We report the results of a lattice quantum chromodynamics calculation of $F_K/F_{\pi}$ using Möbius domain-wall fermions computed on gradient-flowed $N_f = 2 + 1 + 1$ highly-improved staggered quark 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 $F_K/F_{\pi}$. We also include an estimate of the strong isospin breaking corrections and arrive at a final result of $F_K/F_{\pi} = 1.1942(45)$ 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 $|V_{us}|/|V_{ud}| = 0.2311(10)$.

Contents

I. Introduction 1
II. Details of the lattice calculation 2
   A. MDWF on gradient-flowed HISQ 2
   B. Correlation function construction and analysis 4
   1. Analysis strategy 4
III. Extrapolation Functions 5
   A. N^2LO chPT 5
   B. Discretization corrections 8
   C. Finite volume corrections 10
   D. N^3LO corrections 11
IV. Extrapolation Details and Uncertainty Analysis 11
   A. Prior widths for LECs 11
   B. Physical point 12
   C. Model averaging procedure 12
   D. Full analysis and uncertainty breakdown 13
      1. Impact of a06m310L ensemble 13
      2. Convergence of the chiral expansion 14
   E. QCD isospin breaking corrections 17
V. Summary and Discussion 18
Acknowledgments 20
A. MODELS INCLUDED IN FINAL ANALYSIS 20
B. NLO MIXED ACTION FORMULAS 20
C. HYBRID MONTE CARLO FOR NEW ENSEMBLES 21
References 24

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
suns 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 con-
stants, $F_K/F_\pi$, respectively, may be related to the ratio
of light and strange CKM matrix elements $|V_{us}|, |V_{ud}|$
via [1, 2],

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

(1.1)

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

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

Using lattice quantum chromodynamics (QCD) calcula-
tions of the ratio of decay constants in the above ex-
pression yields one of the most precise determina-
tions 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.

$F_K/F_\pi$ is a so-called gold-plated quantity [7] for cal-
culating 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 statisti-
cal fluctuations between numerator and denominator, as
down the lack of signal-to-noise issues associated with
calculations involving, for instance, nucleons. Lattice
QCD calculations of $F_K/F_\pi$ are now a mature endeavor,
with state-of-the-art calculations determining this quan-
tity 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 phys-
ical point by multiple groups, quotes a value of

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

(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 fer-
nions [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 $F_K/F_\pi$
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 quan-
tity also allows for precise testing of new lattice actions,
and the demonstration of control over systematic un-
certainties for a given action. FLAG also reports av-

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

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

(1.4)

where the first uncertainty in the first line is the combi-
nation 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 extrac-
ting 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 conver-
gence 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 ac-
TABLE I. Input parameters for our lattice action. The abbreviated ensemble name [23] 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 $m_L = 4$.

| Ensemble | $\beta$ | $N_{t_{cf}}$ | volume | $am_s$ | $am_{c}\times 10^{-2}$ | $L_2/a$ | $aM_s$ | $b_5-c_5$ | $am_{val}$ | $am_{res}\times 10^4$ | $am_{val}$ | $am_{res}\times 10^4$ | $\sigma$ | $N_{res}$ |
|----------|--------|-------------|--------|--------|-------------------|-------|-------|-----------|---------|------------------|---------|------------------|-------|---------|
| a15m400  | 5.80   | 1000       | 16 x 48| 0.0217 | 0.065  | 0.838          | 12    | 1.3   | 1.50, 0.50 | 0.0278 | 9.365(67)       | 0.0902  | 6.937(63)        | 3.0   | 30      |
| a15m350  | 5.80   | 1000       | 16 x 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      |
| a15m310  | 5.80   | 1000       | 16 x 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      |
| a12m220  | 5.80   | 1000       | 24 x 48| 0.0064 | 0.064  | 0.828          | 16    | 1.3   | 1.75, 0.75 | 0.0172 | 7.536(38)       | 0.0902  | 3.890(25)        | 4.5   | 60      |
| a15m15XL | 5.80   | 1000       | 48 x 64| 0.00242 | 0.06730 | 0.8447        | 24    | 1.3   | 2.25, 1.25 | 0.00237 | 2.706(08)       | 0.0945  | 1.860(09)        | 3.0   | 30      |
| a12m400  | 6.00   | 1000       | 24 x 64| 0.0170 | 0.0599 | 0.635          | 8     | 1.2   | 1.25, 0.25 | 0.0219 | 7.337(50)       | 0.0903  | 5.129(35)        | 3.0   | 30      |
| a12m350  | 6.00   | 1000       | 24 x 64| 0.0130 | 0.0599 | 0.635          | 8     | 1.2   | 1.25, 0.25 | 0.0166 | 7.579(52)       | 0.0903  | 5.062(34)        | 3.0   | 30      |
| a12m310  | 6.00   | 1000       | 24 x 64| 0.0102 | 0.0599 | 0.635          | 8     | 1.2   | 1.25, 0.25 | 0.0126 | 7.702(52)       | 0.0903  | 4.950(35)        | 3.0   | 30      |
| a12m220S| 6.00   | 1000       | 24 x 64| 0.00507| 0.0597 | 0.628          | 12    | 1.2   | 1.50, 0.50 | 0.00600 | 3.990(42)       | 0.0903  | 2.390(24)        | 6.0   | 90      |
| a12m220  | 6.00   | 1000       | 32 x 64| 0.00507| 0.0597 | 0.628          | 12    | 1.2   | 1.50, 0.50 | 0.00600 | 4.050(20)       | 0.0903  | 2.364(15)        | 6.0   | 90      |
| a12m220L| 6.00   | 1000       | 40 x 64| 0.00507| 0.0597 | 0.628          | 12    | 1.2   | 1.50, 0.50 | 0.00600 | 4.400(26)       | 0.0903  | 2.361(19)        | 6.0   | 90      |
| a12m130  | 6.00   | 1000       | 48 x 64| 0.00184| 0.0597 | 0.628          | 20    | 1.2   | 2.00, 1.00 | 0.00195 | 1.642(09)       | 0.0903  | 0.945(08)        | 3.0   | 30      |
| a09m400  | 6.30   | 1201       | 32 x 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      |
| a09m500  | 6.30   | 1201       | 32 x 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      |
| a09m310  | 6.30   | 780        | 32 x 64| 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     |
| a09m220  | 6.30   | 1001       | 48 x 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      |
| a09m135  | 6.30   | 1010       | 64 x 96| 0.001326| 0.03636 | 0.4313        | 12    | 1.1   | 1.50, 0.50 | 0.00152 | 0.938(06)       | 0.0475  | 0.418(04)        | 3.5   | 45      |
| a06m310L | 6.72   | 1000       | 72 x 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      |

* 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 [24, 25]. These configurations are available to any interested party upon request, and will be available for easy anonymous downloading—hopefully soon.
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 Ref. [22] and [38, 39]. 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]

$$F_{q_1 q_2} = \frac{z_{0p}^{q_1 q_2} m_{q_1} + m_{q_1} z_{0s}^{q_1 q_2} + m_{q_2} + m_{q_2} z_{0s}^{q_1 q_2}}{(E_{q_1 q_2}^0)^{3/2}},$$  \hspace{1cm} (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

$$C_{(ss)ps} (t) = \sum_n z_{n(s)p}^{q_1 q_2} e^{-E_{n}^{q_1 q_2} t} + e^{-E_{n}^{q_1 q_2} (T-t)}$$  \hspace{1cm} (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 $m_{\text{res}}$ is calculated by the ratio of two-point correlation functions evaluated at the mid point of the fifth dimension $L_5/2$ and bounded on the domain wall [31]

$$m_{\text{res}}(t) = \frac{\sum x \langle \pi(t,x,L_5/2) \pi(0,0,0) \rangle}{\sum x \langle \pi(t,x,0) \pi(0,0,0) \rangle},$$  \hspace{1cm} (2.3)

where $\pi(t,x,s) \equiv \bar{q}(t,x,w) \gamma_5 q(t,x,w)$ is the pseudoscalar interpolating operator at time $t$, space $x$ and fifth dimension $s$. We extract $m_{\text{res}}$ 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 fm and 0.12 fm lattice spacing two-point correlation functions, and separations between 0.75 and 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 de-
violation. 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 and 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. Fig. 1 shows the stability of the determination of $E_0$ for the pion and kaon on the a12m130 ensemble versus $t_{\text{min}}$ 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 determinations of $F_K/F_{\pi}$ (a15m135XL, a12m130, and a09m135, see Tab. 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 MALQCD action [52] in which $O(m_{\pi}^2) \sim O(a^2\Lambda_{\text{QCD}}^2)$ are all treated as small scales. For the quark mass expansion, the dimensionless small parameters $(m_p/4\pi F)^2$ naturally emerge from $\chi$PT where $P \in \{\pi, K, \eta\}$. For the discretization corrections, while $a\Lambda_{\text{QCD}}^2$ 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 $\Lambda_{\text{QCD}}$ is a QCD scale that does not emerge in $\chi$PT.

