses the continuum χPT fit function through N^{3}LO with discretization corrections added as in Eqs. (3.32) and (3.41). The NLO contributions are kept in a ratio form, Eq. (3.17), and we have included the corresponding N^{2}LO ratio correction $\delta^{\mathrm{N^2LO}}_{\mathrm{ratio}}$. The finite volume corrections have been included at NLO. The discretization terms at N^{2}LO include the $\alpha_S\epsilon_a^2(\epsilon_K^2-\epsilon_\pi^2)$ counterterm. The renormalization scale appearing in the logs is $\mu = 4\pi F_\pi$ as indicated by _PP, and we have included the corresponding N^{2}LO correction $\delta^{\mathrm{N^2LO}}_{F_\pi}$, Eq. (3.29), to hold the actual renormalization scale fixed at $\mu_0 = 4\pi F_0$.

When _ct appears in the model name, the only N^{2}LO terms that are added are from the local counterterms while all chiral log corrections are set to zero.

APPENDIX B: NLO MIXED ACTION FORMULAS

The expression for $d\ell_\pi$ arises from the integral

$$d\ell_\pi = \int_R \frac{d^dk}{(2\pi)^d}\frac{i}{(k^2-m_\pi^2)^2} = \frac{1+\ln(m_\pi^2/\mu^2)}{(4\pi)^2}, \tag{B1}$$

which has been regulated and renormalized with the standard χPT modified dimensional-regularization scheme [48]. The finite volume corrections to $\delta\ell_\pi$ are given by

$$\delta^{\mathrm{FV}} d\ell_\pi = \sum_{|\mathbf{n}|\neq 0}\frac{c_n}{(4\pi)^2}\left[\frac{2K_1(mL|\mathbf{n}|)}{mL|\mathbf{n}|} - K_0(mL|\mathbf{n}|) - K_2(mL|\mathbf{n}|)\right]. \tag{B2}$$

The expression for $\mathcal{K}_{\phi_1\phi_2}$ arises from the integral

$$\mathcal{K}_{\phi_1\phi_2} = (4\pi)^2\int_R \frac{d^dk}{(2\pi)^d}\frac{i}{(k^2-m_{\phi_1}^2)(k^2-m_{\phi_2}^2)} = \frac{\ell_{\phi_2}-\ell_{\phi_1}}{\epsilon_{\phi_2}^2-\epsilon_{\phi_1}^2}. \tag{B3}$$

Similarly, $\mathcal{K}^{(2,1)}_{\phi_1\phi_2}$ is given by

$$\mathcal{K}^{(2,1)}_{\phi_1\phi_2} = \int_R \frac{d^dk}{(2\pi)^d}\frac{i(4\pi)^2(4\pi F)^2}{(k^2-m_{\phi_1}^2)^2(k^2-m_{\phi_2}^2)} = \frac{\ell_{\phi_2}-\ell_{\phi_1}}{(\epsilon_{\phi_2}^2-\epsilon_{\phi_1}^2)^2} - \frac{d\ell_{\phi_1}}{\epsilon_{\phi_2}^2-\epsilon_{\phi_1}^2}. \tag{B4}$$

Finally, $\mathcal{K}_{\phi_1\phi_2\phi_3}$ is given by

$$\mathcal{K}_{\phi_1\phi_2\phi_3} = \int_R \frac{d^dk}{(2\pi)^d}\frac{i(4\pi)^2(4\pi F)^2}{(k^2-m_{\phi_1}^2)(k^2-m_{\phi_2}^2)(k^2-m_{\phi_3}^2)} = \frac{\ell_{\phi_1}}{(\epsilon_{\phi_1}^2-\epsilon_{\phi_2}^2)(\epsilon_{\phi_1}^2-\epsilon_{\phi_3}^2)} + \frac{\ell_{\phi_2}}{(\epsilon_{\phi_2}^2-\epsilon_{\phi_1}^2)(\epsilon_{\phi_2}^2-\epsilon_{\phi_3}^2)} + \frac{\ell_{\phi_3}}{(\epsilon_{\phi_3}^2-\epsilon_{\phi_1}^2)(\epsilon_{\phi_3}^2-\epsilon_{\phi_2}^2)}. \tag{B5}$$

TABLE VIII. Input parameters and measured acceptance rate for the new HISQ ensembles. In addition to the columns standardly reported by MILC (see Table IV of Ref. [43]), we list the abbreviated ensemble name, the number of streams N_{stream}, and the total number of configurations N_{cfg}. For a given ensemble, each stream has an equal number of configurations. The gauge coupling, light, strange, and charm quark masses on each ensembles are given as well as the tadpole factor u_0 and the Naik-term added to the charm quark action ϵ_N. Here s denotes the total length in molecular dynamics time units (MDTU) between each saved configuration, Len. denotes the length between accept/reject steps (in MDTU), and Acc. denotes the fraction of trajectories accepted. The microstep size ϵ used in the HMC is provided as Len./N_{steps} which was input with single precision. The average acceptance rate over all streams is listed as well as the number of streams.

Ensemble	$10/g^2$	am_l	am_s	am_c	u_0	ϵ_N	s	Len.	ϵ	Acc.	N_{stream}	N_{cfg}
a15m135XL	5.80	0.002426	0.06730	0.8447	0.85535	−0.35892	5	0.2	0.2/150	0.631	4	2000
a09m135	6.30	0.001326	0.03636	0.4313	0.874164	−0.11586	6	1.5	1.5/130	0.693	2	1010
a06m310L	6.72	0.0048	0.024	0.286	0.885773	−0.05330	6	2.0	2.0/120	0.765	2	1000

In each of these expressions, the corresponding expression including FV corrections are given by replacing $\ell_\phi \to \ell_\phi^{\text{FV}}$, Eq. (3.39).