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

$$
\epsilon_p^2 = \left( \frac{m_p}{\Lambda_\chi} \right)^2 = \left( \frac{m_p}{4\pi F} \right)^2, \quad \epsilon_a^2 = \left( \frac{a}{2w_0} \right)^2.
$$

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 $\Lambda_\chi$. This particular choice of $\epsilon_a$ is chosen such that the range of values of this small parameter roughly corresponds to $\epsilon_a^2 \leq \epsilon_p^2 \leq \epsilon_K^2$ as the lattice spacing is varied, similar to the variation of $\epsilon_K^2$ itself over the range of pion masses used, see Tab. II. As we will discuss in Sec. IV, this choice of $\epsilon_a$ 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 $g_A$ [38, 39].

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

$$
\begin{align*}
\text{NLO: } & O(\epsilon^2_p) \sim O(\epsilon_a^2), \\
\text{N}^3\text{LO: } & O(\epsilon^6_p) \sim O(\epsilon^4_p\epsilon^2_a) \sim O(\epsilon^4_a), \\
\text{N}^4\text{LO: } & O(\epsilon^8_p) \sim O(\epsilon^4_p\epsilon^4_a) \sim O(\epsilon^6_a).
\end{align*}
$$

(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 $O(\epsilon^2_p)$ correction as it must accompany terms which vanish in the SU(3)$_V$ limit, such as $\epsilon^2_p - \epsilon^2_a$. Therefore, at NLO, there cannot be any counterterms proportional to $\epsilon^2_a$ 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 $\chi$PT theory expression for $F_K/F_\pi$ in the continuum limit [64–67].

#### A. N$^2$LO $\chi$PT

The analytic expression for $F_K/F_\pi$ up to N$^2$LO is [67]

$$
\frac{F_K}{F_\pi} = 1 + 5 \frac{\ell_\pi}{8} - \frac{1}{4} \frac{\ell_\pi}{\ell_K} + \frac{3}{8} \frac{\ell_\pi}{\ell_\eta} + 4 \ell_5 (\epsilon^2_p - \epsilon_a^2)
\begin{align*}
&+ \epsilon^4_p F_F \left( \frac{m_p^2}{m_K^2} \right) + \hat{K}^+_1 \lambda_\pi^2 + \hat{K}^+_2 \lambda_\pi \lambda_K \\
&+ \hat{K}^+_3 \lambda_\pi \lambda_\eta + \hat{K}^+_4 \lambda_K^2 + \hat{K}^+_5 \lambda_\eta + \hat{K}^+_6 \lambda_\eta^2 \\
&+ \hat{C}^+_1 \lambda_\pi + \hat{C}^+_2 \lambda_K + \hat{C}^+_3 \lambda_\eta + \hat{C}^+_4.
\end{align*}
$$

(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

$$
\epsilon^2_p = \left( \frac{m_p}{4\pi F_{\pi}(m_p)} \right)^2
$$

(3.4)

where $F_{\pi}(m_p)$ is the “on-shell” pion decay constant at the masses $m_p$. The quantities $\ell_p$ are defined as

$$
\ell_p = \epsilon^2_p \ln \left( \frac{m_p^2}{\mu^2} \right)
$$

(3.5)
TABLE II. Extracted masses and decay constants from correlation functions. An HDF5 file is provided with this publication which includes the generated bootstrap samples which can be used to construct the correlated uncertainties. The small parameters in this table are defined as $\epsilon_{\pi,K} = m_{\pi,K}/(4\pi F_\pi)$, $\epsilon_{\alpha} = a/(2\pi)$. The normalization of $\epsilon_{\alpha}$ is chosen such that as a small parameter, it spans the range of $\epsilon_{\pi}^2 \lesssim \epsilon_{\pi}^2 \lesssim \epsilon_{\pi}^2$.

| Ensemble | $a_{m_{\pi}}$ | $a_{m_{K}}$ | $\epsilon_{\pi}^2$ | $\epsilon_{K}^2$ | $m_{\pi} L$ | $\alpha_S$ | $a_{F_\pi}$ | $a_{F_K}$ | $F_{K}/F_\pi$ |
|----------|----------------|----------------|----------------|----------------|----------------|----------|-------------|-------------|----------------|
| a15m400  | 0.30281(31)    | 0.42723(27)    | 0.09216(33)    | 0.18344(62)    | 4.85           | 0.19378(13) | 0.58801    | 0.07938(00) | 0.08504(09)   | 1.0713(09)   |
| a15m350  | 0.26743(30)    | 0.41369(28)    | 0.07505(28)    | 0.18326(60)    | 4.24           | 0.19378(13) | 0.58801    | 0.07960(11) | 0.08370(09)   | 1.0848(09)   |
| a15m310  | 0.23601(29)    | 0.40457(25)    | 0.06223(17)    | 0.18254(48)    | 3.78           | 0.19378(13) | 0.58801    | 0.07579(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)   |
| a15m15XL| 0.10923(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)   |

where $\mu$ is a renormalization scale. The coefficient $\bar L_5 = (4\pi)^2 L_5^I(\mu)$ 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:

$$L_i \equiv (4\pi)^2 L_i^I(\mu).$$  \hspace{1cm} (3.6)

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

$$m_\eta^2 \equiv \frac{4}{3} m_K^2 - \frac{1}{3} m_\pi^2,$$  \hspace{1cm} (3.7)

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

$$\lambda_P \equiv \ln \left( \frac{m_\eta^2}{\mu^2} \right).$$  \hspace{1cm} (3.8)

The $\ln^2$ terms are encapsulated in the $F_F(x)$ function, defined in Eqs. (8-17) of Ref. [67], and the $K_i^\tau_\alpha \lambda_P \lambda_P$, terms whose coefficients are given by

$$K_1^\tau = \frac{11}{24} \frac{2\pi^2}{\epsilon_K^2} - \frac{131}{192} \epsilon_K^4,$$  \hspace{1cm} (3.9)

The single log coefficients $\hat C_{i-3}$ are combinations of the NLO Gasser-Leutwyler coefficients

$$\hat C_i^\tau = c_i^\pi \pi^A + c_i^K K^A + c_i^K K^A,$$  \hspace{1cm} (3.10)

The single log coefficients $\hat C_{i-3}$ are combinations of the NLO Gasser-Leutwyler coefficients

$$\hat C_i = c_i^\pi \pi^A + c_i^K K^A + c_i^K K^A,$$  \hspace{1cm} (3.10)

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 **file provided by J. Bijnens to CHERON** [69], the package for two-loop $\chi$PT 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 $K_i^\tau$ term presented in Ref. [67]: a simple power-counting reveals the $\xi_K = K^A$ accompanying this term should not be there.
where

\[ c_1^{\pi} = -\frac{113}{72} - 2(2L_1 + 5L_2) - \frac{13}{2}L_3 + \frac{21}{2}L_5, \]
\[ c_1^{K} = -\frac{7}{9} - \frac{11}{2}L_5, \]
\[ c_1^{KK} = c_2^{\pi} = 0, \]
\[ c_2^{K} = \frac{209}{144}, \]
\[ c_2^{KK} = \frac{53}{96} + 2(2L_1 + 5L_2) + 5L_3 - 5L_5, \]
\[ c_3^{\pi} = \frac{19}{288} + \frac{1}{6}L_3 + \frac{11}{6}L_5 - 8(2L_7 + L_8), \]
\[ c_3^{K} = -\frac{4}{9} - \frac{4}{3}L_3 - \frac{25}{6}L_5 + 16(2L_7 + L_8), \]
\[ c_3^{KK} = \frac{13}{18} - \frac{8}{3}L_3 + \frac{2}{3}L_5 + 6(2L_7 + L_8). \] (3.11)

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

\[ \hat{C}_4^{\pi} = (c_4^K - c_4^\pi) \left[(A_4^K + L_4^K)c_4^K + (A_4^\pi + L_4^\pi)c_4^\pi\right] \] (3.12)

which are linear combinations of the N\(^2\)LO counterterms

\[ A_4^K = 16(4\pi)^4(C_{14}^\pi + C_{15}^\pi), \]
\[ A_4^\pi = 8(4\pi)^4(C_{15}^\pi + 2C_{17}^\pi), \] (3.13)

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

\[ L_4^K = 8L_5(8(L_4 - 2L_6) + 3L_5 - 8L_8) \]
\[ - 2L_1 - L_2 - \frac{1}{18}L_3 + \frac{4}{3}L_5 - 8(2L_7 + L_8), \]
\[ L_4^\pi = 8L_5(4(L_4 - 2L_6) + 5L_5 - 8L_8) \]
\[ - 2L_1 - L_2 - \frac{5}{18}L_3 - \frac{4}{3}L_5 + 8(2L_7 + L_8). \] (3.14)

There were several nonunique choices that went into the determination of Eq. (3.3). When working with the full N\(^2\)LO \( \chiPT \) 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 \( F_K/F_\pi \)

\[ \frac{F_K}{F_\pi} = 1 + \delta F_{KLO}^{NLO} + \delta F_{KLO}^{N^{2} LO} + \cdots \]
\[ = 1 + \delta F_{KLO}^{NLO} + \delta F_{KLO}^{N^{2} LO} \]
\[ + (\delta F_{KLO}^{NLO})^2 - \delta F_{KLO}^{N^{2} LO} \delta F_{KLO}^{N^{3} LO} + \cdots, \] (3.15)

where the \( \cdots \) represent higher-order terms in the expansion and \( \delta F_{KLO}^{N^{3} LO} = \delta F_{KLO}^{N^{2} LO} - \delta F_{KLO}^{NLO} \). Eq. (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 \( \eta \) mass through the GMO relation. The NLO expressions are the standard ones [68]

\[ \delta F_{KLO}^{NLO} = -\frac{3}{8} \ell_\pi - \frac{3}{4} \ell_K - \frac{3}{8} \ell_\eta + 4L_5\epsilon^2_K \]
\[ + 4L_4(\epsilon^2_K + 2\epsilon^2_K), \]
\[ \delta F_{KLO}^{N^{2} LO} = -\ell_\pi - \frac{1}{4} \ell_K + 4L_5\epsilon^2_K + 4L_4(\epsilon^2_K + 2\epsilon^2_K), \]
\[ \delta F_{KLO}^{N^{3} LO} = 5\ell_\pi - \frac{3}{4} \ell_K - \frac{3}{8} \ell_\eta + 4L_5(\epsilon^2_K - \epsilon^2_\pi). \] (3.16)

The \( \delta F_{KLO}^{N^{3} LO} \) 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 \( (\delta F_{KLO}^{NLO})^2 - \delta F_{KLO}^{N^{2} LO} \delta F_{KLO}^{N^{3} LO} \) 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 \( \delta F_{KLO}^{N^{3} LO} \) 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\(^3\)LO is

\[ \frac{F_K}{F_\pi} [(3.3)] = \frac{1 + \delta F_{KLO}^{NLO}}{1 + \delta F_{KLO}^{N^{2} LO}} + \delta F_{KLO}^{N^{2} LO} + \delta F_{KLO}^{N^{3} LO} + \cdots, \] (3.17)

where \( \delta F_{KLO}^{N^{2} LO} \) is the full N\(^2\)LO expression in Eq. (3.15)

\[ \delta F_{KLO}^{N^{2} LO} = \delta F_{KLO}^{N^{2} LO} + (\delta F_{KLO}^{N^{2} LO})^2 - \delta F_{KLO}^{N^{2} LO} \delta F_{KLO}^{N^{3} LO}, \] (3.18)

and the ratio correction is given by

\[ \frac{\delta F_{KLO}^{N^{2} LO}}{\delta F_{KLO}^{N^{3} LO}} = \delta F_{KLO}^{N^{2} LO} - (\delta F_{KLO}^{N^{2} LO})^2. \] (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 \( \chi PT \)'s \( F_0 \) and \( B_m \), which must be determined through extrapolation analysis). Equally valid, one could have chosen \( F \to F_K \) or \( F^2 \to F_K F_K \). 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

\[ \epsilon^2_F = \frac{m_F^2}{(4\pi F_F)^2} \]
\[ = \frac{m_F^2}{(4\pi F_F)^2} F_K F_\pi = \frac{m_F^2}{(4\pi F_F)^2} F_\pi \left(1 + \delta F_{KLO}^{N^{2} LO}\right), \]
\[ = \frac{m_F^2}{(4\pi F_F)^2} F_K F_\pi = \frac{m_F^2}{(4\pi F_F)^2} F_\pi \left(1 + 2\delta F_{KLO}^{N^{3} LO}\right), \] (3.20)

plus higher-order corrections.
Related to this choice, Eq. (3.3) is implicitly defined at the standard renormalization scale \[ \mu_0 = 4\pi F_0, \]
(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
\[
\frac{F_K}{F_\pi} = 1 + 5\frac{\mu_\pi}{8\pi} - \frac{1}{4} \ell_\mu^{\mu_\pi} - \frac{3}{8\pi} \ell_\mu^{\mu_\pi} + 4(\epsilon_\pi^2 - \epsilon_\pi^2)\tilde{L}_5(\mu_0) \\
= + \frac{5}{8\pi} \epsilon_\pi^2 \ln \left( \frac{\mu_\pi^2}{\mu_0^2} \right) - \frac{1}{4} \epsilon_\pi^2 \ln \left( \frac{\mu_\pi^2}{\mu_0^2} \right) \\
- \frac{3}{8\pi} \epsilon_\pi^2 \ln \left( \frac{\mu_\pi^2}{\mu_0^2} \right) + 4(\epsilon_\pi^2 - \epsilon_\pi^2)\tilde{L}_5(\mu_0) \\
= 1 + 5\frac{\mu_\pi}{8\pi} - \frac{1}{4} \ell_\mu^{\mu_\pi} - \frac{3}{8\pi} \ell_\mu^{\mu_\pi} + 4(\epsilon_\pi^2 - \epsilon_\pi^2)\tilde{L}_5(\mu_0) \\
+ \ln \left( \frac{\mu_\pi^2}{\mu_0^2} \right), \quad (3.23)
\]
where we have introduced the notation
\[
\ell_\mu^{\mu_\pi} = \epsilon_\pi^2 \ln \left( \frac{\mu_\pi^2}{\mu_0^2} \right). \quad (3.24)
\]
If we chose the renormalization scale \( \mu_\pi \) and add the second term of the last equality, then this expression is equivalent to working with the scale \( \mu_0 \) through N^2LO. The convenience of this choice becomes clear as \( \mu_\pi/\mu_0 \) has a familiar expansion
\[
\frac{\mu_\pi}{\mu_0} = 1 + \delta F_\pi^{\text{NLO}} + \ldots . \quad (3.25)
\]
Using the GMO relation Eq. (3.7) and expanding \( \ln(1+x) \) for small \( x \), this expression becomes
\[
\frac{F_K}{F_\pi} = 1 + 5\frac{\mu_\pi}{8\pi} - \frac{1}{4} \ell_\mu^{\mu_\pi} - \frac{3}{8\pi} \ell_\mu^{\mu_\pi} + 4(\epsilon_\pi^2 - \epsilon_\pi^2)\tilde{L}_5(\mu_0) \\
- \frac{3}{2} (\epsilon_\pi^2 - \epsilon_\pi^2)\delta F_\pi^{\text{NLO}} . \quad (3.26)
\]

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

If we temporarily expose the implicit dependence of the expression for \( \frac{F_K}{F_\pi} \) on the choices of \( F \) and \( \mu \), such that Eq. (3.3) is defined as
\[
\frac{F_K}{F_\pi}[(3.3)] = \frac{F_K}{F_\pi}(\mu_0, \mu_\pi), \quad (3.27)
\]
then the following expressions are all equivalent at N^2LO
\[
\frac{F_K}{F_\pi}(\mu_\pi, \mu_0) = \frac{F_K}{F_\pi}(\mu_\pi, \mu_K) + \delta F_\pi^{\text{NLO}} \\
= \frac{F_K}{F_\pi}(\mu_\pi, \mu_\pi) + \delta F_\pi^{\text{NLO}} , \quad (3.28)
\]
where
\[
\delta F_{\pi K} = - \frac{3}{2}(\epsilon_\pi^2 - \epsilon_\pi^2)\delta F_\pi^{\text{NLO}} \quad (3.29)
\]
and the LECs in these expressions are related to those at the standard scale by evolving them from \( \mu_0 \) to \( \mu_\pi \) with their known scale dependence \[ [68]. \] Implicit in these expressions is the normalization of the small parameters
\[
\epsilon_\pi^2 = \begin{cases} 
\frac{m_\pi^2}{(4\pi F_\pi)^2}, & \text{for } F \rightarrow F_\pi \\
\frac{m_\pi^2}{(4\pi F_K F_\pi)}, & \text{for } F \rightarrow \sqrt{F_\pi F_K} . \quad (3.30)
\end{cases}
\]
We have described several choices one can make in parameterizing the \( \chi \)PT formula for \( \frac{F_K}{F_\pi} \). The key point is that if the underlying chiral expansion is well behaved, the formulas resulting from each choice are all equivalent through N^2LO in the SU(3) chiral expansion, with differences only appearing at N^3LO 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
TABLE III. Extracted masses of the mixed MDWF-HISQ mesons. We use the notation from Ref. [71] in which $m_a$ and $m_b$ denote the masses of the valence pion and kaon and $j$ and $r$ denote the light and strange flavors of the sea quarks while $v$ and $n$ 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 $m_{\mu}^2 - m_i^2$ and other splittings provide an estimate of the additive mixed-meson mass splitting due to discretization effects, $a^2 \Delta_{\text{Mix}}^{\text{LO}}$ [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 sea-quark mass is increased.

| Ensemble | $a_{\mu,j}^2$ | $a_{\mu,n}^2$ | $a_{\mu,j}^2$ | $a_{\mu,n}^2$ | $a_{\mu,j}^2$ | $a_{\mu,n}^2$ | $a_{\mu,j}^2$ | $a_{\mu,n}^2$ | $a_{\mu,j}^2$ | $a_{\mu,n}^2$ |
|----------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| a1m400   | 0.359(17)      | 0.4586(24)     | 0.278(17)      | 0.3577(11)     | 0.5219(02)     | 0.0486(15)     | 0.0359(28)     | 0.0516(23)     | 0.0440(16)     | 0.112(14)     |
| a1m350   | 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)     |
| a1m310   | 0.3060(17)     | 0.4345(16)     | 0.4590(12)     | 0.5188(02)     | 0.0489(13)     | 0.0324(18)     | 0.0511(17)     | 0.0416(16)     | 0.112(14)     |
| a1m220   | 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)     |
| a1m1530  | 0.2323(12)     | 0.4058(56)     | 0.3337(84)     | 0.5560(31)     | 0.5257(02)     | 0.0559(75)     | 0.0187(59)     | 0.0489(94)     | 0.0423(45)     | 0.112(14)     |