APPENDIX C: HMC FOR NEW ENSEMBLES

We present various summary information for the three new ensembles used in this work, a06m310L, a15m135XL and a09m135. In Table VIII, we list the parameters of the HISQ ensembles used in the HMC. In Fig. 11, we show the MDTU history of the ΔS for the three ensembles. For the a15m135XL ensemble, we reduced the trajectory length significantly compared to the a15m130 from MILC to overcome spikes in the HMC force calculations. To compensate, we lowered the acceptance rate to encourage the HMC to move around parameter space with

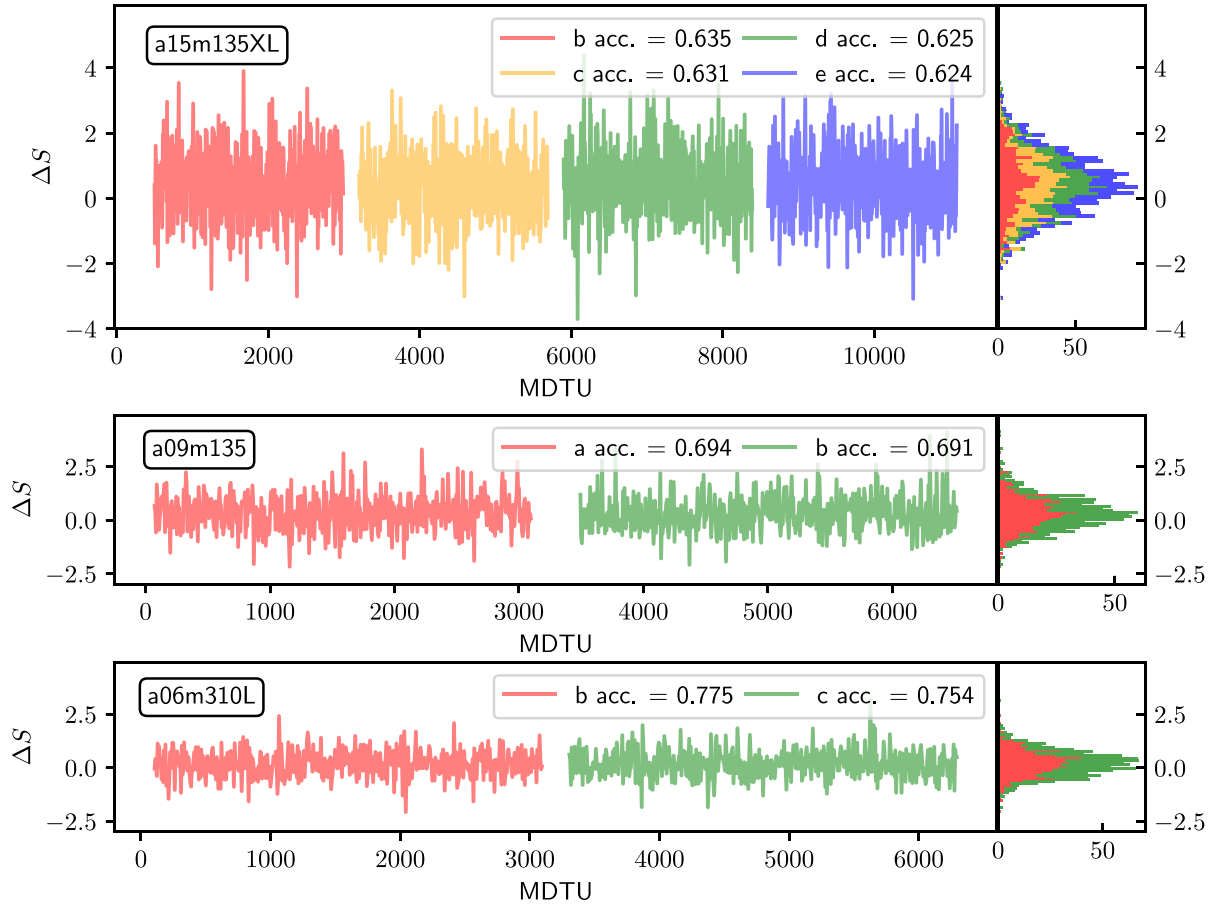

FIG. 11. The ΔS values computed in the accept/reject step of the HMC versus MDTU. The different colors correspond to the different streams which are separated and shifted in MDTU for clarity.

TABLE IX. Average values of the quark-antiquark condensate $\bar{\psi}\psi$ with statistical errors and integrated autocorrelation times τ measured with the Γ-method analysis. Each value is averaged over all the available streams (which are all statistically compatible). The integrated autocorrelation time is reported in units of MDTU. The a15m135XL results are obtained from the second half of each stream because we have more measurements.

Ensemble	N_{stream}	$\bar{\psi}\psi_l$	τ_l	$\bar{\psi}\psi_s$	τ_s	$\bar{\psi}\psi_c$	τ_c
a15m135XL	4	0.02390(2)	11(2)	0.08928(2)	71(26)	0.4800580(5)	<1
a09m135	2	0.005761(6)	9(2)	0.003935(3)	32(9)	0.3205399(5)	1.5(1)
a06m310L	2	0.006599(4)	30(8)	0.002356(2)	34(8)	0.2275664(4)	2.0(2)

larger jumps in an attempt to reduce the autocorrelation time. We ran 25 HMC accept/reject steps before saving a configuration for a total trajectory length of 5.