$\lim_{\mu \to m_a} F_K = 1$, (3.33)

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 at any lattice spacing.
from Ref. [55], the NLO MA EFT expressions are

$$\delta F^\text{MA}_\pi = -\ell_{ju} - \frac{\ell_{ru}}{2} + 4\bar{L}_5\pi^2 + 4\bar{L}_4(c_\pi^2 + 2c_K^2) + c_a^2\bar{L}_a,$$

$$\delta F^\text{MA}_K = -\ell_{ju} + \ell_\pi + \frac{\ell_{ru} - \ell_{js} - \ell_{rs} - \ell_{ss}}{4} + \frac{3\ell_X}{8} + 4\bar{L}_5c_K^2 + 4\bar{L}_4(c_\pi^2 + 2c_K^2) + c_a^2\bar{L}_a$$

$$- \frac{\delta^2_{ju}}{8}(dL_\pi - 2K_{\pi X}) - \frac{\delta^4_{ju}}{24}K_{\pi X}^{(2,1)}$$

$$+ \frac{\delta^2_{rs}}{4}(K_{ssX} - 2(\ell_\pi + c_a^2K_{\pi X}^{(2,1)}))$$

$$+ \frac{\delta^2_{ju}\delta^2_{rs}}{12}(K_{ssX}^{(2,1)} - 2K_{ssX}). \quad (3.34)$$

In these expressions, we use the partially quenched flavor notation [71] in which

$$\pi: \text{valence-valence pion}$$

$$K: \text{valence-valence kaon}$$

$$u: \text{valence light flavor quark}$$

$$j: \text{sea light flavor quark}$$

$$s: \text{valence strange flavor quark}$$

$$r: \text{sea strange flavor quark}$$

$$X: \text{sea-sea eta meson}$$

and so, for example

$$\ell_{ju} = \frac{m^2_{ju}}{(4\pi F_\pi)^2} \ln \left( \frac{m^2_{ju}}{(4\pi F_\pi)^2} \right), \quad (3.35)$$

where $m_{ju}$ is the mass of a mixed valence-sea pion. The partial quenching parameters $\delta_{ju}^2$ and $\delta_{rs}^2$ 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 splitting in the quark masses plus a discretization correction arising from the taste-identity splitting

$$\delta_{ju} = 2B_0(m_j - m_u) + a^2\Delta_1$$

$$\delta_{rs} = 2B_0(m_s - m_s) + a^2\Delta_1. \quad (3.37)$$

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 $m_s = m_s$ within 1%-2%. The sea-sea eta mass in this tuning is given at LO in MA EFT as

$$m_X^2 = \frac{4}{3}m_K^2 - \frac{1}{3}m_\pi^2 + a^2\Delta_1. \quad (3.38)$$

These parameters, and the corresponding meson masses are provided in Tab. III. The expressions for $dL_\pi$, $K_{\phi_1\phi_2}$, $K_{\phi_1\phi_2\phi_3}$ and $K_{\phi_1\phi_2\phi_3}$ are provided in Appendix B.

At NLO in the MA EFT, the LECs which contribute to $\delta F_K$ and $\delta F_\pi$ 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 $\chi$PT and MA expressions. The well-known modification to the integral can be expressed as [75–77]

$$\ell_{jP}^{\mu,\text{FV}} = \ell_{jP}^{\mu} + 4e_P^2 \sum_{|n|\neq0} \frac{c_n}{m_PL|n|} K_1(m_PL|n|), \quad (3.39)$$

where the sum runs over all nonzero integer three-vectors. Each value of $|n|$ 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 $|n|$ from triplets of integers, see Tab. 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

$$\delta_{FV}\ell_{jP} \equiv \ell_{jP}^{\text{FV}} - \ell_{jP} = e_P^2 \sqrt{2\pi} \frac{e^{-m_PL}}{(m_PL)^{3/2}}$$

$$+ e_P^2 \times O\left( \frac{e^{-m_PL \sqrt{2}}}{(m_PL)^{5/2}}, \frac{e^{-m_PL}}{(m_PL)^{5/2}} \right). \quad (3.40)$$

The full finite volume corrections to the continuum formula are also known at N$^2$LO [78] as well as in the partially quenched $\chi$PT [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 Tab. II). MILC generated three volumes for this a12m220 ensemble series to study FV corrections. Fig. 2 shows a comparison of the results from the a12m220L, a12m220, and a12m220S along with the predicted volume corrections arising from NLO in $\chi$PT. The uncertainty band arises from an N$^2$LO fit using the full N$^2$LO continuum $\chi$PT 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.

| | 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 |

**TABLE IV.** Multiplicity factors for the finite volume corrections of the first 10 vector lengths, $|n|$. |
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

\[
\delta^{N^{3}LO} = (\epsilon^K - \epsilon^\pi) \left\{ \epsilon^4 a^4 A^6 + \epsilon^K (A^K s + A^K s + A^K s + c^K \epsilon^K) \right\} + A^K s + c^K \epsilon^K + (\epsilon^K - \epsilon^\pi) (A^K s + A^K s + A^K s + c^K \epsilon^K) \right\}. \tag{3.41}
\]

In principle, we could also add counterterms proportional to higher powers of $\alpha_S$ but with four lattice spacings, we would not be able to resolve the difference between the complete set of operators including all possible additional $\alpha_S$ 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 parameterize 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 | $N_{L_i}$ | $N_X$ | $N_a$ |
|-------|-----------|-------|-------|
| 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, $N_X$ the number of counterterms associated with the continuum $\chi$PT 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 $\mu_r = 770$ 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 $\mu_r$ to $\mu_{F_0}$:

\[
L'_i(\mu_2) = L_i(\mu_1) - \frac{\Gamma_i}{(4\pi)^2} \ln \left( \frac{\mu_2}{\mu_1} \right), \tag{4.1}
\]

with the values of $\Gamma_i$ listed in Table V for convenience. We use $F_0 = 80$ MeV, which is the value adopted by FLAG [8]. We set the central value of all the $L_i$ with this procedure and the widths are set as described in Tab. V.

Next, we must determine priors for the $N^2$LO and $N^3$LO local counterterm coefficients, $A^K_{\pi,\pi,s}$. 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 $\chi$PT 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 $A^K_{\pi,\pi,s}$ LECs equal to 2 results in good fits with near-optimal values. This provides evidence the normalization of small parameters we have chosen for $\epsilon^K$ and $\epsilon^K$, Eq. (3.1), is “natural” and supports the 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 $\chi$PT LECs vary from model to model with 5 a reasonable value for all.

As a result of this search, we pick as our priors

\[
\begin{align*}
A^K_{\pi,\pi,s} = 0 \pm 2, \quad A^K = 0 \pm 2, \\
A^K_{\pi,\pi,s} = 0 \pm 5, \quad A^K_{\pi,\pi,s} = 0 \pm 5. \tag{4.2}
\end{align*}
\]
TABLE V. $\Gamma_i$ 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 priori width in the minimization: If a value of $L_i$ is less than $0.5 \times 10^{-3}$, we assign it a 100% uncertainty at the scale $\mu = 770$ MeV; If the value is larger than $0.5 \times 10^{-3}$, we assign it the larger of 0.5 or 1/3 of the mean value.

| $L_i$ | $L_1$ | $L_2$ | $L_3$ | $L_4$ | $L_5$ | $L_6$ | $L_7$ | $L_8$ |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $10^3 L_i(m_\rho)$ | 3/32 | 3/16 | 0 | 1/8 | 3/8 | 111/44 | 0 | 5/48 |
| $10^3 L_i(m_0)$ | 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) |
| $10^2 L_i(0.5\mu)$ | 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) |

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

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

The values of $F_{\pi^+}$ 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)

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

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

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

C. Model averaging procedure

Our model average is performed under a Bayesian framework following the procedure described in [39, 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’ theorem, as

$$P(M_k|D) = \frac{P(D|M_k) P(M_k)}{\sum_i P(D|M_i) P(M_i)}. \tag{4.7}$$

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:

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

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

$$P(M_k|D) = \frac{\exp(\log \text{BF})_{M_k}}{\sum_{i=1}^K \exp(\log \text{BF})_{M_i}} \tag{4.9}$$

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 $Y$.

$$\mathbb{E}[Y] = \sum_k \mathbb{E}[Y|M_k] P(M_k|D) \tag{4.10}$$
$$\text{Var}[Y] = \left[ \sum_k \text{Var}[Y|M_k] P(M_k|D) \right] + \left[ \sum_k (\mathbb{E}[Y|M_k] P(M_k|D) - \mathbb{E}[Y|D]) \right]. \tag{4.11}$$

The variance $\text{Var}[Y]$ 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 $\chi$PT or MA EFT model consist of:

- $\times 2$: $\chi$PT or MA EFT at NLO
- $\times 3$: use $F_2 = \{F_2^2, F_2, F_K^2, F_K^2\}$ in defining $\epsilon_P^2$
- $\times 2$: fully expanded (3.15) or ratio (3.17) form
- $\times 2$: at N$^2$LO, use full $\chi$PT or just counterterms
- $\times 2$: include or not an $\alpha_S$ term at N$^2$LO
- $\times 2$: include or not the NLO FV corrections
- $\times 2$: include or not 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:

- $\times 2$: work to N$^2$LO or N$^3$LO
- $\times 3$: use $F_2 = \{F_2^2, F_2, F_K^2, F_K^2\}$ in defining $\epsilon_P^2$
- $\times 2$: include or not an $\alpha_S$ term at N$^2$LO
- $\times 2$: 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, $\chi^2$) 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 $\chi$PT form at NLO.