For each accept/reject step we also measure the quark-antiquark condensate $\bar{\psi}\psi$ using a stochastic estimate with 5 random sources that are averaged together. We compute it for each of the quark masses am_l, am_s, and am_c. On the a15m135XL we have measured $\bar{\psi}\psi$ only on every saved configuration for the first half of each stream, while we measured it at each accept/reject step for the second half. The integrated autocorrelation time, as well as the average and statistical errors of $\bar{\psi}\psi$, are computed using the Γ-method analysis [107] with the Python package UNEW [108]. We report the results in Table IX. In Fig. 12 we report the value of the $\bar{\psi}\psi$ on each saved configuration for the three quark masses on each ensemble.

Because we observe a long autocorrelation time of the $\langle\bar{\psi}_s\psi_s\rangle$ on the a15m135XL ensemble, we also studied the uncertainty on the extracted pion and kaon effective masses as a function of block size to check for possible longer autocorrelations than usual, with blocking lengths of 10, 25, and 100 MDTU (Fig. 13). We observe that these hadronic quantities have a much shorter autocorrelation time as the uncertainty is independent of τ_b and consistent with the unblocked data. On this a15m135XL ensemble, while we have generated 2000 configurations, we have only utilized

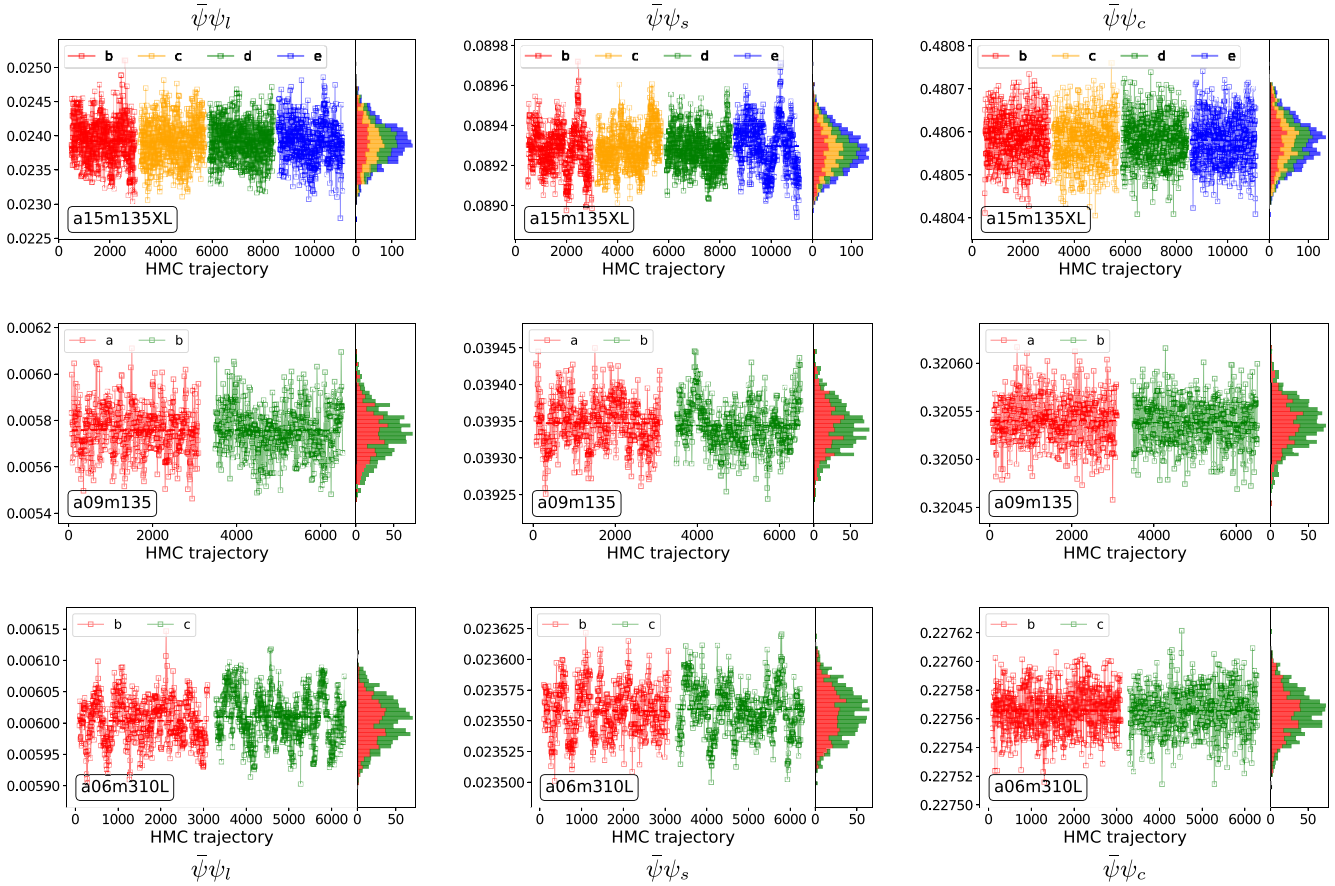

FIG. 12. The quark-antiquark condensate on each configuration of the three ensembles. The different streams are plotted separately for clarity. The plots in each column correspond to the light, strange and charm quark masses, respectively.

TABLE X. Values of the overrelaxed stout smearing parameters used to measure the topological charge Q and the resulting mean ($\bar{Q}$) and width (σ) of the distribution for each stream. The last column reports the integrated autocorrelation time in units of MDTU using the Γ-method analysis. These measurements were performed with QUDA which is now available via the `su3_test` test executable in the `develop` branch [39,40].

			$\bar{Q}_{\text{stream}}(\sigma)$						
Ensemble	ρ	N_{step}	a	b	c	d	e	$\bar{Q}_{\text{all}}(\sigma)$	$\tau_{\text{all}}(\sigma)$
a15m135XL	0.068	2000	⋯	−10(34)	−3(35)	−2(33)	5(32)	−3(33)	15(3)
a09m135	0.065	2000	0.5(12.0)	2(12)	⋯	⋯	⋯	1(12)	18(4)
a06m310L	0.066	1800	⋯	4(12)	−1.2(7.4)	⋯	⋯	1(10)	420(198)

1000 in this paper (the first half from each of the four streams).

Finally, in Table X, we list the parameters of the overrelaxed stout smearing used to measure the topological charge Q on each configuration [109] and we show the resulting Q distributions in Fig. 14. While the Q-distribution on the a06m310L ensemble is less than ideal and the integrated autocorrelation time is long, the volume is sufficiently large ($aL = 72a \simeq 4.1$ fm) that we do not anticipate any measurable impact from the poorly distributed Q-values, which nonetheless average to nearly 0.

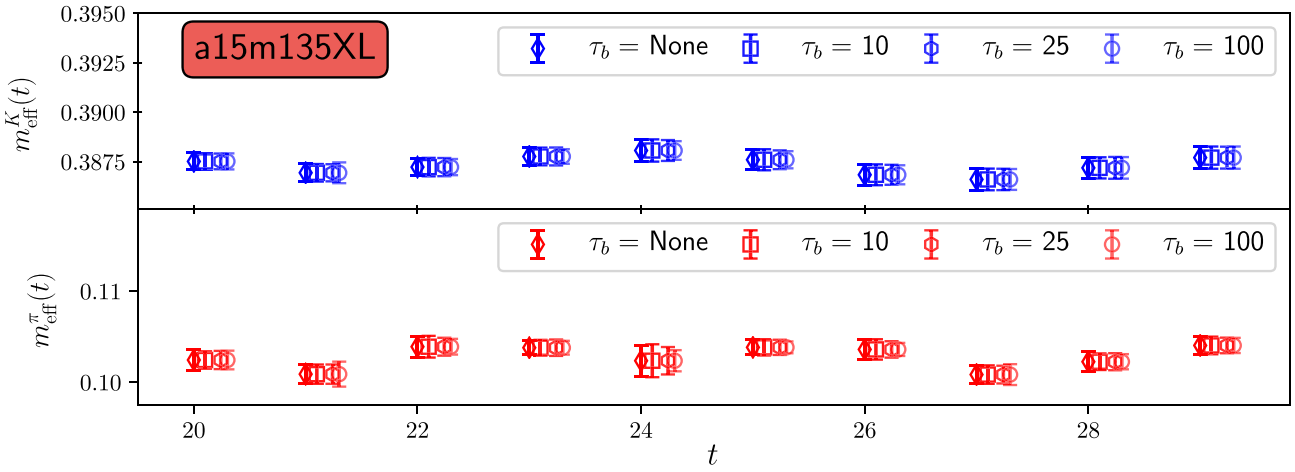

FIG. 13. The effective mass in the mid- to long-time region of the kaon (top) and pion (bottom) on the a15m135XL ensemble are plotted as a function of the blocking time, τ_b in MDTU. For example, $\tau_b = 10$ blocks nearest neighbor configurations while $\tau_b = 100$ is blocking in groups of 20 configurations. That the uncertainty is independent of τ_b indicates the autocorrelation time for these hadronic quantities is very short.

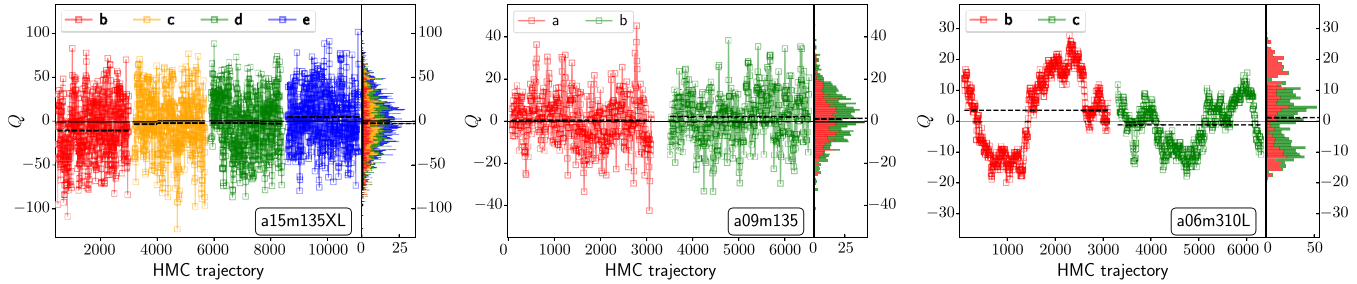

FIG. 14. Distribution of topological charge Q measurements on each configuration. The Q-values were determined by using the overrelaxed stout smearing technique outlined in [109] with weight parameters ρ given in Table X and $\varepsilon = -0.25$. We cross-checked a sample of our stout smeared measurements with the more expensive Symanzik flow technique and saw good agreement between the two. We determined the ρ parameter and the number of steps to perform on an ensemble-to-ensemble basis, i.e., for a handful of configurations per ensemble we choose a spread of ρs and step numbers and observe which combination gives the best plateau. These values of ρ and step number (ε is always -0.25) are then applied to the entire ensemble.