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

$$\frac{F_K}{F_\pi} = 1 + \bar{L}_S(\epsilon_R^2 - \epsilon_P^2)\{1 + t_{\text{FV}} \sum_{|n| \neq 0} \frac{c_n}{m_\pi L|n|}K_1(m_\pi L|n|)\} + \cdots \quad (4.12)$$

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 $\chi^2$ values). Cutting all of these variations reduces our final set of models to be N$^3$LO $\chi$PT with the following variations:

- $\times 3$: use $F_2 = \{F_2^2, F_2, F_K^2, F_K^2\}$ in defining $\epsilon_P^2$
- $\times 2$: fully expanded (3.15) or ratio (3.17) form
- $\times 2$: at N$^2$LO, use full $\chi$PT or just counterterms
- $\times 2$: include or not an $\alpha_S$ 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 Tab. 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 ($\chi$), 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:

\[ (c_K^2 - c_a^2) \times \{ a^2, \alpha S c_s^2, c_a^4 \}, \]

and thus, the statistical uncertainty of the results grows as well as the model variance, with a total uncertainty growth from \( \sim 0.0044 \) to \( \sim 0.0057 \), 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

\[ \frac{F_K}{F_\pi} = 1.1964(44) \rightarrow \frac{F_K}{F_\pi} \bigg|_{\text{no a06}} = 1.1941(57). \] (4.13)

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

- **Left:** all ensembles, \( N^3 \)LO \( \chi \)PT with only counterterms at \( N^2 \)LO and \( N^3 \)LO and \( F = F_\pi \);
- **Middle:** no a06m310L, \( N^3 \)LO \( \chi \)PT with only counterterms at \( N^2 \)LO and \( N^3 \)LO and \( F = F_\pi \);
- **Right:** no a06m310L, \( N^2 \)LO \( \chi \)PT 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 \( c_a^4 \) and \( c_s^2 \) contributions and are perfectly consistent with the result on the a06m310L ensemble. They are also consistent with an \( N^2 \)LO fit (no \( c_a^4 \) 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 \( F_K/F_\pi \) 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 \( N^3 \)LO and higher, it is interesting to study the convergence of the chiral expansion by studying the fits which use the full \( \chi \)PT expression at \( N^2 \)LO.

In Fig. 6, we show the resulting light quark mass dependence using the \( N^3 \)LO extrapolation with the full \( N^2 \)LO \( \chi \)PT formula. After the analysis is performed, the results from each ensemble are shifted to the physical kaon mass point, leaving only dependence upon \( c_\pi^2 \) and \( c_a^2 \) as well as dependence upon the \( \eta \) 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 \( c_\pi^2 \) 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 \( \sim 0.16 \) compared to 1 at \( \Lambda \), the \( N^2 \)LO contributions add another \( \sim 0.04 \), and the \( N^3 \)LO 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 \( N^3 \)LO in the noncounterterm contributions) are smaller than the uncertainty in our results. It is also interesting to note that choosing \( F_\pi, F_K \) is penalized by the analysis, indicating the numerical results prefer larger expansion parameters. In Tab. VI, we show the resulting \( \chi \)PT LECs determined in this analysis for the two choices \( F = \{ F_\pi, F_K \} \), as well as whether the ratio form of the fit is used, Eq. (3.17). For the Gasser-Leutwyler LECs, we evolve the values back from \( \mu_0 \rightarrow \mu_p \) 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
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 $F_K/F_\pi$. 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, $\exp\{\log\text{GBF}_i - \log\text{GBF}_{\text{max}}\}$. 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 N$_3$LO and N$_2$LO analysis can be compared which indicates a large preference for the N$_3$LO 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 N$_2$LO and N$_3$LO priors from those chosen in Eq. (4.2).

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 $\chi$PT 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 $\chi$PT 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 Tab. V, we find the contribution from the Gasser-Leutwyler LECs to these N$_2$LO counterterms,
PDFs from N$^2$LO ct or full XPT

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.

| LEC     | prior | $F^2 = F^2_{\pi \pi}$ ratio | $F^2 = F_{\pi \pi}F_K$ ratio |
|---------|-------|-------------------------------|-------------------------------|
| $\mu = 770$ |       |                               |                               |
| $10^4 L_1$ | 0.53(50) | 0.47(49) | 0.50(49) | 0.45(49) | 0.48(49) |
| $10^4 L_2$ | 0.81(50) | 0.77(46) | 0.84(46) | 0.69(44) | 0.77(45) |
| $10^4 L_3$ | -3.02(85) | -2.84(86) | -3.26(81) | -3.05(82) |
| $10^4 L_4$ | 0.30(30) | 0.24(29) | 0.14(29) | 0.24(29) | 0.16(29) |
| $10^4 L_5$ | 0.01(50) | 0.48(35) | 0.52(34) | 0.40(33) | 0.47(34) |
| $10^4 L_6$ | 0.14(14) | 0.14(14) | 0.14(14) | 0.14(14) | 0.14(14) |
| $10^4 L_7$ | -0.34(34) | -0.55(32) | -0.57(32) | -0.52(33) | -0.53(33) |
| $10^4 L_8$ | 0.47(47) | 0.30(46) | 0.28(46) | 0.35(46) | 0.32(46) |

As the $A_j^4$ terms are priored at $0(2)$, it is sufficient to rerun the analysis by simply setting $L_j^4 = 0$. 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 [10] and so it is not clear Eq. (3.14), are given by

$$L_4^4 = 0.3(1.3) , \quad L_\pi^4 = -0.64(94) . \quad (4.14)$$
if other groups observe the same preference for counter-
terms only at NNLO or not.

If the Taylor expansion fits (pure counterterm) were
good and favored over the $\chi$PT fits, this could be a
sign that the $SU(3)$ $\chi$PT formula was failing to de-
scribe the lattice results. However, we have to include
the NLO $\chi$PT expression, including its predicted (counter-
term free) volume dependence to describe the numer-
ical results. It would be nice to have the full $N^3$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 $N^3$LO MA EFT to better describe the results.
However, at two loops in $\chi$PT, the universality of MA
EFT expressions [58] breaks down such that the MA EFT
expression can no longer be “derived” from the corre-
sponding PQ$\chi$PT one (which is known for $F_K/F_\pi$ at
two loops [79, 86–88]). It is therefore unlikely that the NNLO
MA EFT expression specific to this MALQCD calcula-
tion 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 ob-
tain a direct determination of $F_{K+/F_\pi+}$ including strong
isospin breaking corrections, but excluding QED correct-
tions. This is the standard value quoted in the FLAG
reviews [8, 81]. Our calculations, like most, are per-
fomed 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 per-
turbative modification of the path integral and correla-
tion functions [90, 91]. In this work, we have not per-
fomed 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 $F_\pi$ including the
strong isospin breaking corrections are given by

$$
\delta F_{\pi^\pm}^{NLO} = \frac{\ell_{\pi^0} - \ell_{\pi^\pm}}{2} - \frac{\ell_{K^0} - \ell_{K^\pm}}{4} + 4\bar{L}_4 (\epsilon_{K^\pm}^2 + \epsilon_{K^\pm}^2) + 4\bar{L}_5 \epsilon_{K^\pm}^2 ,
$$

$$
\delta F_{K^\pm}^{NLO} = \frac{\ell_{\pi^0} - \ell_{\pi^\pm}}{8} - \frac{3\ell_{K^0} - \ell_{K^\pm}}{4} + \frac{1}{4} (\epsilon_{K^0}^2 - \epsilon_{K^\pm}^2) \frac{\ell_{\pi^0} - \ell_{\pi^\pm}}{\epsilon_{\pi^0} - \epsilon_{\pi^\pm}} + 4\bar{L}_4 (\epsilon_{\pi^\pm}^2 + \epsilon_{K^\pm}^2 + \epsilon_{K^0}^2) + 4\bar{L}_5 \epsilon_{K^\pm}^2 ,
$$

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 break-
ing 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 symmet-
ic 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 $\chi$PT
$B(m_d - m_u) = (M_{K_0}^2 - M_{K^\pm}^2)$.