[1] W. J. Marciano, Precise Determination of $|V_{us}|$ from Lattice Calculations of Pseudoscalar Decay Constants, Phys. Rev. Lett. **93**, 231803 (2004).

[2] C. Aubin, C. Bernard, C. E. DeTar, J. Osborn, S. Gottlieb, E. B. Gregory, D. Toussaint, U. M. Heller, J. E. Hetrick, and R. Sugar (MILC Collaboration), Light pseudoscalar decay constants, quark masses, and low energy constants from three-flavor lattice QCD, Phys. Rev. D **70**, 114501 (2004).

[3] R. Decker and M. Finkemeier, Short and long distance effects in the decay $\tau \to \pi\nu_\tau(\gamma)$, Nucl. Phys. **B438**, 17 (1995).

[4] M. Finkemeier, Radiative corrections to π_{l2} and K_{l2} decays, Phys. Lett. B **387**, 391 (1996).

[5] V. Cirigliano and H. Neufeld, A note on isospin violation in $P_{l2(\gamma)}$ decays, Phys. Lett. B **700**, 7 (2011).

[6] M. Tanabashi *et al.* (Particle Data Group), Review of particle physics, Phys. Rev. D **98**, 030001 (2018).

[7] C. T. H. Davies *et al.* (HPQCD, UKQCD, MILC, and Fermilab Lattice Collaboration), High Precision Lattice QCD Confronts Experiment, Phys. Rev. Lett. **92**, 022001 (2004).

[8] S. Aoki *et al.* (Flavour Lattice Averaging Group), FLAG review 2019, Eur. Phys. J. C **80**, 113 (2020).

[9] N. Carrasco *et al.*, Leptonic decay constants f_K, f_D, and f_{D_s} with $N_f = 2 + 1 + 1$ twisted-mass lattice QCD, Phys. Rev. D **91**, 054507 (2015).

[10] R. Frezzotti and G. C. Rossi, Twisted mass lattice QCD with mass nondegenerate quarks, Nucl. Phys. B, Proc. Suppl. **128**, 193 (2004).

[11] R. Frezzotti and G. C. Rossi, Chirally improving Wilson fermions. 1. $O(a)$ improvement, J. High Energy Phys. 08 (2004) 007.

[12] R. J. Dowdall, C. T. H. Davies, G. P. Lepage, and C. McNeile, V_{us} from π and K decay constants in full lattice QCD with physical u, d, s, and c quarks, Phys. Rev. D **88**, 074504 (2013).

[13] A. Bazavov *et al.*, B- and D-meson leptonic decay constants from four-flavor lattice QCD, Phys. Rev. D **98**, 074512 (2018).

[14] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu, H. Trottier, and K. Wong (HPQCD and UKQCD Collaborations), Highly improved staggered quarks on the lattice, with applications to charm physics, Phys. Rev. D **75**, 054502 (2007).

[15] E. Follana, C. T. H. Davies, G. P. Lepage, and J. Shigemitsu (HPQCD and UKQCD Collaborations), High Precision Determination of the π, K, D, and D_s Decay Constants from Lattice QCD, Phys. Rev. Lett. **100**, 062002 (2008).

[16] A. Bazavov *et al.* (MILC Collaboration), Results for light pseudoscalar mesons, Proc. Sci., LATTICE2010 (**2010**) 074 [arXiv:1012.0868].

[17] S. Durr, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, T. Kurth, L. Lellouch, T. Lippert, A. Ramos, and K. K. Szabo, The ratio F_K/F_π in QCD, Phys. Rev. D **81**, 054507 (2010).

[18] T. Blum *et al.* (RBC and UKQCD Collaboration), Domain wall QCD with physical quark masses, Phys. Rev. D **93**, 074505 (2016).

[19] S. Dürr *et al.*, Leptonic decay-constant ratio f_K/f_π from lattice QCD using $2 + 1$ clover-improved fermion flavors with 2-HEX smearing, Phys. Rev. D **95**, 054513 (2017).

[20] V. G. Bornyakov, R. Horsley, Y. Nakamura, H. Perlt, D. Pleiter, P. E. L. Rakow, G. Schierholz, A. Schiller, H. Stüben, and J. M. Zanotti (QCDSF–UKQCD Collaboration), Flavor breaking effects in the pseudoscalar meson decay constants, Phys. Lett. B **767**, 366 (2017).

[21] B. Blossier *et al.* (ETM Collaboration), Pseudoscalar decay constants of kaon and D-mesons from $N_f = 2$ twisted mass Lattice QCD, J. High Energy Phys. 07 (2009) 043.

[22] E. Berkowitz, C. Bouchard, C. C Chang, M. A. Clark, B Joo, T. Kurth, C. Monahan, A. Nicholson, K. Orginos, E. Rinaldi, P. Vranas, and A. Walker-Loud, Möbius domain-wall fermions on gradient-flowed dynamical HISQ ensembles, Phys. Rev. D **96**, 054513 (2017).

[23] K. Symanzik, Continuum limit and improved action in lattice theories. 1. Principles and ϕ^4 theory, Nucl. Phys. **B226**, 187 (1983).

[24] K. Symanzik, Continuum limit and improved action in lattice theories. 2. $O(N)$ non-linear Sigma model in perturbation theory, Nucl. Phys. **B226**, 205 (1983).

[25] D. B. Renner, W. Schroers, R. Edwards, G. T. Fleming, P. Hagler, J. W. Negele, K. Orginos, A. V. Pochinski, and D. Richards (LHP Collaboration), Hadronic physics with domain-wall valence and improved staggered sea quarks, Nucl. Phys. B, Proc. Suppl. **140**, 255 (2005).

[26] R. C. Brower, H. Neff, and K. Orginos, Mobius fermions: Improved domain wall chiral fermions, Nucl. Phys. B, Proc. Suppl. **140**, 686 (2005).

[27] R. C. Brower, H. Neff, and K. Orginos, Mobius fermions, Nucl. Phys. B, Proc. Suppl. **153**, 191 (2006).

[28] R. C. Brower, H. Neff, and K. Orginos, The Möbius domain wall fermion algorithm, Comput. Phys. Commun. **220**, 1 (2017).

[29] R. Narayanan and H. Neuberger, Infinite N phase transitions in continuum Wilson loop operators, J. High Energy Phys. 03 (2006) 064.

[30] M. Lüscher and P. Weisz, Perturbative analysis of the gradient flow in non-Abelian gauge theories, J. High Energy Phys. 02 (2011) 051.

[31] M. Lüscher, Chiral symmetry and the Yang–Mills gradient flow, J. High Energy Phys. 04 (2013) 123.

[32] M. Lüscher, Properties and uses of the Wilson flow in lattice QCD, J. High Energy Phys. 08 (2010) 071.

[33] R. Lohmayer and H. Neuberger, Continuous smearing of Wilson loops, Proc. Sci., LATTICE2011 (**2011**) 249 [arXiv:1110.3522].

[34] A. Nicholson, E. Berkowitz, H. Monge-Camacho, D. Brantley, N. Garron, C. C. Chang, E. Rinaldi, M. A. Clark, B. Joó, T. Kurth, B. Tiburzi, P. Vranas, and A. Walker-Loud, Heavy Physics Contributions to Neutrinoless Double Beta Decay from QCD, Phys. Rev. Lett. **121**, 172501 (2018).

[35] E. Berkowitz, C. Bouchard, D. B. Brantley, C. C. Chang, M. A. Clark, N. Garron, B. Joó, T. Kurth, C. Monahan, H. Monge-Camacho, A. Nicholson, K. Orginos, E. Rinaldi, P. Vranas, and A. Walker-Loud, An accurate calculation of