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

$$
\delta F_{K^-\pi^-}^{iso} = \frac{F_{K+}^{iso}}{F_{\pi+}^{iso}} - \frac{F_{K}^{iso}}{F_{\pi}^{iso}} = -\frac{1}{4} (\ell_{K^+} - \ell_{K^-}) + 4\bar{L}_5 (\epsilon_{K^+}^2 - \epsilon_{K^-}^2)
$$

$$
+ \frac{1}{4} (\epsilon_{K^0}^2 - \epsilon_{K^\pm}^2) \frac{\ell_{\pi^0} - \ell_{\pi^\pm}}{\epsilon_{\pi^0} - \epsilon_{\pi^\pm}} .
$$

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

$$
\delta F_{K^-\pi^-}^{iso} = -\frac{1}{6} \frac{\epsilon_{K^0}^2 - \epsilon_{K^\pm}^2}{\epsilon_{\pi^0} - \epsilon_{\pi^\pm}} \left[ 4 \left( \frac{F_{K}^{iso}}{F_{\pi}^{iso}} - 1 \right) + \epsilon_{\pi^0}^2 \ln \left( \frac{\epsilon_{\pi^0}^2}{\epsilon_{\pi^\pm}^2} \right) - \epsilon_{K^-}^2 + \epsilon_{K^+}^2 \right] .
$$

In this expression, we have utilized the two relations

$$
\ell_{K^+} - \ell_{K^-} = -\frac{2}{3} \frac{\epsilon_{K^0}^2 - \epsilon_{K^\pm}^2}{\epsilon_{\pi^0}^2 - \epsilon_{\pi^\pm}^2} (\epsilon_{K^0}^2 - \epsilon_{K^\pm}^2) (\ln \epsilon_{K^0}^2 + 1)
$$

$$
\epsilon_{\pi^0}^2 - \epsilon_{\pi^\pm}^2 = \frac{4}{3} (\epsilon_{K^0}^2 - \epsilon_{K^\pm}^2) .
$$

At this order, both Eqs. (4.16) and (4.17) are equiva-
lent. 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 addi-
tional 25% uncertainty for $SU(3)$ truncation errors. In
Sec. IV D 2 we observe the $N^3$LO correction is $\sim 25\%$ of
the NLO correction (while NLO is $\sim 16\%$ of LO).

In order to evaluate these expressions, we have to def-
ine the physical point with strong isospin breaking and
without QED isospin breaking. We employ the values from FLAG[2017] \cite{FLAG} (except \(M_{\pi^0} = 134.6(3)\) MeV):
\[
\begin{align*}
\hat{M}_{\pi^0} &= \hat{M}_{K^+} = 134.8(3)\text{ MeV}, \\
\hat{M}_{K^0} &= 497.2(4)\text{ MeV}, \\
\hat{M}_{K^+} &= 491.2(5) \text{ MeV}.
\end{align*}
\]
With this definition of the physical point, we find (under the same model average as Tab. VII)
\[
\begin{align*}
\delta F^{\text{iso}}_{K - \pi} &= -0.00188(51), \\
\delta F^{\text{iso'}}_{K - \pi} &= -0.00215(24),
\end{align*}
\]
resulting in our estimated strong isospin breaking correction
\[
\frac{F_{K^+}}{F_{\pi^+}} - \frac{F_K}{F_\pi} = -0.00215(72),
\]
and our final result as reported in Eq. (1.4)
\[
\frac{F_{K^+}}{F_{\pi^+}} = 1.1942(44)(07)^{\text{iso}}
\]
\[
= 1.1942(5).
\]
where the first uncertainty in the first line is the combination of those in Eq. (1.3).

V. SUMMARY AND DISCUSSION

The ratio \(F_K/F_\pi\) 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 \cite{18}
\[
\frac{|V_{us}|}{|V_{ud}|} = 0.2760(4),
\]
where strong isospin breaking effects must be included for direct comparison with experimental data. Combining this expression with our final result, we find
\[
\frac{|V_{us}|}{|V_{ud}|} = 0.2311(10).
\]

Utilizing the current global average, \(|V_{ud}| = 0.97420(21)\), extracted from superallowed nuclear beta decays \cite{18} results in
\[
\frac{|V_{us}|}{|V_{ud}|} = 0.2251(10).
\]

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
\[
|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 0.99977(59).
\]

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 \(|V_{us}| = 0.2231(7)\) \cite{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
\[
|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 0.9812(95),
\]
leading to a roughly \(2\sigma\) tension with unitarity. Our result, along with 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.

Another motivation for this work was to precisely test (below 1%) whether the action we have used for our nucleon structure calculations \cite{20, 21} 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 \cite{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 \cite{93, 94}. While there has been speculation that this discrepancy is due to the continuum extrapolations \cite{94}, new work suggests the original work underestimated the systematic uncertainty in the correlation.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure9}
\caption{Result for the ratio of CKM matrix elements, \(|V_{us}|/|V_{ud}|\), extracted from the ratio \(F_K/F_\pi\) reported in this work (red band). The global lattice average for \(|V_{us}|\) extracted from a semileptonic decay form factor, \(f_+(0)\) \cite{8}, is shown as a horizontal blue band, while the global experimental average for \(|V_{ud}|\) from nuclear beta decay \cite{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.}
\end{figure}
function analysis, and when accounted for, the tension between our results goes away [95].

In either case, to obtain a subpercent determination of $g_A$, which is relevant for trying to shed light on the neutron lifetime discrepancy [96], 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 $F_K/F_\pi$ involves the same axial current that is relevant for the computation of the nucleon matrix element used to compute $g_A$.

Fig. 4 shows the continuum extrapolation of $F_K/F_\pi$ from our analysis. The size of the discretization effects are noticeably larger than we observed in our calculation of $g_A$ [39]. In Sec. IV D 1, we demonstrated that, while helpful, the a06m310L ensemble is not necessary to achieve a subpercent determination of $F_K/F_\pi$. 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 [92].

In Ref. [92], 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 $a \sim 0.06$ 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 $F_K/F_\pi$. We would like to thank J. Bijnens for helpful correspondence on $\chi$PT 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 [24, 25] 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 Award 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 Tab. VII. For example, the model

xpt-ratio_mmnlo_FV_alphaS_PP

indicates the model uses the continuum $\chi$PT 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 $\delta_{\text{ratio}}^{N^2\text{LO}}$. The finite volume corrections have been included at NLO. The discretization terms at N$^2$LO include the $\alpha_s\epsilon^2_a(\epsilon^2_K-\epsilon^2_\pi)$ 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 $\delta_{\pi}^{N^2\text{LO}}$, Eq. (3.29), to hold the actual renormalization scale fixed at $\mu_0 = 4\pi F_\pi$.

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

$$d\ell_\pi = \int_R \frac{d^dk}{(2\pi)^d} \frac{i}{(k^2-m^2_\pi)^2} = 1 + \ln(m^2_\pi/\mu^2) \frac{(4\pi)^2}{(4\pi)^2},$$  \hspace{1cm} (B1)
which has been regulated and renormalized with the standard $\chi$PT modified dimensional-regularization scheme [48].

The finite volume corrections to $\delta\ell_\varphi$ are given by

$$\delta_{\text{FV}}^{\ell_\varphi} = \sum_{n\neq 0} \frac{c_n}{(4\pi)^2} \left[ \frac{2K_1(mL|n|)}{mL|n|} - K_0(mL|n|) - K_2(mL|n|) \right]$$

The expression for $K_{\phi_1\phi_2}$ arises from the integral

$$K_{\phi_1\phi_2} = (4\pi)^2 \int_R \frac{d^d k}{(2\pi)^d} \frac{i}{(k^2 - m_{\phi_1}^2)(k^2 - m_{\phi_2}^2)} = \frac{\ell_{\phi_2} - \ell_{\phi_1}}{\epsilon_{\phi_2}^2 - \epsilon_{\phi_1}^2}. \quad (B3)$$

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

$$K^{(2,1)}_{\phi_1\phi_2} = \int_R \frac{d^d k}{(2\pi)^d} \frac{i(4\pi)^2(4\pi F^2)}{(k^2 - m_{\phi_1}^2)(k^2 - m_{\phi_2}^2)(k^2 - m_{\phi_3}^2)} = \frac{\ell_{\phi_3} - \ell_{\phi_2}}{(\epsilon_{\phi_3}^2 - \epsilon_{\phi_2}^2)(\epsilon_{\phi_3}^2 - \epsilon_{\phi_1}^2)} - \frac{\ell_{\phi_1}}{\epsilon_{\phi_1}^2 - \epsilon_{\phi_2}^2}. \quad (B4)$$

Finally, $K_{\phi_1\phi_2\phi_3}$ is given by

$$K_{\phi_1\phi_2\phi_3} = \int_R \frac{d^d k}{(2\pi)^d} \frac{i(4\pi)^2(4\pi F^2)}{(k^2 - m_{\phi_1}^2)(k^2 - m_{\phi_2}^2)(k^2 - m_{\phi_3}^2)} = \frac{\ell_{\phi_1}}{(\epsilon_{\phi_1}^2 - \epsilon_{\phi_2}^2)(\epsilon_{\phi_1}^2 - \epsilon_{\phi_3}^2)} + \frac{\ell_{\phi_2}}{(\epsilon_{\phi_2}^2 - \epsilon_{\phi_1}^2)(\epsilon_{\phi_2}^2 - \epsilon_{\phi_3}^2)} + \frac{\ell_{\phi_3}}{(\epsilon_{\phi_3}^2 - \epsilon_{\phi_1}^2)(\epsilon_{\phi_3}^2 - \epsilon_{\phi_2}^2)}. \quad (B5)$$

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

### Appendix C: HYBRID MONTE CARLO FOR NEW ENSEMBLES

We present various summary information for the three new ensembles used in this work, a06m310L, a15m135XL and a09m135. In Tab. VIII, we list the parameters of the HISQ ensembles used in the hybrid Monte Carlo (HMC).