the nucleon axial charge with lattice QCD, arXiv:1704.01114.
[36] C. C. Chang, A. Nicholson, E. Rinaldi, E. Berkowitz, N. Garron, D. A. Brantley, H. Monge-Camacho, C. J. Monahan, C. Bouchard, M. A. Clark, B. Joó, T. Kurth, K. Orginos, P. Vranas, and A. Walker-Loud, A per-cent-level determination of the nucleon axial coupling from quantum chromodynamics, Nature (London) **558**, 91 (2018).
[37] E. Berkowitz, M. A. Clark, A. Gambhir, K. McElvain, A. Nicholson, E. Rinalid, P. Vranas, A. Walker-Loud, C. C. Chang, B. Joó, T. Kurth, and K. Orginos, Simulating the weak death of the neutron in a femtoscale universe with near-exascale computing, arXiv:1810.01609.
[38] T. Bhattacharya, V. Cirigliano, S. Cohen, R. Gupta, A. Joseph, H.-W. Lin, and B. Yoon (PNDME Collaboration), Iso-vector and Iso-scalar tensor charges of the nucleon from lattice QCD, Phys. Rev. D **92**, 094511 (2015).
[39] M. A. Clark, R. Babich, K. Barros, R. C. Brower, and C. Rebbi, Solving lattice QCD systems of equations using mixed precision solvers on GPUs, Comput. Phys. Commun. **181**, 1517 (2010).
[40] R. Babich, M. A. Clark, B. Joo, G. Shi, R. C. Brower *et al.*, Scaling lattice QCD beyond 100 GPUs, in *SC'11: Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis* (2011), pp. 1–11, https://doi.org/10.1145/2063384.2063478.
[41] O. Bar, G. Rupak, and N. Shoresh, Simulations with different lattice Dirac operators for valence and sea quarks, Phys. Rev. D **67**, 114505 (2003).
[42] A. Bazavov *et al.* (MILC Collaboration), Scaling studies of QCD with the dynamical HISQ action, Phys. Rev. D **82**, 074501 (2010).
[43] A. Bazavov *et al.* (MILC Collaboration), Lattice QCD ensembles with four flavors of highly improved staggered quarks, Phys. Rev. D **87**, 054505 (2013).
[44] D. B. Kaplan, A method for simulating chiral fermions on the lattice, Phys. Lett. B **288**, 342 (1992).
[45] Y. Shamir, Chiral fermions from lattice boundaries, Nucl. Phys. **B406**, 90 (1993).
[46] V. Furman and Y. Shamir, Axial symmetries in lattice QCD with Kaplan fermions, Nucl. Phys. **B439**, 54 (1995).
[47] P. Langacker and H. Pagels, Chiral perturbation theory, Phys. Rev. D **8**, 4595 (1973).
[48] J. Gasser and H. Leutwyler, Chiral perturbation theory to one loop, Ann. Phys. (N.Y.) **158**, 142 (1984).
[49] H. Leutwyler, On the foundations of chiral perturbation theory, Ann. Phys. (N.Y.) **235**, 165 (1994).
[50] S. R. Sharpe and R. L. Singleton, Jr, Spontaneous flavor and parity breaking with Wilson fermions, Phys. Rev. D **58**, 074501 (1998).
[51] O. Bar, G. Rupak, and N. Shoresh, Chiral perturbation theory at $\mathcal{O}(a^2)$ for lattice QCD, Phys. Rev. D **70**, 034508 (2004).
[52] O. Bar, C. Bernard, G. Rupak, and N. Shoresh, Chiral perturbation theory for staggered sea quarks and Ginsparg-Wilson valence quarks, Phys. Rev. D **72**, 054502 (2005).
[53] B. C. Tiburzi, Baryons with Ginsparg-Wilson quarks in a staggered sea, Phys. Rev. D **72**, 094501 (2005); Erratum, Phys. Rev. D **79**, 039904 (2009).
[54] J.-W. Chen, D. O'Connell, R. S. Van de Water, and A. Walker-Loud, Ginsparg-Wilson pions scattering on a staggered sea, Phys. Rev. D **73**, 074510 (2006).
[55] J.-W. Chen, D. O'Connell, and A. Walker-Loud, Two meson systems with Ginsparg-Wilson valence quarks, Phys. Rev. D **75**, 054501 (2007).
[56] K. Orginos and A. Walker-Loud, Mixed meson masses with domain-wall valence and staggered sea fermions, Phys. Rev. D **77**, 094505 (2008).
[57] F.-J. Jiang, Mixed action lattice spacing effects on the nucleon axial charge, arXiv:hep-lat/0703012.
[58] J.-W. Chen, D. O'Connell, and A. Walker-Loud, Universality of mixed action extrapolation formulae, J. High Energy Phys. 04 (2009) 090.
[59] J.-W. Chen, M. Golterman, D. O'Connell, and A. Walker-Loud, Mixed action effective field theory: An addendum, Phys. Rev. D **79**, 117502 (2009).
[60] T. Blum *et al.*, Quenched lattice QCD with domain wall fermions and the chiral limit, Phys. Rev. D **69**, 074502 (2004).
[61] Y. Aoki *et al.*, Domain wall fermions with improved gauge actions, Phys. Rev. D **69**, 074504 (2004).
[62] S. Syritsyn and J. W. Negele, Oscillatory terms in the domain wall transfer matrix, Proc. Sci., LAT2007 (**2007**) 078.
[63] S. Borsanyi *et al.*, High-precision scale setting in lattice QCD, J. High Energy Phys. 09 (2012) 010.
[64] G. Amoros, J. Bijnens, and P. Talavera, Two point functions at two loops in three flavor chiral perturbation theory, Nucl. Phys. **B568**, 319 (2000).
[65] B. Ananthanarayan, J. Bijnens, and S. Ghosh, An analytic analysis of the pion decay constant in three-flavoured Chiral perturbation theory, Eur. Phys. J. C **77**, 497 (2017).
[66] B. Ananthanarayan, J. Bijnens, S. Friot, and S. Ghosh, Analytic representations of m_K, F_K, m_η and F_η in two loop $SU(3)$ chiral perturbation theory, Phys. Rev. D **97**, 114004 (2018).
[67] B. Ananthanarayan, J. Bijnens, S. Friot, and S. Ghosh, Analytic representation of F_K/F_π in two loop chiral perturbation theory, Phys. Rev. D **97**, 091502 (2018).
[68] J. Gasser and H. Leutwyler, Chiral perturbation theory: Expansions in the mass of the strange quark, Nucl. Phys. **B250**, 465 (1985).
[69] J. Bijnens, CHIRON: A package for ChPT numerical results at two loops, Eur. Phys. J. C **75**, 27 (2015).
[70] S. R. Beane, P. F. Bedaque, K. Orginos, and M. J. Savage (NPLQCD Collaboration), f_K/f_π in full QCD with domain wall valence quarks, Phys. Rev. D **75**, 094501 (2007).
[71] J. Balog, F. Niedermayer, and P. Weisz, Logarithmic corrections to $O(a^2)$ lattice artifacts, Phys. Lett. B **676**, 188 (2009).
[72] J. Balog, F. Niedermayer, and P. Weisz, The Puzzle of apparent linear lattice artifacts in the 2d nonlinear sigma-model and Symanzik's solution, Nucl. Phys. **B824**, 563 (2010).
[73] N. Husung, P. Marquard, and R. Sommer, Asymptotic behavior of cutoff effects in Yang-Mills theory and in Wilson's lattice QCD, Eur. Phys. J. C **80**, 200 (2020).
[74] J.-W. Chen and M. J. Savage, Baryons in partially quenched chiral perturbation theory, Phys. Rev. D **65**, 094001 (2002).

[75] J. Gasser and H. Leutwyler, Light quarks at low temperatures, Phys. Lett. B **184**, 83 (1987).

[76] G. Colangelo and C. Haefeli, An asymptotic formula for the pion decay constant in a large volume, Phys. Lett. B **590**, 258 (2004).

[77] G. Colangelo, S. Durr, and C. Haefeli, Finite volume effects for meson masses and decay constants, Nucl. Phys. **B721**, 136 (2005).

[78] J. Bijnens and T. Rössler, Finite volume at two loops in chiral perturbation theory, J. High Energy Phys. 01 (2015) 034.

[79] J. Bijnens and T. Rössler, Finite volume for three-flavour partially quenched chiral perturbation theory through NNLO in the meson sector, J. High Energy Phys. 11 (2015) 097.

[80] J. Bijnens and G. Ecker, Mesonic low-energy constants, Annu. Rev. Nucl. Part. Sci. **64**, 149 (2014).

[81] S. Aoki *et al.*, Review of lattice results concerning low-energy particle physics, Eur. Phys. J. C **77**, 112 (2017).

[82] R. E. Kass and A. E. Raftery, Bayes factors, J. Am. Stat. Assoc. **90**, 773 (1995).

[83] P. Lepage, gplepage/lsqfit: lsqfit version 11.5.1 (2020), https://github.com/gplepage/lsqfit.

[84] W. I. Jay and E. T. Neil, Bayesian model averaging for analysis of lattice field theory results, arXiv:2008.01069.

[85] C. Bernard, Effective field theories and lattice QCD, Proc. Sci., CD15 (**2015**) 004 [arXiv:1510.02180].

[86] J. Bijnens, N. Danielsson, and T. A. Lahde, The pseudoscalar meson mass to two loops in three-flavor partially quenched chiPT, Phys. Rev. D **70**, 111503 (2004).

[87] J. Bijnens and T. A. Lahde, Decay constants of pseudoscalar mesons to two loops in three-flavor partially quenched chiral perturbation theory, Phys. Rev. D **71**, 094502 (2005).

[88] J. Bijnens, N. Danielsson, and T. A. Lahde, Three-flavor partially quenched chiral perturbation theory at NNLO for meson masses and decay constants, Phys. Rev. D **73**, 074509 (2006).

[89] D. Giusti, V. Lubicz, G. Martinelli, C. T. Sachrajda, F. Sanfilippo, S. Simula, N. Tantalo, and C. Tarantino, First Lattice Calculation of the QED Corrections to Leptonic Decay Rates, Phys. Rev. Lett. **120**, 072001 (2018).

[90] G. M. de Divitiis, R. Frezzotti, V. Lubicz, G. Martinelli, R. Petronzio, G. C. Rossi, F. Sanfilippo, S. Simula, and N. Tantalo (RM123 Collaboration), Leading isospin breaking effects on the lattice, Phys. Rev. D **87**, 114505 (2013).

[91] D. Giusti, V. Lubicz, C. Tarantino, G. Martinelli, F. Sanfilippo, S. Simula, and N. Tantalo, Leading isospin-breaking corrections to pion, kaon and charmed-meson masses with twisted-mass fermions, Phys. Rev. D **95**, 114504 (2017).

[92] T. Bhattacharya, V. Cirigliano, S. Cohen, R. Gupta, H.-W. Lin, and B. Yoon, Axial, scalar and tensor charges of the nucleon from $2+1+1$-flavor lattice QCD, Phys. Rev. D **94**, 054508 (2016).

[93] R. Gupta, Y.-C. Jang, B. Yoon, H.-W. Lin, V. Cirigliano, and T. Bhattacharya, Isovector charges of the nucleon from $2+1+1$-flavor lattice QCD, Phys. Rev. D **98**, 034503 (2018).

[94] Y.-C. Jang, R. Gupta, B. Yoon, and T. Bhattacharya, Axial Vector Form Factors from Lattice QCD that Satisfy the PCAC Relation, Phys. Rev. Lett. **124**, 072002 (2020).

[95] A. Czarnecki, W. J. Marciano, and A. Sirlin, Neutron Lifetime and Axial Coupling Connection, Phys. Rev. Lett. **120**, 202002 (2018).

[96] A. Bazavov *et al.* (Fermilab Lattice and MILC Collaborations), Charmed and light pseudoscalar meson decay constants from four-flavor lattice QCD with physical light quarks, Phys. Rev. D **90**, 074509 (2014).

[97] CalLat repository https://github.com/callat-qcd/project_fkfpi. See also DOI: https://doi.org/10.5281/zenodo.3979560.

[98] P. Lepage, gplepage/gvar: gvar version 11.2 (2020), https://github.com/gplepage/gvar.

[99] LALIBE: CalLat Collaboration code for lattice QCD calculations, https://github.com/callat-qcd/lalibe.

[100] R. G. Edwards and B. Joo (SciDAC, LHPC, and UKQCD Collaborations), The Chroma software system for lattice QCD, Nucl. Phys. B, Proc. Suppl. **140**, 832 (2005).

[101] The HDF Group, Hierarchical Data Format, version 5 (1997-NNNN), http://www.hdfgroup.org/HDF5/.

[102] T. Kurth, A. Pochinsky, A. Sarje, S. Syritsyn, and A. Walker-Loud, High-performance I/O: HDF5 for lattice QCD, Proc. Sci., LATTICE2014 (**2015**) 045 [arXiv:1501.06992].

[103] E. Berkowitz, METAQ: Bundle supercomputing tasks (2017), https://github.com/evanberkowitz/metaq.

[104] E. Berkowitz, G. R. Jansen, K. McElvain, and A. Walker-Loud, Job management and task bundling, EPJ Web Conf. **175**, 09007 (2018).

[105] C. C. Chang, C. Körber, and A. Walker-Loud, EspressoDB: A scientific database for managing high-performance computing workflow, J. Open Source Software **5**, 2007 (2020).

[106] MILC Collaboration code for lattice QCD calculations, https://github.com/milc-qcd/milc_qcd.

[107] U. Wolff (ALPHA Collaboration), Monte Carlo errors with less errors, Comput. Phys. Commun. **156**, 143 (2004); Erratum, Comput. Phys. Commun. **176**, 383 (2007).

[108] B. De Palma, M. Erba, L. Mantovani, and N. Mosco, A Python program for the implementation of the Γ -method for Monte Carlo simulations, Comput. Phys. Commun. **234**, 294 (2019).

[109] P. J. Moran and D. B. Leinweber, Overimproved stout-link smearing, Phys. Rev. D **77**, 094501 (2008).