| Model | $\chi^2$ | $Q$ | logGBF | weight | $F_K/F_\pi$ |
|-------|---------|-----|--------|--------|-------------|
| xpt-nnlo_FV.ct.PP | 0.847 | 0.645 | 77.728 | 0.273 | 1.1968(40) |
| xpt-nnlo_FV.alphaS.ct.PP | 0.843 | 0.650 | 77.551 | 0.229 | 1.1962(46) |
| xpt-nnlo_FV.ct.PP | 0.908 | 0.569 | 76.830 | 0.111 | 1.1974(40) |
| xpt-nnlo_FV.alphaS.ct.PP | 0.902 | 0.576 | 76.668 | 0.095 | 1.1966(40) |
| xpt-nnlo_FV.ct.PK | 1.014 | 0.439 | 76.343 | 0.068 | 1.1952(37) |
| xpt-nnlo_FV.alphaS.ct.PK | 1.100 | 0.344 | 74.421 | 0.010 | 1.1960(37) |
| xpt-nnlo_FV.ct.KK | 1.244 | 0.215 | 74.004 | 0.007 | 1.1912(39) |
| xpt-nnlo_FV.alphaS.PK | 1.150 | 0.295 | 73.780 | 0.005 | 1.1959(41) |
| xpt-nnlo_FV.KK | 1.288 | 0.184 | 72.757 | 0.002 | 1.1938(36) |
| xpt-nnlo_FV.alphaS.KK | 1.273 | 0.194 | 72.718 | 0.002 | 1.1930(40) |
| xpt-nnlo_FV.KK | 1.338 | 0.152 | 72.343 | 0.006 | 1.1952(37) |
| xpt-nnlo_FV.alphaS.KK | 1.322 | 0.162 | 72.323 | 0.001 | 1.1929(39) |
| xpt-nnlo_FV.KK | 1.536 | 0.068 | 71.459 | 0.001 | 1.1900(38) |
| xpt-nnlo_FV.KK | 1.558 | 0.061 | 71.430 | 0.001 | 1.1909(35) |
| Bayes Model Average | | | | | 1.1964(42) |

TABLE VII. List of models used in final result, as described in the text.
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 $N_{\text{stream}}$, and the total number of configurations $N_{\text{cfg}}$. For a given ensemble, each stream has an equal number of configurations. The gauge coupling, light, strange, and charm quark masses on each ensemble are given as well as the tadpole factor $u_0$ and the Naik-term added to the charm quark action $\epsilon_N$. Here $s$ denotes the total length in molecular 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 $\epsilon$ used in the HMC is provided as Len./$N_{\text{steps}}$ which was input with single precision. The average acceptance rate over all streams is listed as well as the number of streams.

| Ensemble     | $10/g^2$ | $a_{\text{m}}$ | $a_{\text{s}}$ | $a_{\text{c}}$ | $u_0$ | $\epsilon_N$ | $s$   | Len. | $\epsilon$ | Acc. | $N_{\text{stream}}$ | $N_{\text{cfg}}$ |
|--------------|----------|----------------|----------------|----------------|------|--------------|------|------|------------|------|-------------------|-----------------|
| 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            |

FIG. 11. The $\Delta S$ 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.

In Fig. 11, we show the MDTU history of the $\Delta S$ 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 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 $a_{\text{m}}, a_{\text{s}},$ and $a_{\text{c}}$. 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 \( \Gamma \)-method analysis [107] with the Python package \texttt{uneq} [108]. We report the results in Tab. 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 \( \langle \bar{\psi}\psi_s \rangle \) 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 \( \tau_s \) and consistent with the unblocked data. On this a15m135XL ensemble, while we have generated 2000 configurations, we have only utilized 1000 in this paper (the first half from each of the four streams).

Finally, in Tab. 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.

### Table IX. Average values of the quark-antiquark condensate \( \bar{\psi}\psi \) with statistical errors and integrated autocorrelation times \( \tau \) measured with the \( \Gamma \)-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     | \( N_{\text{stream}} \) | \( \bar{\psi}\psi_l \) | \( \tau_l \) | \( \bar{\psi}\psi_s \) | \( \tau_s \) | \( \bar{\psi}\psi_c \) | \( \tau_c \) |
|--------------|----------------|----------------|------------|----------------|------------|----------------|------------|
| 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)     |

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.
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, $\tau_b$ 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 $\tau_b$ indicates the autocorrelation time for these hadronic quantities is very short.

![FIG. 13](image)

TABLE X. Values of the overrelaxed stout smearing parameters used to measure the topological charge $Q$ and the resulting mean $\langle Q \rangle$ and width $\sigma$ of the distribution for each stream. The last column reports the integrated autocorrelation time in units of MDTU using the $\Gamma$-method analysis. These measurements were performed with QUDA which is now available via the su3.test test executable in the develop branch [24, 25].

| ensemble     | $\rho$ | $N_{\text{step}}$ | $Q_{\text{stream}}(\sigma)$ | $Q_{\text{all}}(\sigma)$ | $\tau_{\text{all}}(\sigma)$ |
|--------------|--------|-------------------|-------------------------------|--------------------------|-------------------------------|
| a15m135XL    | 0.068  | 2000              | a -10(34) b -3(35) c -2(33) d 5(32) e -3(33)                         | 15(3)                    |
| a09m135      | 0.065  | 2000              | 0.5(12.0) a 2(12) b -1(35) c -2(33) d 5(32) e -3(33)                   | 18(4)                    |
| a06m310L     | 0.066  | 1800              | 0.5(12.0) a 4(12) b -3(12.0) c -2(33) d 5(32) e -3(33)                 | 420(198)                 |

(aL = 72a ≃ 4.1 fm) that we do not anticipate any measurable impact from the poorly distributed $Q$-values, which nonetheless average to nearly 0.

![FIG. 14](image)

[1] William J. Marciano, “Precise determination of $\overline{V}(us)$ from lattice calculations of pseudoscalar decay constants,” Phys. Rev. Lett. 93, 231803 (2004).
[2] C. Aubin, C. Bernard, Carleton E. DeTar, J. Osborn, Steven Gottlieb, E. B. Gregory, D. Toussaint, U. M. Heller, J. E. Hetrick, and R. Sugar (MILC), “Light pseudoscalar decay constants, quark masses, and low energy constants from three-flavor lattice QCD,” Phys. Rev. D70, 114501 (2004), arXiv:hep-lat/0407028 [hep-lat].

[3] Roger Decker and Markus Finkemeier, “Short and long distance effects in the decay tau → pi tau-neutrino (gamma),” Nucl. Phys. B438, 17–53 (1995), arXiv:hep-ph/9403385 [hep-ph].

[4] Markus Finkemeier, “Radiative corrections to pi(12) and K(12) decays,” 2nd Workshop on Physics and Detectors for DAPHNE (DAPHNE 95) Frascati, Italy, April 4-7, 1995, Phys. Lett. B387, 391–394 (1996), arXiv:hep-ph/9505434 [hep-ph].

[5] Vincenzo Cirigliano and Helmut Neufeld, “A note on isospin violation in Pi2(gamma) decays,” Phys. Lett. B700, 7–10 (2011), arXiv:1102.0563 [hep-ph].

[6] M. Tanabashi et al. (Particle Data Group), “Review of Particle Physics,” Phys. Rev. D98, 070001 (2018).

[7] C. T. H. Davies et al. (HPQCD, UKQCD, MILC, Fermilab Lattice), “High precision lattice QCD confronts experiment,” Phys. Rev. Lett. 92, 022001 (2004), arXiv:hep-lat/0304004 [hep-lat].

[8] S. Aoki et al. (Flavour Lattice Averaging Group), “FLAG Review 2019,” (2019), arXiv:1902.08191 [hep-lat].

[9] N. Carrasco et al., “Leptonic decay constants fK, fD, and fD with Nf = 2+1+1 twisted-mass lattice QCD,” Phys. Rev. D91, 054507 (2015), arXiv:1411.7908 [hep-lat].

[10] R. Frezzotti and G. C. Rossi, “Twisted mass lattice QCD with nondegenerate quarks,” Lattice hadron physics. Proceedings, 2nd Topical Workshop, LHPC 2003, Cairns, Australia, July 22-30, 2003, Nucl. Phys. Proc. Suppl. 128, 193–202 (2004), [193(2003)], arXiv:hep-lat/0311008 [hep-lat].

[11] R. Frezzotti and G. C. Rossi, “Chirally improving Wilson fermions. 1. O(a) improvement,” JHEP 08, 007 (2004), arXiv:hep-lat/0306014 [hep-lat].

[12] R. J. Dowdall, C. T. H. Davies, G. P. Lepage, and C. McNeile, “Vus from pi and K decay constants in full lattice QCD with physical u, d, s and c quarks,” Phys. Rev. D88, 074504 (2013), arXiv:1303.1670 [hep-lat].

[13] A. Bazavov et al., “B- and D-meson leptonic decay constants from four-flavor lattice QCD,” Phys. Rev. D98, 074512 (2018), arXiv:1712.09262 [hep-lat].

[14] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu, H. Trottier, and K. Wong (HPQCD, UKQCD), “Highly improved staggered quarks on the lattice, with applications to charm physics,” Phys. Rev. D75, 054502 (2007), arXiv:hep-lat/0610092 [hep-lat].

[15] E. Follana, C.T.H. Davies, G.P. Lepage, and J. Shigemitsu (HPQCD, UKQCD), “High Precision determination of the pi, K, D and D(s) decay constants from lattice QCD,” Phys. Rev. Lett. 100, 062002 (2008), arXiv:0706.1726 [hep-lat].

[16] A. Bazavov et al. (MILC), “Results for light pseudoscalar mesons,” PoS LATTICE2010, 074 (2010), arXiv:1012.0868 [hep-lat].
Richard C. Brower, Harmut Neff, and Kostas Orginos, “The moebius domain wall fermion algorithm,” (2012), arXiv:1206.5214 [hep-lat].

R. Narayanan and H. Neuberger, “Infinite N phase transitions in continuum Wilson loop operators,” JHEP 0603, 064 (2006), arXiv:hep-th/0601210 [hep-th].

M. Lüscher and P. Weisz, “Perturbative analysis of the gradient flow in non-abelian gauge theories,” JHEP 1102, 051 (2011), arXiv:1101.0963 [hep-th].

Martin Lüscher, “Chiral symmetry and the Yang–Mills gradient flow,” JHEP 1304, 123 (2013), arXiv:1302.5246 [hep-lat].

Martin Lüscher, “Properties and uses of the Wilson flow in lattice QCD,” JHEP 1008, 071 (2010), arXiv:1006.4518 [hep-lat].

Robert Lohmayer and Herbert Neuberger, “Continuous smearing of Wilson Loops,” PoS LATTICE2011, 249 (2011), arXiv:1110.3522 [hep-lat].

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), arXiv:1805.02634 [nucl-th].

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,” (2017), arXiv:1704.01114 [hep-lat].

C. C. Chang, A. Nicholson, 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 percent-level determination of the nucleon axial coupling from quantum chromodynamics,” Nature 558, 91–94 (2018), arXiv:1805.12130 [hep-lat].

Evan Berkowitz, M. A. Clark, A. Gambhir, K. McElvain, A. Nicholson, E. Rinaldi, P. Vranas, A. Walker-Loud, C. C. Chang, B. Joó, T. Kurth, and K. Orginos, “Simulating the weak death of the neutron in a femtoscopic universe with near-Exascale computing,” (2018), arXiv:1810.01609 [hep-lat].

Oliver Bar, Gautam Rupak, and Noam Shoresh, “Simulations with different lattice Dirac operators for valence and sea quarks,” Phys. Rev. D67, 114505 (2003), arXiv:hep-lat/0210050 [hep-lat].

A. Bazavov et al. (MILC), “Scaling studies of QCD with the dynamical HISQ action,” Phys. Rev. D82, 074501 (2010), arXiv:1004.0342 [hep-lat].

A. Bazavov et al. (MILC), “Lattice QCD ensembles with four flavors of highly improved staggered quarks,” Phys. Rev. D87, 054505 (2013), arXiv:1212.4768 [hep-lat].

David B. Kaplan, “A Method for simulating chiral fermions on the lattice,” Phys. Lett. B288, 342–347 (1992), arXiv:hep-lat/9206013 [hep-lat].

Yigal Shamir, “Chiral fermions from lattice boundaries,” Nucl. Phys. B406, 90–106 (1993), arXiv:hep-lat/9303005 [hep-lat].

Vadim Furman and Yigal Shamir, “Axial symmetries in lattice QCD with Kaplan fermions,” Nucl. Phys. B439, 54–78 (1995), arXiv:hep-lat/9405004 [hep-lat].

Paul Langacker and Heinz Pagels, “Chiral perturbation theory,” Phys. Rev. D8, 4595–4619 (1973).

J. Gasser and H. Leutwyler, “Chiral Perturbation Theory to One Loop,” Annals Phys. 158, 142 (1984).

H. Leutwyler, “On the foundations of chiral perturbation theory,” Annals Phys. 235, 165–203 (1994), arXiv:hep-th/9311274 [hep-ph].

Stephen R. Sharpe and Robert L. Singleton, Jr, “Spontaneous flavor and parity breaking with Wilson fermions,” Phys. Rev. D58, 074501 (1998), arXiv:hep-lat/9804028 [hep-lat].

Oliver Bar, Gautam Rupak, and Noam Shoresh, “Chiral perturbation theory at Ω(a²) for lattice QCD,” Phys. Rev. D70, 034508 (2004), arXiv:hep-lat/0306021 [hep-lat].

Oliver Bar, Claude Bernard, Gautam Rupak, and Noam Shoresh, “Chiral perturbation theory for staggered sea quarks and Ginsparg-Wilson valence quarks,” Phys. Rev. D72, 054502 (2005), arXiv:hep-lat/0503009 [hep-lat].

Brian C. Tiburzi, “Baryons with Ginsparg-Wilson quarks in a staggered sea,” Phys. Rev. D72, 094501 (2005). [Erratum: Phys. Rev.D79,039904(2009)], arXiv:hep-lat/0508019 [hep-lat].

Jiunn-Wei Chen, Donal O’Connell, Ruth S. Van de Water, and Andre Walker-Loud, “Ginsparg-Wilson pions scattering on a staggered sea,” Phys. Rev. D73, 074510 (2006), arXiv:hep-lat/0510024 [hep-lat].

Jiunn-Wei Chen, Donal O’Connell, and Andre Walker-Loud, “Two Meson Systems with Ginsparg-Wilson Valence Quarks,” Phys. Rev. D75, 054501 (2007), arXiv:hep-lat/0611003 [hep-lat].

Kostas Orginos and Andre Walker-Loud, “Mixed meson masses with domain-wall valence and staggered sea fermions,” Phys. Rev. D77, 094505 (2008), arXiv:0705.0572 [hep-lat].

Fu-Jiun Jiang, “Mixed Action Lattice Spacing Effects on the Nucleon Axial Charge,” (2007), arXiv:hep-lat/0703012 [hep-lat].

Jiunn-Wei Chen, Donal O’Connell, and Andre Walker-Loud, “Universality of mixed action extrapolation formulae,” JHEP 04, 090 (2009), arXiv:0706.0035 [hep-lat].

Jiunn-Wei Chen, Maarten Golterman, Donal O’Connell, and Andre Walker-Loud, “Mixed Action Effective Field Theory: An Addendum,” Phys. Rev. D79, 117502 (2009), arXiv:0905.2566 [hep-lat].

T. Blum et al., “Quenched lattice QCD with domain wall fermions and the chiral limit,” Phys. Rev. D69, 074502 (2004), arXiv:hep-lat/0007038 [hep-lat].

Y. Aoki et al., “Domain wall fermions with improved gauge actions,” Phys. Rev. D69, 074504 (2004), arXiv:hep-lat/0211023 [hep-lat].

Sergey Syritsyn and John W. Negele, “Oscillatory terms in the domain wall transfer matrix,” Proceedings, 25th International Symposium on Lattice field theory (Lattice 2007): Regensburg, Germany, July 30-August 4, 2007), International Symposium on Lattice field theory (Lattice 2007), 39–50 (2008), arXiv:0710.0425 [hep-lat].

Szabolcs Borsanyi et al., “High-precision scale setting in lattice QCD,” JHEP 09, 010 (2012), arXiv:1203.4469 [hep-lat].

Gabriel Amoros, Johan Bijnen, and P. Talavera, “Two point functions at two loops in three flavor chiral per-
[98] Peter 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] Robert G. Edwards and Balint Joo (SciDAC Collaboration, LHPC Collaboration, UKQCD Collaboration), “The Chroma software system for lattice QCD,” Nucl.Phys.Proc.Suppl. 140, 832 (2005), arXiv:hep-lat/0409003 [hep-lat].

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

[102] Thorsten Kurth, Andrew Pochinsky, Abhinav Sarje, Sergey Syritsyn, and André Walker-Loud, “High-Performance I/O: HDF5 for Lattice QCD,” Proceedings, 32nd International Symposium on Lattice Field Theory (Lattice 2014): Brookhaven, NY, USA, June 23-28, 2014, PoS LATTICE2014, 045 (2015), arXiv:1501.06992 [hep-lat].

[103] Evan Berkowitz, “METAQ: Bundle Supercomputing Tasks,” (2017), https://github.com/evanberkowitz/metaq, arXiv:1702.06122 [physics.comp-ph].

[104] Evan Berkowitz, Gustav R. Jansen, Kenneth McElvain, and André Walker-Loud, “Job Management and Task Bundling,” EPJ Web Conf. 175, 09007 (2018), arXiv:1710.01986 [hep-lat].

[105] Chia Cheng Chang, Christopher Körber, and André Walker-Loud, “EspressoDB: A Scientific Database for Managing High-Performance Computing Workflow,” J. Open Source Softw. 5, 2007 (2020), arXiv:1912.03580 [hep-lat].

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

[107] Ulli Wolff (ALPHA), “Monte Carlo errors with less errors,” Comput. Phys. Commun. 156, 143–153 (2004), [Erratum: Comput.Phys.Commun. 176, 383 (2007)], arXiv:hep-lat/0306017.

[108] Barbara De Palma, Marco Erba, Luca Mantovani, and Nicola Mosco, “A Python program for the implementation of the Γ -method for Monte Carlo simulations,” Comput. Phys. Commun. 234, 294–301 (2019), arXiv:1703.02766 [hep-lat].

[109] Peter J. Moran and Derek B. Leinweber, “Over-improved stout-link smearing,” Phys. Rev. D 77, 094501 (2008), arXiv:0801.1165 [hep-lat].